 DOI 10.1002/mnfr.200900304 *Mol. Nutr. Food Res.* 2010, *54*, 240–248

REVIEW

Transcriptome analysis in benefit–risk assessment of micronutrients and bioactive food components

Jaap Keijer[1], *Yvonne G. J. van Helden*[1,2,3], *Annelies Bunschoten*[1] *and Evert M. van Schothorst*[1]

[1] Human and Animal Physiology, Wageningen University, Wageningen, The Netherlands
[2] Department of Health Risk Analysis and Toxicology, Research Institute NUTRIM, Maastricht University, Maastricht, The Netherlands
[3] RIKILT-Institute of Food Safety, Wageningen, The Netherlands

The establishment of functional effects due to variation in concentrations of micronutrients in our diet is difficult since they are often not immediately recognized as being healthy or unhealthy. Indeed, effects induced by micronutrients are hard to identify and therefore the establishment of the recommended daily intake, the optimal intake and the upper limit pose a challenge. For bioactive food components this is even more complicated. Whole-genome transcriptome analysis is highly suitable to obtain unbiased information on potential affected biological processes on a whole-genome level. Here, we will describe and discuss several aspects of transcriptome analysis in benefit–risk assessment, including effect size, sensitivity and statistical power, that have to be taken into account to faithfully identify functional effects of micronutrients and bioactive food components.

Received: June 30, 2009
Revised: September 3, 2009
Accepted: September 7, 2009

Keywords:
DNA microarray analysis / Methodological considerations / Nutrigenomics / Power / Sensitivity

1 Introduction

It is well established that different intakes of micronutrients affect our health. How they do this over a range of concentrations is less clear. This information is important for determining optimal intakes and for the application of micronutrients in dietary strategies to increase a healthy life span. Clearly, deficient as well as excessive intakes of micronutrients can result in physiological changes and ultimately in pathological effects. Between these extremes, many physiological effects over a range of concentrations can occur, with a certain optimal beneficial intake. Benefit–risk assessment means to establish the borders where shortage changes into enough, defined as the recommended daily intake (RDI), and where enough changes into too much, defined as the upper limit (UL), as well as the establishment of the optimal beneficial intake (Fig. 1, [1]). The RDI is established based on the minimal intake that is needed to avoid deficiency symptoms in 97% of the population. While in some cases this is already difficult to establish, it is even more difficult to determine the optimal beneficial concentration, supporting maximal health over a lifetime. Indeed for most micronutrients this is not known. To be able to assess this properly, an overview of all possible effects is needed and functional genomic technologies are potentially suitable to obtain such an overview. What is true for micronutrients is also true for bioactive food components with claimed beneficial health effects as a consequence of habitual intake, such as polyphenols and carotenoids. Of course, for bioactive food components functionality has to be established first.

The question can be raised whether a slightly higher intake than the RDI can be of additional functional importance. That this may be true has been shown for the micronutrient selenium, which, to our knowledge, is the only example available at present. It is well known that

Correspondence: Professor Jaap Keijer, Human and Animal Physiology, Wageningen University, P. O. Box 338, 6700 AH Wageningen, The Netherlands
E-mail: jaap.keijer@wur.nl
Fax: +31-317-484077

Abbreviations: CR, calorie restriction; **EGCg,** epigallocatechin-gallate; **Q,** quercetin; **R,** resveratrol; **RDI,** recommended daily intake; **UL,** upper limit

(severe) selenium deficiency results in Keshan disease, Kashin–Beck disease, and even symptoms of hypothyroidism due to its function in thyroid metabolism (reviewed by [2]), whereas an excess intake results in toxicity known as selenosis [3]. Although epidemiological studies showed positive health effects of high selenium intake, the underlying mechanisms were not clear, making it difficult to causally link intake to effect. Gene expression analysis of human lymphocytes showed an increase in the protein translation machinery when the selenium status was increased from just below to a level just above the RDI [4]. The same result, an increase in the protein translation machinery, was also found when gene expression in the colon of wildtype C57BL/6J mice on marginal deficient and adequate selenium diets were compared [5], thereby validating the observations made in humans. At present it is not clear whether the effect on the protein translation machinery is related to the main physiological effects associated with deficiency (*e.g.* Keshan disease), but its identification allows to establish not only this causality, but also the minimal, and possibly optimal, required selenium intake for optimal protein translation capacity, making it a potential biomarker.

This example indicates that understanding the mechanisms of nutrient-induced functional effects is important, since they may affect fitness, well-being and healthy aging. Moreover, understanding the mode of action will help to define susceptible population groups. By definition, no overt pathology is involved at micronutrient intakes between the RDI and UL and we are thus faced with the task to identify small functional and molecular effects. This poses constraints on the experimental setup and on the analysis techniques. Furthermore, we do not know which effects will occur and the use of an initial unbiased approach is essential. Whole-genome transcriptome analysis is highly suited for this. In this review we describe this technique and several technical as well as experimental considerations with regard to its application in nutritional studies in general and in benefit–risk assessment in more detail.

2 Whole-genome transcriptome analysis as a tool for benefit–risk analysis

A highly suitable tool toward the identification of mechanisms underlying nutrient-induced functional effects is whole-genome expression profiling (also known as whole-genome DNA microarray analysis). This is a highly comprehensive profiling technique, able to screen for differences in gene expression between (groups of) individuals. This technique has become robust over the past years, and is able to detect small differences between individuals and thus suitable as a first phase screen in benefit–risk analysis.

