DOI 10.1002/mnfr.200500033 Mol. Nutr. Food Res. 2005, *49*, 908 – 917

Review

Impact of microarray technology in nutrition and food research

Bettina Spielbauer[1] and Frank Stahl[2]

[1]Neuro and Sensory Physiology, University of Göttingen, Göttingen, Germany
[2]Institute for Technical Chemistry, University of Hannover, Hannover, Germany

Microarrays have become standard tools for gene expression profiling as the mRNA levels of a large number of genes can be measured in a single assay. Many technical aspects concerning microarray production and laboratory usage have been addressed in great detail, but it remains still crucial to establish this technology in new research fields such as human nutrition and food-related areas. The correlation between diet and inter-individual variation in gene expression is an important and relatively unexplored issue in human nutrition. Therefore, nutritionists changed their research field dramatically from epidemiology and physiology towards the 'omics' sciences. Nutrigenomics as a field of research is based on the complete knowledge of the human genome and refers to the entire spectrum of human genes that determine the interactions of nutrition with the organism. Nutrigenetics is based on the inter-individual, genetically determined differences in metabolism. Nutrigenomics and nutrigenetics carry the hope that individualized diet can improve human health and prevent nutrition-related diseases. In this article we give an overview of current DNA and protein microarray techniques (including fabrication, experimental procedure and data analysis), we describe their applications to nutrition and food research and point out the limitations, problems and pitfalls of microarray experiments.

Keywords: Microarray / Gene expression / Biosensor / Functional food / Review

Received: March 7, 2005; revised: July 14, 2005; accepted: July 17, 2005

1 Introduction

In the 'post-genome era' – as the Human Genome Project (HGP) has come to completion – whole-genome studies are becoming more and more important as the number of DNA sequences available to researchers has been rising exponentially [1–5]. For the understanding of biological systems with up to 30 000 genes, the measurement of mRNA levels for a complete set of transcripts of an organism will be necessary. An ideal tool for such measurements is the DNA microarray or so-called DNA chip technology, a high-throughput method that has become a standard tool for gene expression profiling as the mRNA levels of a large number of genes can be measured in a single assay. This is the great advantage over traditional methods for gene expression analysis such as RT-PCR, Northern blotting or more advanced techniques such as differential display, which have also enabled the discovery of novel differentially expressed genes. However, only few samples can be analyzed simultaneously using these well-established methods. While conventional methods are confining themselves to the examination of single genes, a DNA chip experiment delivers a complete gene expression pattern of the cell [6–10]. In the nutritional sciences, whole-genome gene expression profiles could provide new insights into nutrient-gene interactions and the diet-related mechanisms underlying alterations in gene expression and functional states of the cell.

Current hybridization arrays can be categorized into macroarrays, microarrays, and microelectronic arrays. Macroarrays are fabricated by spotting probes on a membrane-based matrix [11, 12]. The term macroarray refers to the lower spot density (spot diameter >300 μm) on these arrays. Currently, DNA clones, PCR products or oli-

Correspondence: Dr. Frank Stahl, Institute for Technical Chemistry, University of Hannover, Callinstr. 3, 30167 Hannover, Germany
E-mail: stahl@iftc.uni-hannover.de
Fax: +49-511-7623004

Abbreviations: ANOVA, analysis of variance; **BSE**, bovine spongiform encephalopathy or mad cow disease; **Cy3**, Cyanine-3; **Cy5**, Cyanine-5; **GEO**, Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo); **GMO**, genetically modified organism; **HGP**, Human Genome Project; **SNP**, single nucleotide polymorphism; **WHO**, World Health Organization

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

gonucleotides are printed onto membranes as probes. Targets for these filter arrays are usually radioactive labeled. Different samples are hybridized to several arrays and the bound target can be detected by phosphoimagers. Microarrays use glass slides as medium that enables high spot density (spot diameter of <250 μm) [13]. Either PCR products or pre-synthesized oligonucleotides are spotted robotically, or short oligonucleotides are synthesized directly on the glass surface. Microelectronic arrays are made up of sets of microelectrodes covered by a thin layer of agarose coupled with an affinity moiety. The generated electric fields are used to immobilize probes and control target hybridization. This technology is still under development.

Predominantly, DNA microarrays are currently applied to analyze RNA expression levels, but DNA arrays can also be used for gene discovery, sequence identification, mapping, and polymorphism detection. Therefore, microarrays contribute to a better understanding of complex issues concerning human health and nutrition. They can tease out the effects of environment and life style, including drugs and nutrition, and help usher in the individualized molecular medicine of the future. For example, the ideal dose for all nutritional supplements is unknown. Microarray experiments could help to define and quantify the impact of these supplements on human health. Large response variabilities among individuals due to genetic and environmental differences require the usage of methods such as microarray technology, leading to custom-made nutritional intake and therapeutic drug doses for the constitution of each individual (individualized and functional food). The effect of nutrition on gene expression is defined as nutrigenomics. In contrast, nutrigenetics addresses the effects of individual genetic variation of response to diet at the level of single nucleotide polymorphisms (SNPs) rather than at the gene level. Nutrigenetic studies may lead to individualized dietary suggestions. Studies on the biological function of genes will help to identify how diet affects gene and protein functions, and why individuals vary in their response to nutrients and diets. When the genetic variation of human genomes among populations is known, a correlation between nutrition and genetic variation might be possible. For these purposes the identification and isolation of mutants that are unable to respond, or that respond abnormally, may provide ways to decipher the mechanisms by which a nutritional signal is transduced into a given response. In the 1980s markers like protein and blood group loci were replaced with DNA polymorphisms. The advantages of SNPs are their higher frequency and higher mutational stability compared to microsatellites. Hence, SNPs are more amenable to automation and DNA chip technology. By the rapid development of the microarray technology within the last few years, this technology is suitable in the medical diagnostics as well as in the fundamental research for routine use. Nevertheless, it remains still crucial to establish this technology in new research fields. Although the real experiment concerning array production and laboratory usage is largely standardized, the array systems have to be optimized for the given experimental question. Further, there are various pitfalls within data analysis and interpretation of results.

2 Fundamental aspects of array technology

2.1 Classification of DNA hybridization arrays

All hybridization arrays are based on the same steps: RNA isolation and labeling with fluorophore or radiolabel, production or purchasing of arrays, hybridization, signal detection and data analysis. The nomenclature of probes and targets that is commonly used reflects the similarity between array and reverse dot-blot technologies. The DNA with a defined identity tethered to a solid medium is referred to as 'probe' and the labeled DNA as 'target' [14].