In transcriptome profiling, global gene expression in one condition is compared with another condition. There are several different microarray platforms, of which those by Affymetrix and Agilent are mostly used. To determine global gene mRNA expression, both systems use a linear amplification and labeling procedure to obtain cRNA, which is hybridized to the microarray. Affymetrix uses statistical ranking of the binding of cRNA to multiple sense *versus* missense probes *per* probe set, with each probe being 25 bases in length. In contrast Agilent probes are 60 bases in length, but only one or a limited number of probes are used to represent an expression product. A second difference is that the Affymetrix platform uses a single dye to label an experimental sample, and normalization is fully statistical, whereas Agilent arrays can hybridize and detect two or even more dyes, for example one for the experimental sample and one for a reference sample. Both types of arrays have become highly robust and allow generation of high quality data [6]. More details are given below. The Agilent platform is used as an example in this paper.

Agilent whole-genome microarrays contain roughly 44 000 probes, including positive and negative controls, which represent over 40 000 different annotations. Since some annotations represent the same gene, the arrays allow to screen for differences in expression of over at least 25 000 unique genes, plus around 1400 annotations of un-assigned genes. Since generally two dyes are used, direct comparison of two different experimental samples is possible, although this approach has two major disadvantages. The first disadvantage is that a dye swap is necessary, since the two dyes give rise to different results, due to chemical differences in the labeling procedure. The second disadvantage is that only pre-defined comparisons can be analyzed and experiments comparing more experimental groups require highly intricate designs. To circumvent these disadvantages a reference pool design can be used. A reference pool design allows for *post hoc* selection of the experimental groups to be compared and a dye swap is not necessary. In a reference pool design, each experimental sample is individually labeled with a fluorescent dye (*e.g.* Cy5-cRNA) and hybridized to a microarray together with the same amount of a mixed sample, ideally containing an equal molar mixture of all experimental samples, labeled with a second dye (*e.g.* Cy3-cRNA) [7, 8]. Five hundred nanogram of total input RNA suffices in the amplification and labeling reaction [8]. After hybridization, the microarrays are washed and scanned. The fluorescent intensity for each probe on the array is then integrated and represents a quantitative measure for the amount of the corresponding mRNA in the sample. The microarrays first pass a quality control pipeline that controls labeling and hybridization. Since values close to the background may give aberrant fold changes between groups, these are generally discarded for further analysis. We normally discard all signal values that are on average below two times the local background values for both the samples and the reference pool, although with present day microarray reproducibility lower cut-offs are possible. In general, at least 50–60% of the probes are being expressed two times above background in any tissue analyzed.

In the next step, expressed probe intensities are normalized by correcting the signals relative to the reference sample probe signals [7]. The use of the reference sample facilitates experimental normalization and, compared with single dye experiments, is less sensitive to technical variation such as hybridization conditions and microarray quality. The use of a method that is low in technical variation for benefit–risk assessment is crucial, since changes in gene expression are small and the used method has to be sensitive enough to separate these changes from the technical noise. In the end, a fluorescent normalized log intensity value is obtained for each of the expressed probes on the array, for each experimental sample. This results in a datasheet that is used for downstream applications, such as statistical analyses, pathway overrepresentation analysis, cluster analysis, and other methods that facilitate data interpretation (see Section 3).

3 Data-confirmation by qRT-PCR

Changes in gene expression identified using whole-genome microarrays are usually confirmed by quantitative RT-PCR (qRT-PCR). In qRT-PCR, specific mRNA's are amplified using intron-spanning primers and amplification is monitored in real time allowing mRNA quantification. To normalize for possible variation in the amount and quality of RNA between different samples, the level of a target gene is compared and expressed relative to that of one or more reference genes. Therefore, qRT-PCR has a semi-quantitative character and the choice of reference genes is particularly important for the validity of the results. The reference genes should be stably expressed over all the conditions that are analyzed. For example, Glyceraldehyde-3-phosphate dehydrogenase, historically used as a reference gene in mRNA Northern blot quantification, shows 1.5 to 3-fold changes in many dietary interventions (Table 1). It is therefore generally not suitable as a reference gene, especially since changes of around 1.3-fold are often encountered as being physiologically relevant in nutritional studies. Other reference gene candidates such as 18S rRNA, general transcription factor II and microarray identified stable expressed genes are better suited as reference candidates. Nevertheless, testing the stability of the expression of different reference genes for the experimental conditions examined remains a sometimes difficult prerequisite and it has therefore been proposed to use the geometrical mean of multiple reference genes instead of using a single reference gene [9].

As for microarray analysis, qRT-PCR analysis needs to meet up to several qualitative important thresholds. We and others use normally a standard curve for each primer pair which should result by linear regression analysis in a R^2 of >0.99 (preferably: >0.995) and a PCR efficiency of $100 \pm 10\%$. More qRT-PCR details and thresholds are recently put forward to be included in analysis and publications (MIQE) similar to the MIAME compliant strategy for microarray analysis [10].

Initially, qRT-PCR validation of microarray results was an absolute necessity, because microarray analysis was insufficiently robust. Present day microarray platforms in the hands of experienced groups are highly robust and technical validation is more a matter of making sure, than of necessity. Indeed, recent experiments in our lab and others have given identical results in a quantitative and qualitative manner, provided that microarray selected reference genes were used in qRT-PCR. It can be argued that data obtained using microarrays are more dependable, rather than less, compared with qRT-PCR, which is particularly important when small differences in gene expression are considered.

The microarray whole-genome transcriptome profiling technique is a powerful tool to measure differences in mRNA quantity. Differences in mRNA expression are of interest, but are only of relevance if they match to functional biology, and thus are linked to changes in either protein or metabolite levels, or to the predicted physiological changes. Differences at protein level can be confirmed using Western blotting [11] or flowmeter based multiplex analysis [12], which both provide quantitative information. Moreover, immunohistochemistry [13] can be used to provide positional information. However, these methods can be difficult or even impossible to perform when a specific and well-characterized antibody is lacking or when protein concentrations are low. Another technique that can be used is 2-D electrophores coupled to mass spectrometry [14]. Using this technique, proteins are separated by their isoelectric point in the first dimension and by the size of protein in the second dimension. Thereafter proteins are stained and quantified

Table 1. Differential gene expression of glyceraldehyde 3-phosphate dehydrogenase (Gapdh) in adipose tissue of wildtype C57BL/6J male adult mice following nutritional intervention ($n = 12$ *per* group) as analyzed by DNA microarray analysis (for details see text and legend of Fig. 2)

	CR *versus* C		LF *versus* C		E *versus* C	
Gene name	Ratio	*p*-Value	Ratio	*p*-Value	Ratio	*p*-Value
Gapdh	3.32	5.4×10^{-9}	1.46	0.001	−1.02	0.69

C = control, a purified 30en% high fat diet; CR = 30% caloric restriction of C; LF = a purified low fat diet (10en% fat, fat substituted by starch); E = C plus 0.5% EGCg.

using image analysis software. Differentially expressed proteins can be cut from the gel to be identified by MS. This thus allows for simultaneous analysis of many proteins in one sample, which is referred to as proteome analysis. Proteomics has the advantage over transcriptome analysis that it directly assesses at a functional biochemical level. The diverse characteristics of proteins, including variation in size, hydrophobicity, abundancy and secondary, tertiary, or quaternary modifications, imply a smaller window of targets that can be analyzed simultaneously. This precludes replacement of transcriptome analysis for genome wide identification of potential functional effects of micronutrients. This will hold true even if robustness and the window of analysis will be improved by implementing other proteomic approaches that are being developed, which include antibody arrays [15] and chromatographical separation [16]. In many ways proteome analysis is complementary to transcriptome analysis, each targeting a different functional level.

4 Magnitude of micronutrient effects

As indicated, benefit–risk assessment will inherently deal with small functional effects, because studies will try to capture stable differences in homeostatic conditions in a relatively short time frame, rather than waiting for a life time outcome of these changes. To exemplify the magnitude of the effects that can be expected, the number (and magnitude) of genes that changed expression was assessed for a mouse dietary intervention with three different bioactive food components and compared with a strong nutritional intervention (calorie restriction, CR) and a widely examined nutritional intervention (high fat *versus* low fat). The 30% CR intervention resulted in 15 041 genes that were significantly ($p<0.05$) differentially expressed compared with the control high fat diet. The gene expression response of a low fat diet was approximately three times less (5603 regulated genes, $p<0.05$), whereas a diet supplemented with one of three polyphenols again showed a more than three fold lower differential expression using the same cut-off ($p<0.05$).1650 genes changed expression by chronic supplementation with 0.5% epigallocatechingallate (EGCg; E), 1683 by supplementation with an equimolar amount of resveratrol (R) and 977 by supplementation with an equimolar amount of quercetin (Q). This shows that effects of bioactive food components are relatively small (Fig. 2), both in terms of the number of differentially regulated genes (Fig. 2A) and in terms of the fold changes observed (Fig. 2B) when compared with a high *versus* low fat diet and especially when compared with CR. It should be realized that a high *versus* low fat diet intervention already has a mild effect on gene expression when compared with a pharmacological intervention or gene-knockout strategy [17].

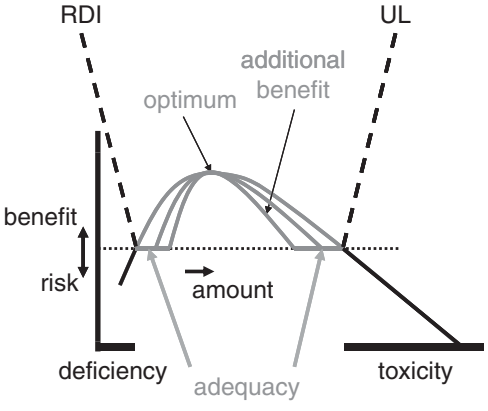

Figure 1. The RDI and the UL border the window of benefit [35]. Deficiency occurs below the RDI and toxicity above the UL. Between these borders the amount of a micronutrient can be adequate to support physiological performance, but additive benefit may occur at certain concentrations or under certain conditions (adapted from [1]).