2.1.1 Fabrication of DNA microarrays

Basically, DNA microarrays can be classified depending on diverse criteria such as quantity of DNA probes (high- and low-density microarrays), features of the attached DNA probes (PCR- or oligonucleotide-based microarrays), or procedure of array manufacturing. Presently, there are three primary technologies used in automated microarray fabrication: photolithography, ink-jet printing, and contact printing, along with derivatives of each. Each of these technologies has specific advantages and disadvantages in microarray manufacturing. The photolithographic approach relies on the *in situ* synthesis of 25-mer oligonucleotides using photomasks. This means that each probe is individually synthesized on the chip surface. Photolitography was developed by Fodor *et al.* [15] and commercialized by Affymetrix. In contrast, the ink-jet and contact printing methods attach pre-synthesized DNA probes to the chip surface. While the *in situ* probe synthesis requires sophisticated and expensive equipment, the contact and non-contact dispensing methods made DNA chips affordable for academic research laboratories. The *in situ* synthesis on the chip is less exact and the products can not be analyzed carefully. In contrast, pre-made oligonucleotides can be fabricated in high quality and defined quantities. Another possibility is the attachment of PCR products. The PCR-based DNA amplification from bacterial clone libraries require time-consuming establishment of appropriate methods and costly high-throughput robotic systems, but can be reasonable, *e.g.*, for the analysis of self-made clone libraries or fabrication of large amounts of arrays.

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Since 1996, many commercial DNA arrayers have become available and the self-spotted glass slide DNA arrays are currently the most popular format for gene expression profiling experiments [16–20]. For the production of self-made DNA microarrays, oligonucleotides or PCR-products were arrayed on silylated microscope slides using high-speed robotics. The glass slides are coated with functional groups enabling the chemical reaction between the spotted DNA targets and the glass surface. After the DNA fixation, free functional groups of the coated slides and side products are reduced, and any salts and impurities are removed. The microarrays can be stored for several months.

2.1.2 Synthesis of labeled target molecules

DNA microarrays are usually hybridized with fluorescently labeled cDNA targets using the co-hybridization technique. This means, one sample is labeled with dye 1, usually Cy3, the other sample with dye 2, usually Cy5, and both labeled cDNAs are co-hybridized on the same DNA chip. One advantage of such two-component system is its independence of the absolute amount of the fixed DNA, as only the relative ratios of the Cy3- *vs*. Cy5-signal intensities are analyzed for each spot separately. Contrarily, single-color systems introduce more variance as compared to two-component systems, and therefore increases the number of replicates required in an experiment.

Besides the array system used, the limiting amount of starting material determines the type of the labeling protocol. For large quantities of total RNA (*e.g.*, preparations from liver tissue or cell culture), direct labeling methods are usually used. For this, fluorescently labeled dCTPs or dUTPs are incorporated during the reverse transcription of RNA. Another possibility is the incorporation of amino-allyl-modified nucleotides during the first-strand cDNA reaction. In a subsequent step, the dyes are chemically coupled to the reactive amino groups in the modified cDNA. Theoretically, such indirect labeling protocol avoid different efficiencies of the incorporation of fluorophores during cDNA synthesis.

If only small amounts of total RNA or mRNA are available, the starting material has to be amplified, whereby alterations in the original expression levels have to be avoided. Appropriate methods amplify either the RNA or the resulting signal intensities. Most RNA amplification protocols combine cDNA synthesis with a template-directed *in vitro* transcription reaction. In the T7 RNA polymerase amplification method, first described by Eberwine *et al*. [21], a T7 polymerase promoter sequence is incorporated into first-strand cDNA. The second strand of cDNA, which serves as the template for the RNA polymerase, can be generated by conventional second-strand synthesis procedure [22], or *e.g.*, by a template switch mechanism at the 5′ end of the RNA (so-called SMART technology) [23], followed by PCR-based amplification by *in vitro* transcription. The resulting antisense-RNA is then used as template for the synthesis of fluorescently labeled cDNA. The fluorescence signal amplification originates from either a chemical enhancement of fluorescence or an enzymatically based signal amplification. Some protocols are, for example, based on hapten-antibody enzymatic labeling or 3-D multi-labeled structures, while other use tyramide signal amplification.

2.1.3 Array processing and scanning

The hybridization scheme depends on the given experimental question. For this, the nature of the starting material has to be considered, and consequently also the number of biological and technical replicates. To draw conclusions from the investigation of a single measured effect on the entire target population of interest, the biological variability from animal to animal has to be taken into account. This means, *e.g.*, for the investigation of a nutritional supplement in an animal model the measurement of several different animals (biological replicates) is needed, whereas the determination of the same effect in HepG2 cells requires only technical replicates. Experimental standard designs are the so-called dye swap design, where one hybridization is repeated with a reverse labeling and, therefore, the dye effect can be minimized, and the so-called common reference design, where for every hybridization always the same reference is labeled with one dye and the samples of interest (*e.g.*, different patients, cell lines or different time points) are labeled with the second dye [24, 25]. The microarrays are hybridized and washed according to the slide manufacturer's recommendations. For signal detection the arrays are scanned in a microarray scanner avoiding signal saturation.

2.1.4 Data processing and statistical analysis

The first step in data analysis is the quantification of signal and background intensities. For the acquisition of these raw data, commercial or free available software packages can be used (open source packages: see [26, 27]). Dye effects – Cy3 produce for example higher signal intensities than Cy5 – and array-to-array variations of the total signal intensities can be corrected by normalization procedures (*e.g.*, LOWESS regression-based normalization [28]).