5 Data-interpretation

To determine statistical differences between average quantitative (log) values *per* expressed probe, different statistical tests can be used depending on the experimental setup and question to be answered. If a pre-defined gene is statistically differentially expressed between two conditions, a Student's *t-test* might be perfectly suitable. In this manner the expression values of a few pre-defined genes can be examined. The data are directly interpreted by the investigator, more or less as an alternative to qRT-PCR with an additional advantage that for the normalization thousands of genes are used and signals are therefore more reliable. However, when testing differences in expression levels of thousands of probe sets, a correction for multiple comparisons is needed. For this, the Benjamini–Hochberg false discovery rate is a frequently used statistical method [18]. But, as stated above, benefit–risk assessment deals mostly with small effects and the use of this very strict statistical method will result in a very rapid decrease in sensitivity, ultimately rendering insufficient information for functional interpretation. This decrease in sensitivity can be circumvented in two ways. First, by increasing the number of individuals or samples. This has practical disadvantages, including costs, time and space limitations. In addition, there may be ethical hesitation when laboratory animals are used. The second approach to increase sensitivity in data analysis is by considering the genes at the biological process or pathway level rather than at the individual gene level. In practice this is done by pre-selecting the genes by a conventional statistical method, the Student's *t-test* or a transcriptome focused modification of this test, such as the Limma *t-test*. Thereafter, these genes are grouped using pathway, network, or gene ontology overrepresentation. To consider genes in the context of a process rather than on an individual basis

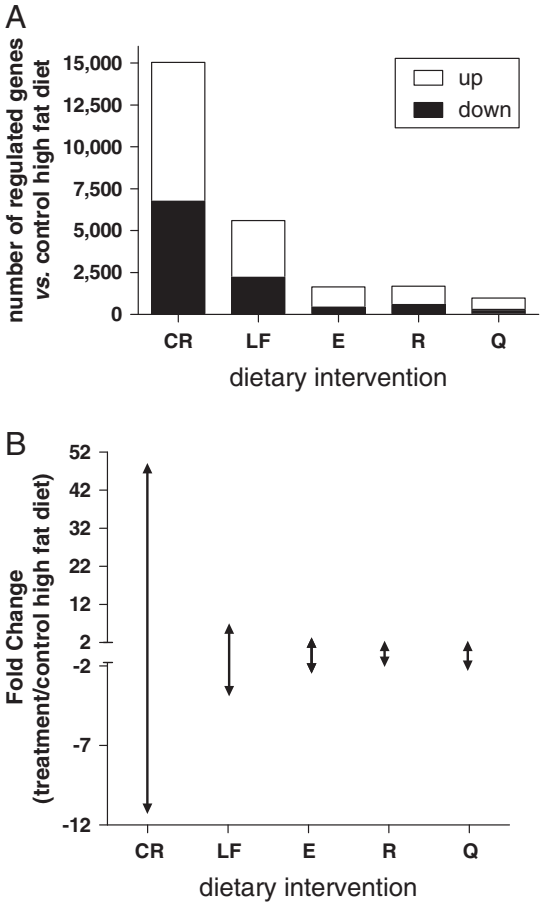

Figure 2. Differential gene regulation by nutritional interventions as analyzed by microarrays. Wildtype C57BL/6J male adult mice ($n = 72$) received a control purified high fat diet (30en% fat) during a 3 wk run-in period, were stratified on body weight, followed by 13 wk dietary intervention with control high fat diet ($n = 12$), 30% caloric restriction (CR, $n = 12$), a purified low fat diet (10en% fat, fat substituted by starch; LF, $n = 12$), a 0.5% EGCg supplemented high fat diet (E, $n = 12$) or equimolar resveratrol (R, $n = 12$) or quercetin (Q, $n = 12$) supplemented high fat diets with dietary groups considered to be independent. The CR diet was adjusted to provide mice with equal micronutrients intake. After the intervention, animals were killed and tissues isolated. All 72 murine epididymal white adipose tissue total RNA samples were analyzed using Agilent 44 K whole-genome microarrays as described in the text. Microarray data were normalized and mean gene expression was calculated *per* dietary group and analyzed *versus* the control high fat diet group. First, the number of significantly upregulated (white bar) and downregulated (black bar) genes was analyzed ($p < 0.05$, panel 2A), and the maximal observed fold change range *per* treatment (treated/control) is shown as double headed arrows in panel 2B (positive numbers: upregulated, negative: downregulated). The animal experiment (DEC2007086) was performed at Wageningen University in The Netherlands fully complying to National and European regulations on animal experimentation. All personnel involved in animal handling have the required experience and certificates.

makes biological sense since an orchestrated regulation of the majority of genes within one pathway seems to have more biological relevance than a single gene within the same pathway. As an example, analysis of rat intestinal infection by Salmonella showed consistent upregulation of chemokines in three repetitions of an oral infection experiment. Although a similar number of chemokines were upregulated in these three consecutive experiments, only a minority of chemokines appeared in all three experiments. These results suggest that a response at the process level is more robust than on individual genes ([19]; unpublished results). Identified pathways have to be investigated subsequently to see whether gene expression changes make biological sense, which is unlikely with false-positive genes.

Another method to group genes is by the use of correlation analysis which is performed using statistical programs that examine the correlation in variation between samples [20]. Correlation analysis takes all genes (above threshold) into account and can thus provide information on non-annotated genes. If their behavior is identical to that of annotated genes, in other words when they group together, they are likely to be functionally correlated [21]. Statistical correlation generally provides limited clues on gene and/or protein function, because this depends on functional knowledge of the genes present within each group.

Although a selection of significantly regulated genes is often used to group genes and to investigate regulated pathways, functional knowledge can also be gained with pathway analysis programs on a selection of, or even the whole data set. These programs are bioinformatical software programs that examine the overrepresentation of certain pathways/networks/gene ontology terms in the data set (public available tools like GSEA, Ermine J, DAVID/EASE, Pathvisio or commercial tools from Ingenuity or MetaCore, among others; [22, 23]). The pathway analysis programs rapidly improve, but still are far from optimal for two main reasons. First, a large number of gene annotations (most) are not recognized neither included in the pathways nor properly translated (*e.g.* murine gene annotation loaded into a human pathway analysis program) [21]. Second, the pathways are built based on existing knowledge, leading to a bias for well studied pathways and a bias for ''classic'' views and questions and against new parts, connections or perspectives. For example, although gastrointestinal hormones can be considered as a common, established denominator of gastrointestinal endocrinology [24], they are not present as an entity in the most frequently used pathway analysis programs. Despite these current drawbacks, pathway analysis programs are an indispensable tool in data interpretation. However, manual inspection of the data and literature mining by the scientist is still required and forms an essential, but time consuming step in the data analysis. It can be envisaged that future analysis will make use of integrated statistical correlation analysis and pathway analysis in a highly structured manner, and a framework of this has been proposed [21].

6 *In vivo* and *in vitro* approaches

Transcriptomics is a potential powerful tool for unbiased identification of functional effects of nutrients. However, human studies are difficult to control, making it more difficult to identify small effects. Differences in behavior and non-nutritional environmental variation add on to difficulties to faithfully monitor consumption patterns as well as consumption of specific food components. The inability to properly monitor the intake of specific food components is in part being tackled by plasma or urine analysis for specific metabolites that can be considered as biomarkers for intake [25]. Another approach is to perform a fully controlled study, but this has its limitations both in terms of duration as well as in terms of number of subjects that can be analyzed. Another disadvantage is that the human population has a high genetic heterogeneity, thus requiring large numbers of study subjects and higher costs. Another important disadvantage in human studies is that only a few tissues can be easily assessed. These are mainly blood, blood cells, urine, saliva and, with more difficulty, muscle, fat and intestinal biopsies. Indeed, peripheral blood mononuclear cells are increasingly used in nutritional transcription profiling, both as a target tissue as well as a surrogate tissue [4, 26, 27].

In vitro approaches, using human immortal cells in culture, can also be used for effect analysis. The main advantages are high-throughput, a well-controlled environment and relatively low costs, as compared with *in vivo* analysis. One major disadvantage is that most cell models are tumor derived or transformed cells that are physiologically compromised. A second disadvantage is that they do not occur in their natural local environment. Furthermore, nutrient metabolism occurring upon ingestion has to be taken into account. For example, most dietary polyphenols are *in vivo* metabolized in the intestinal mucosa, with functional consequences [28]. As a result of metabolism in intestinal mucosa and liver, peripheral tissues are hardly, or not at all, in contact with the unmetabolized polyphenols. Nevertheless, many *in vitro* studies using cell cultures from peripheral origin were performed exposing the cells to these unmetabolized polyphenols, which results in physiologically irrelevant studies. Another, more practical disadvantage is that many physiological and molecular parameters are strongly dependent on the exact cell culture conditions, such as cell density. Transcriptome profiling of *in vitro* dose–response analysis therefore requires multiple replications and even then the results have to be interpreted with extreme caution. Moreover, an *in vitro* approach is unable to detect effects due to interactions between different organs, as occurs *in vivo*. Our own experience trying to confirm new leads that were obtained using *in vitro* models was highly disappointing, making it more logical to first start *in vivo* and then, knowing the validity of the target and the response, proceed using *in vitro* or *ex vivo* systems that show the relevant response.

7 Animal models and diets

Micronutritional benefit–risk studies can also be performed in model animals, with rodents being mostly used. Rodent nutritional studies have the advantage that the environment (temperature, stress), as well as diet and food intake can be highly controlled. Moreover, genetic background can, dependent on the strain, largely be controlled and effects on individual tissues can be assessed in relation to each other. Of importance for benefit–risk assessment is that physiology is largely conserved between rodents and humans [29], although the necessity remains to validate observed effects in humans. An additional advantage of mice is the availability of a large number of strains that lack a specific functional gene (knock-out), either completely, or in a tissue specific or temporal manner. In addition, many transgenic strains exist that express a specific gene at higher level. Both types of strains allow for analysis of the functionality of the gene of interest. For both mice and rats a large number of experimental tools are available, including species-specific gene expression microarrays and many characterized antibodies. Rat and murine genomes are both sequenced [30, 31] and are moderately (rat) to well (mouse) annotated, and many pathways have been investigated, described and are available in pathway analysis tools. Together, this makes these rodents the preferred species of choice for mechanistic studies. In case of a question that focuses on one specific physiological aspect, other species may have a preference over rodents. For example, ferrets display a carotenoid metabolism that is more similar to humans as compared with rodents. In rodents, β-carotene, also known as provitamin A, is almost completely cleaved by carotenoid-monooxygenase in the intestine, while less enzyme activity is observed in the ferret and, like in humans, where intact β-carotene is detected in plasma and peripheral tissues. We have successfully used ferrets to examine effects of β-carotene in the lung [32], but our molecular analyses were frustrating because of a limited availability of annotated ferret sequences (at present, only 311 nucleotide sequences are available in the NCBI database, of which only 111 are annotated as gene (mRNA) which is only a tiny fraction compared with the 265 635 murine mRNA sequences). Rats clearly have the advantage of the larger amount of tissue that is available for analysis and the higher availability of physiological knowledge compared with the mouse. For this and other reasons, rats have become the species of choice in toxicological and physiological studies. Mice were, and still are, the favorite species of molecular geneticists, because knock-out models were earlier and easily available. Also, mice can be maintained at lower costs. Since many nutritional studies focus on disease prevention, nutritionists adopted the mouse as their model of choice following the medical field. As a consequence, mechanistic benefit–risk studies are being performed in rats as well as in mice, often depending on the background of the laboratory involved.