For the selection of the relevant differentially expressed gene candidates from the huge quantity of raw data, especially from studies with high-density arrays, the statistical analysis using multiple testing procedures is essential [29–41]. Otherwise, solely the identification of relevant data by fold change rather produces lists of arbitrary genes than true results. Analysis of variance (ANOVA) models have been proposed and successfully applied to the analysis of

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

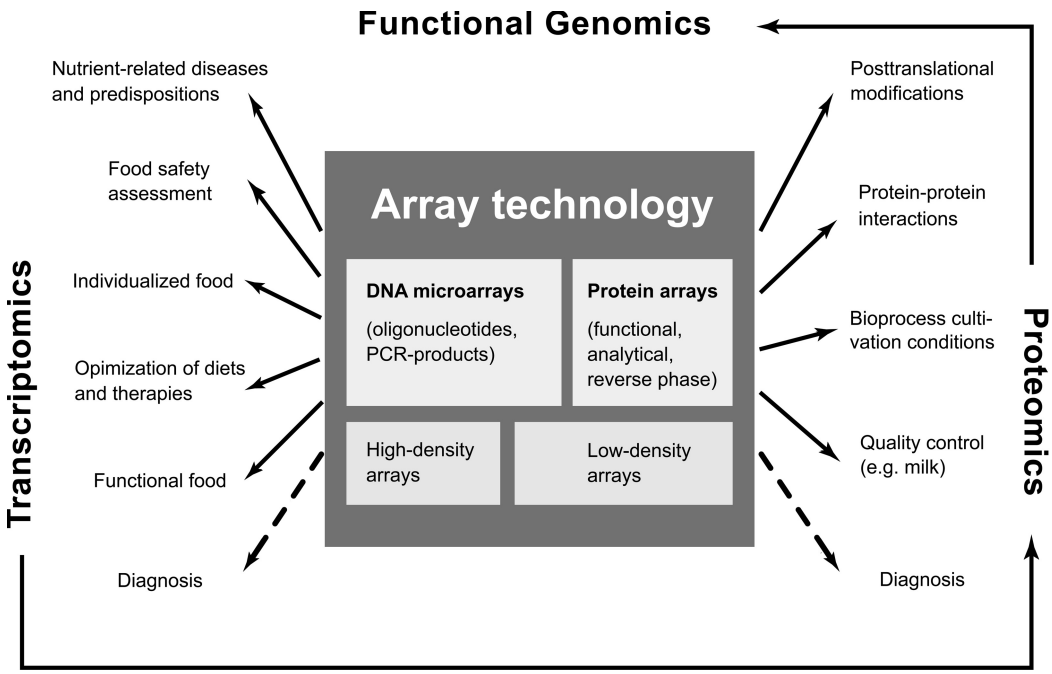

Figure 1. Schematic overview of potential applications of biochips in nutrition and food research. The array technology plays a key role in the scientific analysis of the interactions between nutrients and genes. Currently, the influence of nutrition on the global gene expression profile can be analyzed by using high-density, whole-genome DNA microarrays. For the characterization of specific effects of single organs or cell systems, usually low-density, custom-made arrays are used. Future prospectives focus on the application of standardized low-density arrays in the medical diagnostics or quality control of food products.

microarray experiments. The ANOVA approach makes no assumptions about whether there are a lot, or a few differentially expressed genes, since it tests each gene independently of the others. A related advantage of the ANOVA approach for two-color arrays is that it does not assume the same dye effect for each gene (or each gene at a particular expression level), like alternative methods do. Generally, ANOVA models can be divided into fixed, random and mixed effects models. For the analysis of microarray experiments including biological replicates, mixed ANOVA models should be favored. Microarray analysis produce huge amounts of raw data that indicate the need for bioinformatic solutions. The evaluation of these data, which can be handled only interdisciplinarily, represents several problems such as the lack of standards for comparing and exchanging such data. Here, MIAME standards (Minimum Information About a Microarray Experiment [42]) are the prerequisite to the worldwide comparability of gene expression data that are accessible through the deposition in NCBIs Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) [43 – 45].

2.2 Protein arrays

The concept and methodical approach of 'functional proteomics' can be explained through a lot of examples from 'functional genomics'. Furthermore, the development of protein microarrays shares the principle of miniaturization, parallelization and automation with DNA microarrays. Protein microarray technology allows the analysis of protein function on a whole genome level [46, 47].

Since the complete DNA sequences of a number of organisms from bacteria to man have been determined, and new techniques like microarrays for monitoring biomolecular interactions have been invented, an important milestone in the genome research, as well as in proteome research, has been achieved. The results of such high-throughput screening approaches can change our fundamental understanding of life's cellular processes on the molecular level. However, gene expression analysis is not sufficient to predict the function of a protein. Monitoring protein interactions is an extremely complex issue, because the proteome is the quantitative representation of the complete protein expression pattern of a cell, a tissue, an organ or an organism under exactly defined conditions. The proteome represents the protein equivalent of the genome, which is determined by the sequence, the type and number of its nucleotides. In contrast to this static nature of the genome, the proteome represents a tremendously dynamic object, which is influenced by a variety of parameters. Not all genes will be switched on at the same time in a cell and the sensitive balance between protein synthesis and protein degradation can vary widely under different metabolic or environmental

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

conditions. Consequently, a repeated analysis of the proteome will be successful only under exactly defined conditions, *e.g.*, cell culture conditions. In practice, it proves to be very difficult, since conditions to which the cell severely reacts are often unknown. However, on the other hand, the sensitive dependence of most different parameters offers the possibility to use specific small changes of the protein expression pattern as a sensitive biosensor.

Ideally, the analysis of the proteome delivers the currently available set of all proteins, if there is a way to maintain the quantitative relations of all proteins during the analysis. Such data cannot be obtained with classical molecular biology, since no strict connection between the amount of mRNA and the amount of protein exists. Parameters like mRNA stability, protein degradation, posttranslational modifications and others consequently prevent a statement over the currently available amount of protein. However, this information is of utmost importance, making a high-throughput analysis necessary. One attractive method is the use of microarrays, whereas the direct extrapolation of DNA microarray techniques for proteomic analysis have limited success, but in principal, the microarray technology enables both, the probing of the genome and the proteome. However, arraying of proteins is more difficult than the arraying of DNA, because they have to maintain their correctly folded conformations. The fabrication of protein arrays is, therefore, particularly challenging and protein arrays have lagged behind so far in development because of the more complex coupling chemistry, the instability of the immobilized protein and the far weaker detection signals [14]. In contrast to these technical problems, genome-wide screens for protein function are of biological importance for many applications such as: analyzing protein expression profiles, monitoring protein-protein interactions, identifying protein posttranslational modifications, screening the substrates of protein kinases, examining the protein targets of small molecules, and proteomic analysis as a function of bioprocess cultivation conditions.