Many different mouse strains are being used. The inbred strain that is most widely used in nutrigenomics is C57BL/6J (''black 6J''). These mice are obesity prone and are able to develop insulin resistance upon a high fat feeding. Its counterpart is the A/J strain, which is resistant to diet-induced obesity. It should be kept in mind that there are more differences between these strains, for example A/J mice are also sensitive to lung disease. Of course many other strains are being used, dependent on the phenotype that has the interest of the scientist. SV129 is generally encountered, because it provides the usual genetic background in which specific gene knock-outs are generated. Most nutritional toxicology and nutritional studies in rats are being performed using the outbred Wistar strain or the inbred Sprague Dawley strain, and to a lesser extent the Fisher or Brown Norway strains.

In nutritional studies using rodents, attention should be paid to the following two aspects. First, the use of purified diets for all animals, including controls. In many studies chow is used as control and high fat diets are either from different origin or are generated by adding fat. In both cases this will result in many differences apart from the amount of fat and carbohydrate. Chow-based diets should be strongly discouraged, since protein and lipid sources may differ from batch to batch and may result in 1000-fold concentration differences of specific nutrients [33], thus masking specific effects of micronutrients or bioactive food components. Therefore, despite higher costs, purified or semi-purified diets of precisely known composition should always be used and described in the materials and methods section. Second, it should be realized that many diets, even purified diets, may contain excessive amounts of certain vitamins and bioactive food components. Generally, vitamin E and β-carotene are present at high levels, mainly to prevent lipid oxidation, which precludes proper analysis of these and related compounds and may strongly affect the results obtained with, for example, other antioxidant compounds. For example, the Vitamin E content in AIN-93G based purified diets is 300 mg/kg, providing mice on average a daily intake of 25 mg/kg body weight, while for humans the RDI is only 0.131 mg/kg body weight. It is therefore important to carefully appraise the diet when designing an experiment.

8 Sensitivity and power

As shown above, functional effects of micronutritional compounds are small and, although well controlled, *in vivo* experiments using rodents have variations close to the nutritional effects. In order to obtain as many statistically significantly regulated genes studies need to be sufficiently powered. To get some insight in the effect of the number of animals used (n) on the number of significantly regulated genes in a micronutritional microarray experiment, we determined the number of significantly regulated genes using a smaller number of animals from the data set described above. This was done by random assignment of the desired number of animals ($n = 4$ up to $n = 11$) out of the group of 12 animals for both the intervention group supplemented with EGCg and the intervention group showing the largest effects (CR) as well as the control group. Since different results can be expected dependent on the specific animals that were randomly assigned to the group, analysis of the number of significantly regulated genes (with a more stringent level of $p < 0.01$) for a certain n was performed 1000 times with randomly drawn groups. As shown in Fig. 3A, the average number of significantly regulated genes of CR *versus* control group dropped from 12 179 ($n = 12$) to just below 6000 regulated genes for $n = 4$. Examining the effect of the supplementation with 0.5% of the food bioactive compound EGCg, the average number of significantly regulated genes ($p < 0.01$) was much lower than upon CR and dropped from 490 ($n = 11$) to 258 ($n = 4$) regulated genes on average with decreasing group size (Fig. 3B). Most striking in these data sets is the large variation in the number of regulated genes when a small number of animals are used (Fig. 3B). This means, for example, that by using $n = 4$ animals 45 regulated genes but also 1500 regulated genes could be identified upon EGCg supplementation, depending on randomly drawn animals in the respective groups. This large variation can result, on the one hand, in a large number of false positives, making it hard to identify ''real'' results, or on the other hand, this can result in an underestimation of the actual number of regulated genes, making it hard to perform pathway analysis to interpret the regulated genes in a systems biology manner. In benefit–risk assessment, often an absolute fold-change is also used besides the p-value as a cut-off value. We therefore re-analyzed the treatments using an absolute fold-change > 1.5 in addition to a $p < 0.01$. As a result, the average number of regulated genes upon CR reaches a plateau at $n > 7$ (Fig. 3C). The average number of regulated genes upon a dietary supplementation with 0.5% EGCg did hardly change for $n = 4$–11 animals, but the degree of variation between each random taken sample dropped dramatically (Fig. 3D). Finally, analyzing the R and Q supplemented groups using the same procedure showed results similar to those of the EGCg supplemented group (data not shown). Altogether these results imply that a high number of false positives or false negatives may be obtained when a low number of animals are being used. The number of animals that should be used depends not only on the criteria that are chosen, such as significance, fold change, minimal number of genes to perform pathway analysis, but also on the effect that is expected. While in the comparison of caloric restriction *versus* control no genes are outside the average number of differentially expressed genes plus or minus 25% when 9 or more animals (in fact: good quality arrays) are used, in the comparison of EGCg supplementation *versus* control still 60% of the genes are outside this range when $n = 9$ animals *per* group are sampled.