Protein microarrays are miniaturized systems either with proteins used as capture probes or proteins used as targets, which should be bound by specific capture probes. Different proteins such as antibodies, antigens, peptides, receptors and enzymes can be bound on several formats (modified glass surfaces, hydrogel, membranes, nanoplates, micro fluidic chips). Protein microarrays consist of a solid support, *e.g.*, glass or synthetic material with a modified or coated surface. Capture probes, *e.g.*, antibodies, can be robotically transferred to this surface in the form of microspots (<200 μm) in rows and columns using special printers (pin-based or ink-jet printing heads). Every microspot contains only one kind of capture probes. These immobilized capture probes bind their specific target molecules from out of a complex solution. Today, the detection of immobilized antigens with antibodies is still the most common application [48, 49]. The bound target molecules can be attested by different detection methods which, *e.g.*, are based on fluorescence, chemiluminescence or radioactivity. Alternatively, protein chips for direct measurement of protein mass by MALDI-TOF mass spectrometry are available. Characterization of patient sera on protein chips will allow the diagnosis of diseases upon the presence of specific antibodies as well as the diagnosis of the vaccination protection of the patient. Once disease-specific antigens are known, it is possible to develop a disease-specific protein chip.

3 Integration of array technology into nutrigenomics and nutrigenetics

It is well known that cells respond with altered gene expression to changing environments and that nutrition can influence the proliferation and differentiation of cells. Alongside metabolome analysis, analysis of various components of the proteome, the transcriptome or the genome will become increasingly valuable. Hence, the integration of microarray analysis into basic and applied nutrition and food research provides new insights into the effects of nutrition and food ingredients like fats, carbohydrates, proteins, carotinoides, vitamins, minerals, flavonoids and xenobiotics at the molecular level [50–52]. The actual way in which some macronutrients and micronutrients [53] such as those from milk, fruits and vegetables, produce desirable changes in metabolism is largely unknown. DNA chip technology will open up new ways of studying nutrition in more depth, and identifying far more possible targets. The study of nutrigenomics, *i.e.*, the relationship between specific nutrients or diet and gene expression, could help to identify these effects and may facilitate the prevention of common diet-related diseases. Further areas of application are food safety, food authenticity, testing for genetically modified organisms (GMOs) and food design.

3.1 Nutrition-related diseases and predispositions

Comparing transcriptomic and proteomic analyses can be used, *e.g.*, to examine diet-related defects on the molecular level. The concept of nutrigenomics arose from the knowledge that diet, as the perhaps most important environmental factor, has a permanent effect on the genome. Therefore, diet directly influences metabolic processes, leading to lasting effects on physical health. Nutritionists consequently start to analyze the molecular effects of nutrients on the cell, the genome and, in the end, on the complete organism. The individual genetic disposition represents another significant factor of diet-related illness. The complete sequencing of human and other genomes revealed, amongst others, genes playing key roles in the metabolism of nutri-

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

ents. The knowledge that diet of western industrial nations can cause various diseases is one reason for the increased interest of nutritionists in the field of genomics. Nowadays, there are two large networks analyzing the effects of diet on the genome: the National Center of Excellence in Nutritional Genomics (NCMHD, http://nutrigenomics.ucdavis.edu/), a division of the National Institutes of Health (NIH, http://www.nih.gov), and the European Nutrigenomics Organization (NuGO, http://www.nugo.org).

Gene expression studies allow the efficient identification of molecular mechanisms underlying the pathogenesis of diet-associated diseases such as obesity, diabetes, osteoporosis, arteriosclerosis, hypertension, birth defects, or food allergies. Some studies focus on the interaction between pro-inflammatory stress and metabolic stress as a key role to understand diet-related diseases, the identification of transcription factors that function as nutrient sensors, fatty acid metabolism, or nutrient-induced alterations to homeostasis [54 – 72]. DNA chip technology, therefore, enables large numbers of genes to be screened simultaneously, giving a comprehensive, detailed picture of changes in gene expression, shedding light on complex regulatory interactions.

3.2 Individualized medicine

Today, one of the central goals of genome research is the study on genetic variance among individuals to improve our understanding and treatment of disease, and to identify the variations in our genes that influence an individual's risk of becoming ill. Because inherited differences influence most common diseases and many drug responses, the knowledge of existing genetic variations in the human population is a promising approach in the development of tailor therapeutics. Of these individual variations, 80% are SNPs, and these SNPs have a direct impact on gene regulation or on a protein product of a gene. SNP means that a single base is swapped for an alternate. In cases where a single SNP is sufficient to cause disease, it is possible to analyze the causal change and to improve our understanding of disease [73].

SNPs are the most common form of genetic variant in the human genome, occurring on average 1 per 1000 bps [74]. It is assumed that of the order of 3 million SNPs will be derived from comparison of genomic sequence from individuals among several populations [75]. SNP diagnostics are coming into the focus of interest, since the influence of SNPs on medical therapy or diet is better understood. In 1999, ten pharmaceutical companies put aside their differences and created the SNP Consortium, an international collaboration of academic institutes, a private foundation and pharmaceutical companies (http://snp.cshl.org; public SNP database: http://www.ncbi.nlm.nih.gov/SNP/). Today, the number of regulatory SNPs is still unknown. However, it is possible to use SNPs to track associations to disease using a phenomenon called linkage disequilibrium (LD). SNP maps will facilitate association studies of unrelated individuals for either disease, drug or diet response, although LD has not been successfully applied in whole-genome studies. The identification of a complete human SNP map would considerably increase the efficiency of drug development and medical treatment of disease. Many genotyping platforms have been developed to detect many SNPs at once in an efficient, cost-effective manner including DNA microarray technologies. By the use of allele-specific oligonucleotides containing the variable base, microarray-based biosensors provide a suitable platform for SNP detection and analysis. Large scale SNP genotyping should lead to a fully understanding of the genetic architecture of common traits underlying disease, drug or diet response. Because these are the most personal data we know, access to these data must be protected avoiding misuse by employers or insurers. The goal of this analysis is to relate SNP variation to disease or diet, leading to individualized dietary recommendations and individualized food. Such engineered food should increase the nutritional value and health effects, and decrease allergenicity.

3.3 Functional food

Due to the knowledge of the human genome sequence, molecular diagnostic methods for the determination of nutritional variation are becoming more relevant, as the genetic basis for variation in nutritional response is discovered, and the shift from single genes to whole genomes is reflected. The identification of the cellular response to a nutritional signal may provide ways to decipher the mechanism by which a nutritional signal is transduced into a given response. Today, little is known about the molecular mechanisms by which the genome perceives nutritional signals and mobilizes the organism to respond. Such information is important in any future attempts to engineer food products that influence many chronic diseases such as cancer, heart disease, diabetes, Alzheimer's disease, osteoporosis, obesity, hypertension, immune system responses by diet, or to design components of food that have the power to provide substantial health benefits.

In contrast to single gene disorders such as cystic fibrosis, multifactorial, nutrition-based diseases like hypertension or obesity are more complex, as disease onset likely depends on changes in several genetic regions as well as in nutritional factors. Due to nutritional differences, only a few of the patients carrying the mutations will develop the disease. The multifactorial complexity will complicate prediction of nutrition response outcome. Nevertheless, the determination of the molecular effects of food components by DNA

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

microarray experiments will in future afford the development of functionalized food, *e.g.*, for cholesterol reduction.

3.4 Food safety

Food safety in regard to diet and potable water is related to food ingredients itself, as well as to external food contaminations so-called food-borne or water-borne pathogens [76]. External contamination occur during harvesting, processing and storage, and include enteric bacteria, viral pathogens, bacterial-derived toxins and by-products as well as residues like herbicides, antibiotics, pharmaceuticals, and so on. These by-products could end up in the food chain and accordingly affect human health.

3.4.1 Detection of pathogens

Using a molecular approach for detection, typing, and identification of microorganisms, the identification of the organisms by microarrays saves time compared to conventional cultivation. The identification of pathogen bacteria by molecular genetics can be advanced further by DNA microarray technology [77, 78]. 16S rRNA has been reported to be a suitable target for identifying various specific microorganisms, because 16S rRNA genes show species-specific polymorphisms. Sometimes, the number of substituted bases between the 16S rRNA genes of closely related bacterial strains is either nonexistent or too small to differentiate between these species because of the extremely slow speed of the molecular evolution of 16S rRNA. Species-specific genes could be used as an alternative to 16S rRNA analysis [79–81]. Microarray experiments are a useful tool for phylogenetic analysis and species identification. For the identification of bacteria, this method may involve the labeling of a specific DNA sequence from a target gene or a specific 16S rRNA sequence, subsequent hybridization to species-specific oligonucleotide probes on a microarray, and detection of the label, usually by fluorescence. Species-specific hybridization patterns on a microarray containing strain-specific probes can differentiate and identify bacteria on the species level. Currently, the former Nutri®Chip in a modified form is used for the detection of the presence of pathogenic bacteria or indicator organisms in the chocolate processing industry.

The microbial decontamination or removal of microbial by-products in water becomes increasingly important. More and more closed water management systems (*e.g.*, hospitals) need such decontamination modules. Different electrochemical or chemical water treatment modules are used for microbial decontamination, and these devices are necessary for an efficient decontamination of all kind of drinking or pharmaceutical water supply systems. A high-level quality control of these devices is possible through the identification of the pathogens by microarrays and proceeds much faster than by conventional cultivation.

3.4.2 Food authenticity

DNA chips are also already being used in food analysis: with the help of the CarnoCheck® it is possible to detect eight different animal species in food products and animal feed. Given the current crisis in agriculture and the food industry as a result of the BSE problem, the issue regarding the monitoring of an accurate food-labeling procedure becomes more urgent. Furthermore, incorrect or vague declarations in regard to foodstuffs can cause allergies. CarnoCheck® allows the quick and efficient identification of eight species in all (horse, ass, beef, pork, goat, sheep, turkey and chicken), even from processed animal feed and food products.

3.4.3 Testing for genetically modified organisms

The GMO means 'an organism, with the exception of human beings, in which the genetic material has been altered in a way that does not occur naturally by mating and/or natural recombination' [82]. GMOs, derived through genetic engineering, may be used as food (*e.g.* GM soybean, maize, or tomatoes), feeds, seeds, forestry materials or pharmaceuticals [82]. The use of GMOs in food industry has been ascending exponentially over the past years. During the period 1996–2003, global area of transgenic crops increased 40-fold, from 1.7 million hectars in 1996, to 67.7 million hectars in 2003, with a dominant trait of herbicide tolerance, followed by insect resistance (http://www.isaaa.org, International Service for the Acquisition of Agri-biotech Applications). Recent GM rules in the EU relate to the authorization of GM food and feed, their labeling and traceability, and will affect shipments of non-GM products which may contain adventitious levels of GMOs. These new regulations for food and feed marketed in the EU [Regulation (EC) No 1829/2003 and No 1830/2003 of the European Parliament and of the Council; http://europa.eu.int/] implicate a credible identification and documentation system for GMOs in food and feed products. Each GMO is defined by a so-called unique identifier, a code of letters and digits, referring to the consent holder, the transformation event and a verification number. These unique identifiers, assigned to a specific GMO, can be used for its detection normally by DNA amplification of the specific DNA fragment. The rapid introduction of novel GMOs associated with the legal requirements in food industry necessitate the cost-efficient and reliable testing of foods for the unique identifiers of GMOs.

The application of array techniques to GMO detection imply the capability of standardization, reproducibility and sensitivity of the detection system. Nowadays, commercia-

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

lized arrays do not fulfill these requirements, in spite of a few promising attempts initialized by biotechnical companies and research projects (*cf.* http://www.molspec.org). The need for exact quantities of GMOs in foods to comply with legal requirements, makes the introduction of biochips as a method for the routine analysis so difficult, particularly with regard to well-established quantitative real-time PCR procedures.

Presently, the DNA microarray technology can only be used to monitor alterations in gene expression levels of GMO-derived food products in comparison to traditional food products. The main limitations of such whole-genome gene expression analysis are the high biological variance of the samples, and the multiple effects of complex substances such as food on various pathways. Moreover, the representative effects of accessible samples, such as blood, instead of primary cells from tissues can not be validated. The characterization of potential advantages or damages caused by the usage of GMO-derived food products, however, can not be analyzed by DNA microarray experiments.