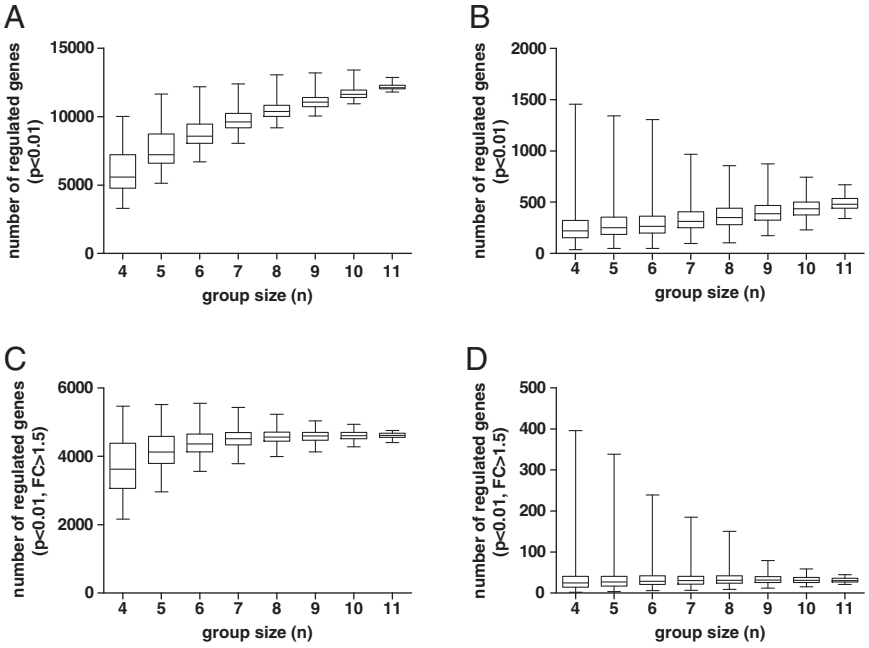

Figure 3. Average number of regulated genes during intervention for different group sizes. Number of regulated genes in a microarray gene-expression experiment of mice fed the control diet (C) compared with mice fed the caloric restriction (CR) or EGCg (E) diet for different group sizes. To obtain a different group size, the group size number of animals was randomly drawn from a group of 12 animals on the control diet and compared with a similar randomly drawn intervention group. This was performed 1000 times *per* group size and the number of significant genes *per* randomly drawn group comparison was calculated and plotted in whiskers box plots (1–99%). A gene was considered significantly different when 3A, $p<0.01$ for C *versus* CR; 3B $p<0.01$ for C *versus* E; 3C, an absolute FC >1.5 and $p<0.01$ for C *versus* CR and 3D, an absolute FC >1.5 and $p<0.01$ for C *versus* E.

9 Concluding remarks

In conclusion, micronutrients and bioactive food components may have functional effects that only become apparent in later life, that is, after chronic differences in intake, or under conditions of stress, affecting fitness and wellbeing. While for intake between RDI and UL no overt differences in health can be expected, these differences may be very important for healthy aging and resistance to disease. They may have particular beneficial effects on the overarching processes; metabolic stress, oxidative stress, inflammatory stress and psychological stress [34]. To be able to establish functional health effects of micronutrients and bioactive food components, an overview of mechanistic effects is needed and transcriptome analysis is highly suited as a first step. When applied correctly, it has the power to detect the relatively small differences in gene expression that are induced by micronutrients and food bioactive compounds. It is however not a miracle technique. It takes experience and hard work to use it properly and to generate true and useful data. But when used sensibly, it is highly powerful and will generate data that will contribute to the establishment of functional effects and mechanisms of action of micronutrients and bioactive food components, a prerequisite for their application in health promoting functional foods.

We thank all group members for their efforts over the past years in development of experimental aspects of the application of transcriptome analysis in nutritional intervention studies. We are members of Mitofood and of NuGO. Y. vH. is funded by the graduate school VLAG.

The authors have declared no conflict of interest.

10 References

[1] Palou, A., Pico, C., Keijer, J., From risk assessment to risk—benefit evaluation: the need for guidelines, concepts and mechanisms. *Crit. Rev. Food Sci. Nutr.* 2009, *49*, 670–680.

[2] Chan, S., Gerson, B., Subramaniam, S., The role of copper, molybdenum, selenium, and zinc in nutrition and health. *Clin. Lab. Med.* 1998, *18*, 673–685.

[3] Koller, L. D., Exon, J. H., The two faces of selenium-deficiency and toxicity--are similar in animals and man. *Can. J. Vet. Res.* 1986, *50*, 297–306.

[4] Pagmantidis, V., Meplan, C., van Schothorst, E. M., Keijer, J., Hesketh, J. E., Supplementation of healthy volunteers with nutritionally relevant amounts of selenium increases the expression of lymphocyte protein biosynthesis genes. *Am. J. Clin. Nutr.* 2008, *87*, 181–189.

[5] Kipp, A., Banning, A., Van Schothorst, E. M., Meplan, C. *et al.*, Four selenoproteins, protein biosynthesis, and Wnt signalling are particularly sensitive to selenium intake in mice colon. *Mol. Nutr. Food Res.* 2009, doi: 10.1002/mnfr.200900105.

[6] de Reynies, A., Geromin, D., Cayuela, J. M., Petel, F. *et al.*, Comparison of the latest commercial short and long oligonucleotide microarray technologies. *BMC Genomics* 2006, *7*, 51.

[7] Pellis, L., Franssen-van Hal, N. L., Burema, J., Keijer, J., The intraclass correlation coefficient applied for evaluation of data correction, labeling methods, and rectal biopsy sampling in DNA microarray experiments. *Physiol. Genomics* 2003, *16*, 99–106.

[8] van Schothorst, E. M., Pagmantidis, V., de Boer, V. C., Hesketh, J., Keijer, J., Assessment of reducing RNA input for Agilent oligo microarrays. *Anal. Biochem.* 2007, *363*, 315–317.

[9] Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B. *et al.*, Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 2002, *3*, RESEARCH0034.

[10] Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J. *et al.*, The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* 2009, *55*, 611–622.