3.5 Safety, ethics and protection of privacy

Microarray experiments allow the comparison of GM food with conventionally produced food in respect of food safety, food authenticity, food contamination during harvesting, processing and storage as well as food-borne pathogens. Microarray analysis can be used for a rigorous evaluation of GMOs, GM food and genetically engineered ingredients relative to human health and environment. The WHO defined the following safety assessments of GM food: (a) toxicity, (b) allergenicity, (c) specific components thought to have nutritional or toxic properties, (d) stability of inserted genes, (e) nutritional effects associated with genetic modification, and (f) any unintended effects.

In regard to allergenicity, the assessment process includes the investigation of the allergenic potential of the protein product as well as the transferred gene. Additionally, contamination of food with herbicides, pharmaceuticals, antimicrobial substances, enteric pathogens and animal cells are important food safety criteria, focusing in particular on long-term effects.

While the nutrigenomic research is still at the beginning, some researchers have already thought about the social, ethical, legal and medical consequences arising from the scientific finding. For example, scientists of the Canadian universities in Guelph and Toronto forecast that nutrigenomics will lead to radical changes at the production, the processing and the consumption of food. In their study 'Nutrition and Genes: Science, Society and the Supermarket' the authors warn against rushing headfirst at the adventure nutrigenomics. The report raises some important questions such as: When are scientific data sufficiently validated to sell diagnostic tests? Who should not have access to nutrigenomic data, who should? How should this use be checked? Which nutrigenomic interests should be supervised and/or regulated?

4 Discussion and conclusions

Genomes can respond in a rapid and specific manner by selectively increasing or decreasing the expression of specific genes, and this responses can be utilized to investigate the molecular events by which the genomes perceive nutritional signals and mobilizes the organism to respond. Food can affect gene expression on the cellular level, *e.g.*, due to the interaction with receptors as well as metabolic pathways and signal pathways. This means that medicine, biology and the food sciences will develop ever more markers for illness symptoms and food contents. Appropriate analytical methods must, therefore, be developed to measure these. Here, simultaneous gene analysis with DNA microarrays provides a rapid and efficient method for large-scale and high-throughput applications, as an expression pattern of thousands of genes can be obtained from a single experiment. They also allow rapid and cost-effective screens for mutations and sequence variations in genomic DNA and provide new insights into fundamental biological processes of human diseases at the molecular level. Many aspects in the pioneering field of studying the direct and indirect interaction of nutrition and gene expression, *e.g.*, diseases like diabetes, cancer, obesity and cardiovascular disease, high triglyceride levels, high blood pressure and high fasting blood glucose levels, allergies and intolerances, as well as microbiological contamination of food, are already analyzed using microarrays. Matsumoto *et al*. [83] investigated the gene expression of cranial sensory ganglia that transmit food intake stimuli, and found several genes differentially expressed in cranial sensory ganglia associated with neurotransmission. Eitsuka *et al.* [84] showed that polyunsaturated fatty acids inhibit telomerase activity, thereby having an inhibiting influence on tumor development, because telomerase activity is mostly only present in cancer cells. Analysis of the effect of hypoallergenic wheat flour by Kato *et al*. [62] demonstrated that the enzymes involved in cholesterol synthesis and catabolism were up-regulated in gluten-fed rats. Nowadays, many diet-regulated genes, *e.g.*, glucose-6-phosphate dehydrogenase or NAD, are known to play a key role in the onset and progression of diseases. Often, these diseases are affected by complex interactions and the balance between genes, environmental factors and the individual SNP map.

This short listing points to the fact that only the synergistic application of different molecular biological techniques in transcriptomics, proteomics and metabolomics will lead to the essential survey of multifactorial, nutritional influences

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

on the human organisms. Diet finally means the incorporation of hundreds of different nutrients. Therefore, statistics and complexity produce not only experimental difficulties, they are also the main problems of nutrigenomics. In contrast to pharmaceuticals, which are well-defined compounds, foods are complex substances and these meet innumerable combinations of genetic variants. Additionally, there are problems with the use of established *in vitro* or xenobiotic metabolism models for this purpose. Primary cells as well as cell lines are available as test systems for studying the effect of nutrients on gene expression, but there are a few drawbacks of both cell types. Primary cells display inter-individual differences caused, for example, by age and gender of the donor, and the widely used HepG2 cells are of no use for investigating nutritional effects [64], because some pathways are different from that in normal hepatocytes and some enzyme systems like the cytochrome P450 system are not active in HepG2 cells after a few days in culture. The disadvantage of xenobiotic models is the purification of the RNA from blood, while little is known about gene expression in peripheral blood. Whitney *et al*. [85] describe the nature of inter-individual and temporal variation of gene expression patterns in human blood. Without sound quality control, microarray experiments may produce useless or, even worse, misleading results.

Despite some limitations concerning the laboratory usage and the data analysis of microarray experiments, this technology has become accessible to a wide community of scientists. DNA chip technology enables large numbers of genes to be screened simultaneously, giving a comprehensive, detailed picture of changes in gene expression, shedding light on complex regulatory interactions, *e.g.*, diet-gene interactions.

Established chip facilities in the academic area and commercial suppliers enable the integration of array experiments into projects without time-consuming establishment of standard protocols or acquisition of expensive equipment. By the assignment of the highly developed DNA chip technology to new areas like human nutrition, protein analytics, environmental analytics and medical diagnostics a simple, effective and cost-efficient analysis of processes like diet-regulated gene disease interactions will be soon possible. Studies in this field have shown first success in cases of hypertension and cardiovascular health. Using microarrays offers a number of advantages in controlling disease by dietary intervention in form of an individualized nutrition. DNA chips can be designed to detect sequence variations that have relevance to disease status, or altered enzyme activity to analyze the influence of food compounds to human health and disease. Research into gene expression and SNPs could help to identify ways for more tailored, effective, personalized food and dietary recommendations and reduce the risk to diseases. Based on these results, functional food with improved nutritional value could be developed. However, such functional food cannot replace balanced diet and physical training. People have to improve their lifestyle, increase their physical fitness, and avoid unhealthy diet, excessive smoking and alcohol consumption. So it will be a long way to go until diet can be coordinated with the results of microarray analysis, and until the testing for individual genetic predispositions will influence the shopping list in future and improve human health in combination with individual healthful lifestyle.

5 References

[1] Ermolaeva, O., Rastogi, M., Pruitt, K. D., Schuler, G. D., Bittner, M. L., Chen, Y., Simon, R. *et al*., *Nat. Genet*. 1998, *20*, 19–23.

[2] Woychik, R. P., Klebig, M. L., Justice, M. J., Magnuson, T. R., Avner, E. D., *Mutat. Res*. 1998, *400*, 3–14.

[3] Vukmirovic, O. G., Tilghman, S. M., *Nature* 2000, *405*, 820.

[4] Tavazoie, S., Hughes, J. D., Campbell, M. J., Cho, R. J., Church, J. M., *Nat. Genet.* 1999, *22*, 281–285.

[5] Schena, M., Shalon, D., Davis, R. W., Brown, P. O., *Science* 1995, *270*, 467–470.

[6] Khan, J., Bittner, M. L., Chen, Y., Meltzer, P. S., Trent, J. M., *Biochim. Biophys. Acta* 1999, *1423*, M17–M28.

[7] Duggan, D. J., Bittner, M., Chen, Y., Meltzer, P., Trent, J. M., *Nat. Genet*. 1999, *21*, 10–14.

[8] Brown, P. O., Botstein, D., *Nat. Genet*. 1999, *21*, 33–37.

[9] Schena, M., Shalon, D., Davis, R. W., Brown, P. O., *Science* 1995, *270*, 467–470.

[10] Southern, E. M., Maskos, U., Elder, J. K., *Genomics* 1992, *13*, 1008–1017.

[11] Grunstein, M., Hogness, D. S., *Proc. Natl. Acad. Sci. USA* 1975, *72*, 3961–3965.

[12] Benton, W. D., Davis, R. W., *Science*. 1977, *196*, 180–182.

[13] Shalon, D., Smith, S. J., Brown, P. O., *Genome Res*. 1996, *6*, 639–645.

[14] *Nat. Genet*. 1999, *21*, 1–60.

[15] Fodor, S. P., Read, J. L., Pirrung, M. C., Stryer, L., Lu, A. T., Solas, D., *Science* 1991, *251*, 767–773.

[16] Lemieux, B., Aharoni, A., Schena, M., *Mol. Breeding* 1998, *4*, 277–289.

[17] Schena, M., Heller, R. A., Theriault, T. P., Konrad, K., Lachenmeier, E., Davis, R. W., *Trends Biotech.* 1998, *16*, 301–306.

[18] Heller, R. A., Schena, M., Chai, A., Shalon, D., Bedilion, T., Gilmore, J., Woolley, D. E., Davis, R. W., *Proc. Natl. Acad. Sci. USA* 1997, *94*, 2150–2155.

[19] Schena, M., Shalon, D., Heller, R., Chai, A., Brown, P. O., Davis, R. W., *Proc. Natl. Acad. Sci. USA* 1996, *93*, 10614–10619.

[20] Schena, M., *Bioessays* 1996, *18*, 427–431.

[21] Eberwine, J., Yeh, H., Miyashiro, K., Cao, Y., Nair, S., Finnell, R., Zettel, M., Coleman, P., *Proc. Natl. Acad. Sci. USA* 1992, *89*, 3010–3014.

[22] Phillips, J., Eberwine, J. H., *Methods* 1996, *10*, 283–288.

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

[23] Matz, M., Shagin, D., Bogdanova, E., Britanova, O., Lukyanov, S., Diatchenko, L., Chenchik, A., *Nucleic Acids Res*. 1999, *27*, 1558 – 1560.

[24] Kerr, M. K., Churchill, G. A., *Biostatistics* 2001, *2*, 183 – 201.

[25] Landgrebe, J., Bretz, F., Brunner, E., *In Silico Biol*. 2004, *4*, 0038.

[26] Katzer, M., Kummert, F., Sagerer, G., *IEEE Trans. Nanobiosci*. 2003, *2*, 202 – 214.

[27] Angulo, J., Serra, J., *Bioinformatics* 2003, *19*, 553 – 562.

[28] Yang, Y. H., Dudoit, S., Luu, P., Lin, D. M., Peng, V., Ngai, J., Speed, T. P., *Nucleic Acids Res*. 2002, *30*, e15.

[29] Kerr, M. K., Martin, M., Churchill, G. A., *J. Comput. Biol*. 2000, *7*, 819 – 837.

[30] Anderssen, R. S., Wu, Y., Dolferus, R., Saunders, I., *Bioinformatics*. 2004, *20*, 1721 – 1727.

[31] Chen, Y., Dougherty, E. R., Bittener, M., *J. Biomed. Opt*. 1997, *2*, 364 – 374.

[32] Cleveland, W. S., Devlin, S. J., *J. Am. Stat. Assoc*. 1988, *83*, 596 – 610.

[33] Dombkowski, A. A., Thibodeau, B. J., Starcevic, S. L., Novak, R. F., *FEBS Lett*. 2004, *560*, 120 – 124.

[34] Eisen, M. B., Spellman, P. T., Brown, P. O., Botstein, D., *Proc. Natl. Acad. Sci. USA* 1998, *95,* 14863 – 14868.

[35] Eisen, M. B., Brown, P. O., *Methods Enzymol*. 1999, *303*, 179 – 205.

[36] Glonek, G. F., Solomon, P. J., *Biostatistics* 2004, *1*, 89 – 111.

[37] Hessner, M. J., Singh, V. K., Wang, X., Khan, S., Tschannen, M. R., Zahrt, T. C., *BMC Genomics* 2004, *5*, 12.

[38] Heyer, L. J., Kruglyak, L., Yooseph, S., *Genome Res.* 1999, *9*, 1106 – 1115.

[39] Kane, M. D., Jatkoe, T. A., Stumpf, C. R., Lu, J., Thomas, J. D., Madore, S. J., *Nucleic Acids Res.* 2000, *28*, 4552 – 4557.

[40] Mah, N., Thelin, A., Lu, T., Nikolaus, S., Kuehbacher, T., Gurbuz, Y., Eickhoff, H., *et al*., *Physiol. Genomics* 2004, *16*, 361 – 370.

[41] Nagpal, S., Karaman, M. W., Timmerman, M. M., Ho, V. V., Pike, B. L., Hacia, J. G., *Nucleic Acids Res*. 2004, *32*, 51.

[42] Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoeckert, C., Aach, J., *et al*., *Nat. Genet*. 2001, *29*, 365 – 371.

[43] Quackenbush, J., *Nat. Genet*. 2001, *2*, 418 – 427.