[11] Mathews, S. T., Plaisance, E. P., Kim, T., Imaging systems for westerns: chemiluminescence vs. infrared detection. *Methods Mol. Biol.* 2009, *536*, 499–513.

[12] Djoba Siawaya, J. F., Roberts, T., Babb, C., Black, G. *et al.*, An evaluation of commercial fluorescent bead-based luminex cytokine assays. *PLoS ONE* 2008, *3*, e2535.

[13] Sullivan, C. A., Chung, G. G., Biomarker validation: in situ analysis of protein expression using semiquantitative immunohistochemistry-based techniques. *Clin. Colorectal Cancer* 2008, *7*, 172–177.

[14] Bonk, T., Humeny, A., MALDI-TOF-MS analysis of protein and DNA. *Neurosci.* 2001, *7*, 6–12.

[15] Zichi, D., Eaton, B., Singer, B., Gold, L., Proteomics and diagnostics: Let's Get Specific, again. *Curr. Opin. Chem. Biol.* 2008, *12*, 78–85.

[16] Conrotto, P., Souchelnytskyi, S., Proteomic approaches in biological and medical sciences: principles and applications. *Exp. Oncol.* 2008, *30*, 171–180.

[17] Patsouris, D., Reddy, J. K., Muller, M., Kersten, S., Peroxisome proliferator-activated receptor alpha mediates the effects of high-fat diet on hepatic gene expression. *Endocrinology* 2006, *147*, 1508–1516.

[18] Hochberg, Y., Benjamini, Y., More powerful procedures for multiple significance testing. *Stat. Med.* 1990, *9*, 811–818.

[19] Rodenburg, W., Bovee-Oudenhoven, I. M., Kramer, E., van der Meer, R., Keijer, J., Gene expression response of the rat small intestine following oral S*almonella infection*. *Physiol Genomics* 2007, *30*, 123–133.

[20] Mansson, R., Tsapogas, P., Akerlund, M., Lagergren, A. *et al.*, Pearson correlation analysis of microarray data allows for the identification of genetic targets for early B-cell factor. *J. Biol. Chem.* 2004, *279*, 17905–17913.

[21] Rodenburg, W., Heidema, A. G., Boer, J. M., Bovee-Oudenhoven, I. M. *et al.*, A framework to identify physiological responses in microarray-based gene expression studies: selection and interpretation of biologically relevant genes. *Physiol Genomics* 2008, *33*, 78–90.

[22] Werner, T., Bioinformatics applications for pathway analysis of microarray data. *Curr. Opin. Biotech.* 2008, *19*, 50–54.

[23] Verducci, J. S., Melfi, V. F., Lin, S., Wang, Z. *et al.*, Microarray analysis of gene expression: considerations in data mining and statistical treatment. *Physiol. Genomics* 2006, *25*, 355–363.

[24] Nauck, M. A., Unraveling the science of incretin biology. *Am. J. Med.* 2009, *122*, S3–S10.

[25] Walsh, M. C., Brennan, L., Pujos-Guillot, E., Sebedio, J. L. *et al.*, Influence of acute phytochemical intake on human urinary metabolomic profiles. *Am. J. Clin. Nutr.* 2007, *86*, 1687–1693.

[26] Afman, L., Muller, M., Nutrigenomics: from molecular nutrition to prevention of disease. *J. Am. Diet. Assoc.* 2006, *106*, 569–576.

[27] Eady, J. J., Wortley, G. M., Wormstone, Y. M., Hughes, J. C. *et al.*, Variation in gene expression profiles of peripheral blood mononuclear cells from healthy volunteers. *Physiol. Genomics* 2005, *22*, 402–411.

[28] de Boer, V. C., de Goffau, M. C., Arts, I. C., Hollman, P. C., Keijer, J., SIRT1 stimulation by polyphenols is affected by their stability and metabolism. *Mech. Ageing Dev.* 2006, *127*, 618–627.

[29] Moreno, C., Lazar, J., Jacob, H. J., Kwitek, A. E., Comparative genomics for detecting human disease genes. *Adv. Genet.* 2008, *60*, 655–697.

[30] Waterston, R. H., Lindblad-Toh, K., Birney, E., Rogers, J. *et al.*, Initial sequencing and comparative analysis of the mouse genome. *Nature* 2002, *420*, 520–562.

[31] Gibbs, R. A., Weinstock, G. M., Metzker, M. L., Muzny, D. M. *et al.*, Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature* 2004, *428*, 493–521.

[32] van Helden, Y. G., Keijer, J., Heil, S. G., Pico, C. *et al.*, Beta-carotene affects oxidative stress related DNA damage in lung epithelial cells and in ferret lung. *Carcinogenesis* 2009, doi:10.1093/carcin/bgp186.

[33] Wang, H., Tranguch, S., Xie, H., Hanley, G. *et al.*, Variation in commercial rodent diets induces disparate molecular and physiological changes in the mouse uterus. *Proc. Natl. Acad. Sci. USA.* 2005, *102*, 9960–9965.

[34] van Ommen, B., Keijer, J., Heil, S. G., Kaput, J., Challenging homeostasis to define biomarkers for nutrition related health. *Mol. Nutr. Food Res.* 2009, *53*, 795–804.

[35] Elliott, R., Pico, C., Dommels, Y., Wybranska, I. *et al.*, Nutrigenomic approaches for benefit–risk analysis of foods and food components: defining markers of health. *Br. J. Nutr.* 2007, *98*, 1095–1100.