[44] Edgar, R., Domrachev, M., Lash, A. E., *Nucleic Acids Res*. 2002, *30*, 207 – 210.

[45] Barrett, T., Suzek, T. O., Troup, D. B., Wilhite, S. E., Ngau, W. C., Ledoux, P., Rudnev, D., *et al*., *Nucleic Acids Res*. 2005, *33 Database Issue*, D562 – D566.

[46] MacBeath, G., *Nat Genet*. 2002, *32 Suppl.*, 526 – 532.

[47] Templin, M. F., Stoll, D., Schrenk, M., Traub, P. C., *Trends Biotechnol*. 2002, *20*, 160 – 166.

[48] Eickhoff, H., Konthur, Z., Lueking, A., Lehrach, H., Walter, G., Nordhoff, E., Nyarsik, L., Bussow, K., *Adv. Biochem. Eng. Biotechnol.* 2002, *77*, 103 – 112.

[49] Bussow, K., Konthur, Z., Lueking, A., Lehrach, H., Walter, G., *Am. J. Pharmacogenomics* 2001, *1*, 37 – 43.

[50] Chatterjee, M., *Mutat. Res*. 2001, *475*, 69 – 87.

[51] Fenech, M., *Mutat. Res*. 2001, *475*, 57 – 67.

[52] Tsuda, S., Matsusaka, N., Ueno, S., Susa, N., Sasaki, Y. F., *Toxicol. Sci*. 2000, *54*, 104 – 109.

[53] Liu-Stratton, Y., Roy, S., Sen, C. K., *Toxicol. Lett*. 2004, *150*, 29 – 42.

[54] Muller, M., Kersten, S., *Nat. Rev. Genet*. 2003, *4*, 315 – 322.

[55] Bauer, M., Hamm, A., Pankratz, M. J., *Biol. Chem*. 2004, *385*, 593 – 596.

[56] van der Meer-van Kraaij, C., Kramer, E., Jonker-Termont, D., Katan, M. B., van der Meer, R., Keijer, J., *Carcinogenesis* 2005, *26*, 73 – 79.

[57] Stover, P. J., *Physiol. Genomics* 2004, *16*, 161 – 165.

[58] Poulsen, L. K., *Mol. Nutr. Food Res.* 2004, *48*, 413 – 423.

[59] Page, G. P., Edwards, J. W., Barnes, S., Weindruch, R., Allison, D. B., *Nutrition* 2003, *19*, 997 – 1000.

[60] Nakagawa, K., Eitsuka, T., Inokuchi, H., Miyazawa, T., *Biofactors* 2004, *21*, 5 – 10.

[61] Matsumoto, I., Nakamura, S., Emori, Y., Arai, S., Abe, K., *Biofactors* 2004, *21*, 15 – 18.

[62] Kato, H., Narasaka, S., Endo, Y., Abe, K., Arai, S., *Biofactors* 2004, *21*, 11 – 13.

[63] Kato, H., Kimura, T., *J. Nutr.* 2003, *133 Suppl. 1*, 2073S – 2077S.

[64] Kaput, J., Rodriguez, R. L., *Physiol. Genomics* 2004, *16*, 166 – 177.

[65] Goi, T., Fujioka, M., Satoh, Y., Tabata, S., Koneri, K., Nagano, H., Hirono, Y., *et al*., *Cancer Res*. 2004, *64*, 1906 – 1910.

[66] Elliott, R., Ong, T. J., *BMJ* 2002, *324*, 1438 – 1442.

[67] Eitsuka, T., Nakagawa, K., Miyazawa, T., *Biofactors* 2004, *21*, 19 – 21.

[68] Daniel, H., tom Dieck, H., *Biol. Chem*. 2004, *385*, 571 – 583.

[69] Daniel, H., *Br. J. Nutr*. 2002, *87 Suppl. 2*, S305 – S311.

[70] Chavez, A., Munoz de Chavez, M., *Eur. J. Clin. Nutr.* 2003, *57 Suppl. 1*, S97 – S100.

[71] Chadwick, R., *Proc. Nutr. Soc. USA* 2004, *63*, 161 – 166.

[72] van Ommen, B., *Nutrition* 2004, *20*, 4 – 8.

[73] Staunton, J. E., Slonim, D. K., Coller, H. A., Tamayo, P., Angelo, M. J., Park, J., Scherf, U., *et al*., *Proc. Natl. Acad. Sci. USA* 2001, *98*, 10787 – 10792.

[74] Collins, F. S., Brooks, L. D., Chakravarti, A., *Genome Res*. 1998, *8*, 1229 – 1231.

[75] Halushka, M. K., Fan, I. B., Bently, K., Hsie, L., Shen, N., Weder, A., Cooper, R., *et al*., *Nat. Genet.* 1999, *22,* 239 – 247.

[76] Wang, D. G., Fan, J. B., Siao, C. J., Berno, A., Young, P., Sapolsky, R., Ghandour, G., *et al., Science* 1998, *280*, 1077 – 1082.

[77] Call, D. R., Brockman, F. J., Chandler, D. P., *Int. J. Food Microbiol*. 2001, *67*, 71 – 80.

[78] Chizhikov, V., Rasooly, A., Chumakov, K., Levy, D. D., *Appl. Environ. Microbiol*. 2001, *67*, 3258 – 3263.

[79] Nagano, I., Kunishima, M., Itoh, Y., Wu, Z., Takahashi, Y., *Microbiol. Immunol*. 1998, *42*, 371 – 376.

[80] Yu, J., Kaper, J. B., *Mol. Microbiol*. 1992, *6*, 411 – 417.

[81] Prager, R., Strutz, U., Fruth, A., Tschape, H., *Int. J. Med. Microbiol*. 2003, *292*, 477 – 486.

[82] Dabrowska, P., *German Law Journal* 2004, *5*, 151 – 166.

[83] Matsumoto, I., Nakamura, S., Emori, Y., Arai, S., Abe, K., *Biofactors* 2004, *21*, 15 – 18.

[84] Eitsuka, T., Nakagawa, K., Miyazawa, T., *Biofactors* 2004, *21*, 19 – 21.

[85] Whitney, A. R., Diehn, M., Popper, S. J., Alizadeh, A. A., Boldrick, J. C., Relman, D. A., Brown, P. O., *Proc. Natl. Acad. Sci. USA* 2003, *100*, 1896 – 1901.

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim