

# Diet–Disease Gene Interactions

Jim Kaput, PhD

*From NutraGenomics, Chicago, Illinois, USA*

---

When we consider the matter, we start to see that we cannot finally separate out any phenomena from the context of other phenomena. We can only speak in terms of relationships... The problem of misperception, which, of course, varies in degree, usually arises because of tendency to isolate aspects of an event or experience and see them as constituting its totality. This leads to a narrowing of perspective and from there to false expectation.

Dalai Lama, in *Ethics for the New Millennium*

## INTRODUCTION

Efforts to explain the causes of human diseases often recapitulate the nature versus nurture debate. The successes of identifying genetic and molecular causes of monogenic diseases during the past century and the completion of the human genome project have led many to a “genome-centric” view and approach to the study of chronic diseases: mutations or misregulation of certain genes cause chronic diseases. The rise in the incidence of obesity and type 2 diabetes mellitus (T2DM) over the past 10 to 15 y, however, suggests that exposure to certain environmental factors or agents causes diseases. Although the naturists and nurturists acknowledge the importance of both influences, experimental designs are often discipline specific: either genes or environment are analyzed but not both simultaneously.

A growing number of scientists are advocating and developing integrative approaches to the study of chronic disease and indeed to biology, because neither nature nor nurture alone can explain the molecular processes of human health or disease.<sup>1</sup> Nutritional genomics, or nutrigenomics, is such an integrative science. The working definition of nutrigenomics is that it seeks to provide a genetic and molecular understanding for how common dietary chemicals (i.e., nutrients) affect the balance between health and disease by altering the expression and/or structure of an individual’s genetic makeup.<sup>2</sup> Dietary chemicals include nutrients and bioactive chemicals that do not directly produce energy but exclude man-made chemicals such as pesticides. This new branch of genomic and nutritional research can best be summarized with the following five tenets:<sup>2</sup>

- Common dietary chemicals act on the human genome, directly or indirectly, to alter gene expression or structure.
- Under certain circumstances and in some individuals, diet can be a serious risk factor for a number of diseases.
- Some diet-regulated genes (and their normal, common variants) are susceptibility genes and likely to play a role in the onset, incidence, progression, and/or severity of chronic diseases.
- The degree to which diet influences the balance between healthy and disease states may depend on an individual’s genetic makeup.

---

This review was supported by the National Center for Minority Health and Health Disparities Center of Excellence in Nutritional Genomics (MD00222) and by NutraGenomics.

Correspondence to: Jim Kaput, PhD, NutraGenomics, 2201 West Campbell Park Drive, Chicago, IL 60612, USA. E-mail: jim@nutrigenomics.com

- Dietary intervention based on knowledge of nutritional requirement, nutrition status, and genotype (i.e., “individualized nutrition”) can be used to prevent, mitigate, or cure chronic disease.

I review key concepts that have emerged from epidemiologic, nutritional, molecular, and genetic experiments examining associations between genes and disease. The results and lessons from these different fields of research will affect the design, strategies, and approaches for nutritional genomic research and specifically for identifying diet-regulated and genotype- × diet-regulated genes involved in susceptibility, onset, incidence, progression, and/or severity of chronic diseases.

## IDENTIFYING CANDIDATE GENES: GENETIC AND MOLECULAR APPROACHES

During the previous century, genetic, biochemical, and molecular biology studies identified almost 1000 mutations that cause diseases in humans. Ninety-seven percent of these diseases are caused by monogenic mutations, i.e., a mutation in one gene is sufficient to cause the disease.<sup>3</sup> Positional cloning technologies developed since about 1980 have accelerated progress in identifying genes that cause these relatively rare diseases.<sup>4,5</sup> Genes involved in chronic diseases such as T2DM, obesity, cardiovascular diseases (CVD), the metabolic syndrome, cancers, and others, however, have not been unequivocally linked to causation of these diseases by using these genetic approaches. More than 600 studies examining associations of polymorphisms (single nucleotide polymorphisms [SNPs] and other types) in candidate genes with disease or with a subphenotype of a chronic disease have been published as of 2002.<sup>6</sup> Only six gene–disease associations were replicated in more than three studies, even though some of the candidate genes produce disease when mutations ablate or overexpress RNA or protein.

Data from quantitative trait loci (QTL) mapping<sup>7,8</sup> illustrates one of the reasons that candidate gene–disease association studies are not replicated: most cases of chronic diseases are not caused by mutations in single genes but rather are due to complex interactions among variants of several to many genes (Figure 1A). QTL mapping can be done in humans, but the process is best explained with laboratory animals. Briefly, two parental inbred mouse strains, selected for differences in some observable or measurable phenotype, are bred to produce an F1 generation. F1 mice are backcrossed to each of the parental strains, thereby producing an F2 generation differing in disease susceptibility because chromosomal rearrangements occur during meiosis. The incidence of disease or severity of subphenotypes of the disease are measured in F2 mice and statistically associated with chromosomal regions from each of the original parental strains. A given pair of inbred mice may have 10 to 15 regions that contribute to the complex phenotype in those strains, and different pairs of strains may reveal new QTLs.<sup>9</sup> Seventy-five QTLs for obesity and 85 for body weight in mice<sup>10</sup> were known as of late 2002. QTLs from different studies overlap, suggesting that some, but not all, of the mechanisms for obesity may be shared. Obesity is not an unique case, because there are many QTL for T2DM, CVDs, lipid levels, insulin levels,

etc. QTLs for many complex traits in mice are found at <http://www.jax.org>.

If one or more of the genes within the QTL regions were mutated, each parental strain would have an increased risk to develop the particular disease. This does not occur. Rather, the sum total of the minor contributions from causative genes within different QTLs produce the specific trait or disease. Two hypotheses were proposed to explain gene variants within QTLs. The common disease/common variant hypothesis<sup>4,5</sup> posits that combinations of normally occurring gene variants produce disease. These gene variants occur in greater than 1% of the population. Others dismiss this theory, suggesting that combinations of rare (<1% in the population) allelic variants cause common diseases. This theory is called the multilocus/multiallele hypothesis.<sup>11</sup>

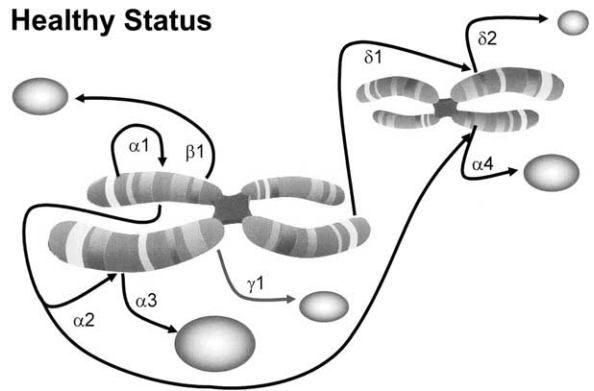
## DIETARY CHEMICALS INFLUENCE EXPRESSION OF GENETIC INFORMATION

Regardless of the outcome of this controversy, many gene-association, QTL studies, and genetic laboratory animal studies neglect to directly test an important variable in the expression of genetic information and a major contributor to disease development, namely the influence of diet on the expression of genetic information. Although many chemicals in foods are nutrients, i.e., are metabolized to energy or involved in key metabolic reactions (e.g., vitamins), some naturally occurring chemicals in foods are ligands for transcription factors and directly alter gene expression, whereas other dietary chemicals alter signal transduction pathways and chromatin structure to indirectly affect gene expression.<sup>2,12</sup> Epidemiologic studies have repeatedly shown that intake of different diets are associated with the incidence and severity of chronic diseases.<sup>13,14</sup> Overconsumption of energy,<sup>14</sup> proteins, types of fats or carbohydrates,<sup>15–17</sup> or lack of key micronutrients<sup>18</sup> are associated with obesity, T2DM, CVD, certain cancers, developmental defects, and neurologic diseases such as Alzheimer's.

Association studies are analyzing diet or responses to dietary changes with SNPs in candidate genes. The candidate genes are usually selected based on known physiology or biochemistry. Variants in these candidate genes are assumed to alter the risk of developing the disease and therefore are considered susceptibility genes. Dietary chemicals may preferentially interact with one or more variants (i.e., susceptibility genes) to increase or decrease disease risk (Figure 1B). Table I lists the candidate genes analyzed for association with response to different diets in about 100 studies. However, the results of these studies duplicate those from the more than 600 gene–disease association studies<sup>6</sup>: few, if any, of the diet–SNP associations are replicated among studies. The same limitations that apply to candidate gene–disease association studies apply to candidate gene–diet studies<sup>19–21</sup>: sample sizes that lack appropriate statistical power, control groups that are not appropriately matched to cases, population stratification that occurs because of genetic admixtures among the study participants, and overinterpreting data (among others).<sup>22–25</sup>

Other variables also may influence candidate gene–disease association studies<sup>1,11,19–21</sup>: 1) the diversity of molecular processes that may produce chronic disease and the consequences of selecting one or a few candidate genes, 2) the physiologic response to the presence of a disease may alter expression of genetic information, 3) genotype  $\times$  environment interactions, and 4) the use of model organisms in the search for candidate genes. These variables illustrate concepts that are fundamental to the field of nutritional genomics and the design of human and laboratory animal studies.

### A) Healthy Status



### B) Unbalanced Nutrition

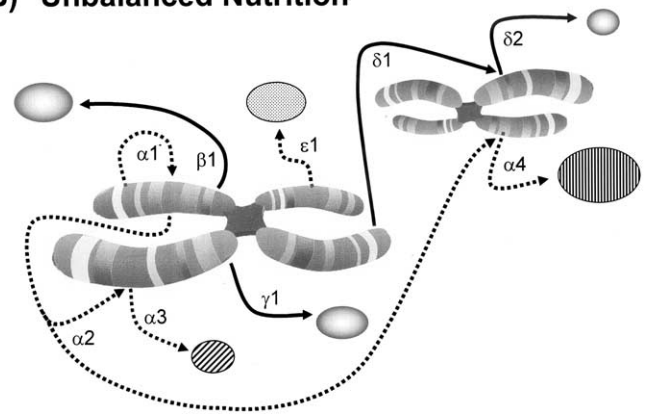


FIG. 1. Genotype  $\times$  environment interactions. Example of gene expression in healthy (A) and unbalanced nutritional (B) states. (A) Gene  $\alpha 1$  is a global transcription factor regulating  $\alpha 2$  (and other transcription) factors. In turn,  $\alpha 2$  regulates  $\alpha 3$  and  $\alpha 4$  expressions. Another transcriptional regulator,  $\delta 1$ , affects  $\delta 2$  expression. Genes encoding  $\beta 1$ ,  $\gamma 1$ ,  $\delta 1$ , and  $\delta 2$  are not regulated by transcription factors influenced by dietary chemicals. (B) Unbalanced nutrition alters the expression (---) of  $\alpha 1$ , ultimately decreasing the amount of  $\alpha 3$  (●) but increasing the expression of  $\alpha 4$  (⊞). Gene  $\epsilon 1$  is expressed in response to changes in metabolism or the altered concentration of a dietary ligand (⊙). Some transcription factors or expression of individual genes may not be affected directly or indirectly by dietary chemicals (e.g.,  $\beta 1$ ,  $\gamma 1$  and  $\delta 1$ ,  $\delta 2$ ), whereas others ( $\alpha 1$  through  $\alpha 4$  and  $\epsilon 1$ ) may be affected. The effects on transcription may result from increased concentrations of ligand derived from the diet acting on a reference or variant (single nucleotide polymorphism) gene sequence. Changing the concentration of effector proteins ( $\alpha 3$ ,  $\alpha 4$ ,  $\epsilon 1$ ) will alter metabolic flux and concentrations of metabolites, inducing further changes in cell physiology.

## CHRONIC DISEASES MAY RESULT FROM DIFFERENT MOLECULAR PATHWAYS

Many experimental designs assume that all cases of a chronic disease result from the same molecular pathway, i.e., all cases of diabetes, CVD, obesity, or other chronic diseases result from mutations in, or misregulation of, only one small set of genes. Genetic and molecular studies over the past 30 y have established that several different pathways may cause a disease. The simplification that all cases of chronic disease occur by the same pathway will confound statistical association studies and limit applicability of results from laboratory animal experiments that rely on a single strain. To illustrate the point, consider oncogenes. Cancer genes are involved in the mechanism of initiating and/or promoting uncontrolled growth, and they usually play a role in regulatory processes involved in growth control. Almost 1700 genes (<http://caroll.vjf.cnrs.fr/cancergene/HOME.html>) fit the description of oncogene, although many may act only in specific tissues

TABLE I.

## GENES ANALYZED IN RESPONSE TO DIET OR IN DIET ASSOCIATION STUDIES

Challenge/association	Gene	OMIM	Reference
Lipid, diet, smoking, sex	Apolipoprotein-A1	107680	21, 43, 45
Lipid, diet, sex	Apolipoprotein-A4	107690	21, 43–45
Lipid, diet, sex	Apolipoprotein-B	107730	21, 43–45
Lipid, diet, smoking, sex	Apolipoprotein-C3	107720	21, 43–45
Lipid, diet, smoking, activity	Apolipoprotein-E	107741	21, 43–45
Smoking	Apolipoprotein-H	138700	21
Lipid, diet, smoking, activity	Lipoprotein lipase	238600	21, 43–45
Lipid, diet, smoking, alcohol, sex	Cholesterol ester transfer protein	118470	21, 43–45
Lipid	Lecithin:cholesterol acyltransferase	606967	45
Lipid, diet	LDL receptor	606945	43, 45
Diet	Hepatic lipase	151670	43, 44
Diet	Cholesterol 7 $\alpha$ -hydroxylase	118455	43
Diet	Intestinal fatty acid-binding protein	600422	43, 44
Diet	Neuropeptidase Y	162640	43
Diet	M/N blood group	111300	43
Alcohol	Alcohol dehydrogenase-3	103730	21
Smoking, activity	Paraoxonase	168820	21
Diet	Microsomal transfer protein	157147	44

LDL, low-density lipoprotein

or cells. Changes in expression or activity levels occur by mutations, DNA rearrangements, and constitutive activation and ultimately may be caused by genomic instability.<sup>26</sup>

Different tumors may be initiated and/or promoted by one or more different oncogenes, each activating its own cascade of altered regulatory processes. Cancers of the same organ or cell type that appear to be morphologically and histologically similar may have unique molecular expression profiles. Glioblastoma provides one example. Data from high-density oligonucleotide arrays showed that epidermal growth factor receptor positive tumors express 90 genes differently from epidermal growth factor receptor negative tumors.<sup>27</sup> Hierarchical cluster analyses also found two additional subtypes of epidermal growth factor receptor negative glioblastoma. Other cancers show similar heterogeneity, e.g., there are three types of classical Hodgkin's disease tumors.<sup>28</sup>

CVD, obesity, diabetes, and other chronic diseases are also caused by multiple and diverse molecular pathways. Mutations in many different genes such as leptin (*Lep* or *ob*), leptin receptor (*Lep* or *db*), agouti signaling peptide (*A<sup>y</sup>*), attractin (*Atrn*), insulin signaling protein (*Tub*), and carboxypeptidase-E (*Cpe*) cause obesity in mice and humans.<sup>10</sup> Comparison of mice deficient in leptin (*Lep<sup>ob/Lepob</sup>*) with mice showing ectopic expression of agouti (*A<sup>y/y-</sup>*) demonstrated that *Lep<sup>ob/Lepob</sup>* mice have smaller lean body mass, are infertile or sterile, and have many other morphologic and physiologic characteristics that are categorically opposite to each other.<sup>29</sup> These single gene mutations produce the same physical appearance of obesity, but by different molecular pathways. Physiologic or molecular comparisons of other diseases such as CVD show similar molecular heterogeneity.<sup>21</sup>

## PRESENCE OF DISEASE MAY ALTER EXPRESSION OF GENETIC INFORMATION

The presence of a disease may also influence expression of genetic information differently depending on genetic background or the presence of a disease phenotype. Stoehr and coworkers<sup>30</sup> introduced the *ob* (mutant leptin) allele into euglycemic and glucose-tolerant, lean BTBR mice by marker-assisted backcrosses with C57BL/6J-*ob/ob* mice. BTBR-*ob* mice were diabetic with espe-

cially severe hyperglycemia relative to B6-*ob* mice. The phenotypic expression of a diabetic subphenotype differs among genotype (BTBR-*ob/ob* versus C57BL/6J-*ob/ob*) and the presence of obesity (BTBR versus BTBR-*ob/ob*). Genes involved in inflammation in adipose tissue and hepatic lipogenic enzymes were expressed more abundantly in obese diabetic (BTBR-*ob/ob*) mice than in obese-non-diabetic mice (C57BL/6J-*ob/ob*), and lipogenic genes in adipose tissue and hepatic gluconeogenic genes were suppressed in diabetic mice.<sup>31</sup> The increased expression of lipogenic genes in liver may protect obese C57BL/6J mice from diabetes development. The implications for human studies are important: not only must individual genotypes be assessed, the health of the individual (e.g., body mass index, serum lipid levels, glucose, and insulin levels) must also be included in analyses of the data for refining association analyses or interpreting data from laboratory animals.

## GENOTYPE $\times$ ENVIRONMENT INTERACTIONS

The concept of gene  $\times$  environment interactions is not new to nutrigenomics,<sup>32</sup> but its definition and use are not always consistent. The precise, statistical definition of gene–environment interaction is “a different effect of an environmental exposure on disease risk in persons with different genotypes” or “a different effect of a genotype on disease risk in persons with different environmental exposures.”<sup>33</sup> Many association and molecular studies have assumed that SNP or haplotype profile is associated with the same phenotype in all individuals. Such an association may be true if the polymorphism produces a major phenotypic effect as occurs in phenylketonuria. Phenylketonuria is a defect in the enzyme phenylalanine hydroxylase that results in an accumulation of phenylalanine in the blood (<http://www.ncbi.nlm.nih.gov/htbin-post/Omim/dispim?261600>), which often results in neurologic damage. Phenylketonuria can be managed with diets low in phenylalanine.

However, SNP or haplotype may be expressed differently within different genotypes. This is best illustrated with data from laboratory animals. C57BL/6J and BTBR parental strains exhibit normal insulin levels and glucose responses, but C57BL/6J  $\times$

BTBR F1 males have elevated fasting insulin levels, are insulin resistant with impaired oral glucose tolerance, and have attenuated insulin-stimulated glucose uptake into muscle tissue.<sup>34</sup> These results may be explained by epistatic or gene–gene interactions between previously silent alleles in B6 and BTBR mice. Expression of these silent alleles altered expression of other genes involved in insulin regulation in F1 mice. Recombinant inbred and F1 mice produced by mating non-diabetic SM/J and A/J also showed symptoms of diabetes.<sup>35</sup>

The molecular explanation for these genetic results is that proteins or enzymes produced by a gene or its variant do not act alone but are usually a part of a pathway, and many pathways are interconnected. A decrease in activity of one member of a pathway may be compensated for by another member of the same pathway or by variations in a connected pathway. Compensation in the activity of parts to maintain the overall balance within the system is called *buffering*.<sup>36,37</sup> Two other uncomplicated examples of compensation are that all autosomal genes in humans come in pairs, one on each chromosome, and both may be expressed, and some genes are duplicated in the genome to produce proteins with overlapping activities.<sup>36</sup>

A dietary chemical may preferentially alter the expression of a susceptibility gene or its variant that in turn affects various gene–gene interactions. The consequence of such gene  $\times$  environment interactions not only may be apparent on the gene of interest but also may affect the action of that gene on other interacting genes. Different functional classes of genes may have greater importance than others: affecting the expression of a nuclear receptor may affect more processes than altering the expression of a gene encoding an enzyme in a metabolic pathway (Figure 1). An example is the phytosterol genistein, which, in addition to inhibiting certain tyrosine kinases, binds to estrogen receptors,<sup>38</sup> and both reactions will affect the expression of different genes. The concentration of other transcriptional ligands is controlled by their *in vivo* synthesis from a dietary chemical precursor.<sup>39</sup> Steroids, for example, are produced through a core of 10 linked reactions from cholesterol. The concentration of any given steroid ligand<sup>40</sup> will be greatly influenced by specific combinations of alleles for the enzymatic steps in biosynthetic, branch, and degradative pathways. Alleles of these interacting and linked genes may be different among individuals, thus creating differences in the levels of steroids and therefore expression of genes that these steroids regulate. Analyzing an SNP or set of SNPs in one gene of this pathway is unlikely to provide a reliable association with the amount of the end product.

Epigenetic processes, such as DNA methylation, also may cause gene expression changes but without altering gene sequences.<sup>41</sup> DNA methylation plays an important role in genetic imprinting. Cooney and colleagues<sup>42</sup> showed that methyl supplements fed to pregnant mice increase the level of DNA methylation in the agouti long terminal repeat (LTR) in offspring, suggesting that dietary effects on gene expression may begin *in utero*. Increased DNA methylation is predicted to decrease gene expression by altering chromatin structure.

The effects of complex mixtures of dietary chemicals on gene expression and epigenetic processes therefore may explain some of the variability observed in human candidate gene association studies. Qualitative and quantitative differences in dietary chemical intake among free-living organisms such as humans are difficult to assess or measure accurately.

## MODEL SYSTEMS AND THE SEARCH FOR NON-OBVIOUS CANDIDATE GENES

Gene–diet interactions have been found in humans.<sup>21,43–45</sup> As one specific example, Krauss<sup>46</sup> relied on phenotype to show genotypic differences. Individuals with small, dense, low-density lipoprotein

particles (phenotype B) have an increased risk of coronary artery disease as compared with those individuals exhibiting large, less dense, low-density lipoprotein particles (phenotype A). The expression of phenotype A depended on diet: 12 of 38 men who switched from a 32% fat diet to a diet containing 10% fat developed the phenotype B pattern.<sup>47</sup> At least three distinct genotypes were present in this small group, one each for the A or B phenotype and a third that was responsive to low-fat, high-carbohydrate diets. This genotype produced the A phenotype when these individuals ate a diet containing 32% fat, but a B phenotype when fed the 10% fat, a result that can be explained by genotype  $\times$  environment interactions.

Humans, however, are not good model organisms because of the genetic variation among individuals, their long lifespan, and difficulty in controlling and monitoring dietary intakes. An alternative approach to the study of diet–gene interactions is to use laboratory animals in which one can control (or choose) the genotype and systematically alter the environmental factors. Many genetic studies using laboratory animals employ only one diet in their studies, thereby reducing the variables that might confound the results. Alternatively, nutritionists typically employ outbred strains or one genotype and change the diet. However, because many cases of chronic diseases are influenced by different diets, environment  $\times$  genotype interactions will not be found unless diet and genotype are controlled and changed in the experimental design.

Inbred strains of mice provide unambiguous evidence of the importance of genotype  $\times$  diet interactions. Nishina and coworkers showed that different inbred mouse strains differ in susceptibility to diet-induced atherosclerosis.<sup>48</sup> The difference in response between a strain that developed atherosclerotic lesions (e.g., B57BL/6J mice) and one that did not (e.g., BALB/cJ) was the effect of cholic acid on high-density lipoprotein cholesterol levels.<sup>49</sup> In addition, cholic acid reduced expression of 7 $\alpha$ -hydroxylase and lecithin:cholesterol acyltransferase activity in B6 mice but not in BALB/c mice. Hence, strains with normal metabolism on a non-atherogenic diet were differentiated by a dietary challenge.

I and my colleagues developed, tested, and use a strategy that changes diet and genotype to identify diet versus genotype versus diet  $\times$  genotype-regulated genes.<sup>50–54</sup> We recently analyzed differences in gene expression in livers of *A/a* versus *A<sup>vy/A</sup>* segregants (i.e., littermates) of BALB/cStCrlfC3H/Nctr  $\times$  VYwffC3Hf/Nctr-*A<sup>vy/a</sup>* matings in response to *ad libitum* and calorically restricted (70% of *ad libitum*) diets (Kaput et al., unpublished observations). The *A/a* (agouti) versus *A<sup>vy/A</sup>* genotypes are identical except for ectopic expression of the agouti gene in *A<sup>vy/A</sup>* mice, which causes obesity, subphenotypes of diabetes, hyperphagia, and susceptibility to cancer in certain tissues. Gene expression in livers was measured with arrays of at least 18 000 cDNAs or expressed sequence tags. Of the 323 genes with known functions found to be reproducibly and differentially regulated across the four groups of mice (Table II), 4.7% were regulated by genotype regardless of the diet, 6.2% were regulated by differences in caloric intake regardless of the genotype, and the remaining genes were differently regulated based on diet and genotype. Table II presents examples of genes regulated by diet only (N-Ras), by genotype (cyclin MCS2), and by genotype  $\times$  diet interactions (Jak2). Diet  $\times$  genotype interactions are highly complex and difficult to predict, thus demonstrating the need for highly controlled genotypes and environmental conditions that allow for identifying different regulatory patterns based on diet and genotype.

As a means to identify genes that may be causal for chronic diseases, we compared the chromosomal position of genes differentially expressed based on diet, genotype, and genotype  $\times$  diet interactions with QTLs for diabetes and obesity. Thirty-two genes overlapped diabetic QTLs. These candidate genes belonged to different functional classes: 10 were enzymes in several pathways, 7 encoded genes involved in cellular structures, 5 were involved in signal transduction, 5 were transcription, replication, or splicing factors, and 1 was an immune response gene. Literature searches

TABLE II.

EXAMPLES OF DIET-GENE INTERACTIONS*					
Gene	Diet		Genotype		Regulated by
	Aal:Acr	Yal:Ycr	Aal:Yal	Acr:Ycr	
Cyclin MCS2			0.4	0.3	Genotype only
N-Ras	2.7	3.2			Diet only
CYP27 (sterol hydroxylase)	0.4				Diet in one genotype
Interleukin-6 receptor				0.4	Genotype in one diet
Janus kinase-2	0.1			10.4	Genotype × environment
Adenosine triphosphate synthase $\beta$ -chain	6.7		3.4	0.2	Genotype × environment

\* Examples of regulation by diet, genotype, and genotype × environment.

Gene expression was measured by phosphorimager analyses of separate filters hybridized with cDNA synthesized in the presence of  $^{33}\text{P}$ -dNTP to cDNA arrays (18 000 mouse genes and essential sequence tags). These genes were expressed in all comparisons, but not at levels of  $\geq 2.5$  or  $\leq 1/2.5$ , levels which were used in the analyses due to the experimental design and technology limitations. Two mice from each group were used and data were analyzed as in Kaput et al. (unpublished results).

A, agouti (*A/a*) mice; Aal:Acr, ratio of expression of the gene in livers from agouti mice fed 100% calories divided by its expression in agouti mice fed 70% of calories; cr, calorie restriction (70% of calories); Y, obese yellow (*A<sup>y</sup>/A*) mice fed ad libitum.

showed that many, but not all, of the genes obtained by this unbiased, stepwise screening procedure were analyzed in models of diabetes or were in pathways known to be affected in diabetic humans or laboratory animals. Our hypothesis was that genes regulated by diet and overlapping QTLs for diseases or phenotypes influenced by diet are better candidate genes than those that are differentially expressed or found within the QTL.

Others have developed a strategy that directly combines gene expression with quantitative trait loci analyses.<sup>55</sup> In brief, RNA from tissues from high-fat-fed C57BL/6 and DBA mice and a subset of their F2 offspring were analyzed with oligonucleotide arrays, and the expression results were treated as quantitative traits. The pattern of expressed QTLs (eQTLs) and the genes within them identified new candidate genes for obesity. This approach links genetic mapping methods with gene expression data and will be a useful tool for identifying gene-gene interactions involved in complex traits. The limitation of this method for nutrigenomics is that each F2 mouse is a genetically unique individual ( $n = 1$ ), and the effects of diet can be compared only with crossover experiments in easily regenerated cells (e.g., white blood cells).

## PROSPECTIVE

The lessons from genetic association studies in humans, laboratory animal experiments, and these early nutrigenomic studies are that identifying genes that cause chronic disease will be challenging because of the complexity of genotypes, diets, and their interactions. These obstacles are not insurmountable but may require a nutrigenomics project on the scale of the human genome project to find genes that cause chronic diseases and the nutrients that regulate or influence their activity. Such a large-scale project could be taken in multiple steps:

1. Identification of genes regulated by diet, genotype, and gene × environment interactions in multiple strains of laboratory animals. The large number of inbred mouse strains (e.g., Jackson Laboratory, <http://www.jax.org>) that are phenotypically characterized (<http://aretha.jax.org/pub/cgi/phenome/mpdcgi?rtn=docs/home>) provides the ability to compare among susceptible and non-susceptible genotypes for most chronic diseases. Outcrossing to wild mice may provide a means to overcome the limited genetic diversity within inbred strains. Defined diets must be used to ensure repro-

ducibility across experiments. Comparative genomic methods being developed in academia and private industry will provide the tools to compare results across species.

2. Identify and validate SNPs in many genes involved in metabolism and those that regulate their expression and activity. The Environmental Genome Project is developing such information for 200 genes that respond to xenobiotics. A similar project is needed for nutritional genomics research efforts.
3. Large-scale human studies. Due to the complexity of genotypes within the population and the cost of assessing haplotypes within individuals, perhaps the best design for human studies would be standard crossover designs to minimize genotypic variation between individuals.<sup>47</sup> This approach nevertheless will require a large sample size and controlled or monitored dietary intakes. Overall physical condition also must be monitored because disease may influence gene expression independently of diet.

Although the complexities of chronic disease research are significant, nutrigenomic approaches offer the best hope for understanding the molecular processes that maintain health and prevent disease development.

## ACKNOWLEDGMENTS

The author thanks Willard Visek, Ray Rodriguez, and Karin Klein for reviewing the manuscript and providing comments and suggestions. Dr. Rodriguez provided the first draft of Figure 1.

## REFERENCES

1. Sing CF, Stengard JH, Kardia SL. Genes, environment and cardiovascular disease. *Arterioscler Thromb Vasc Biol* 2003;23:1190
2. Kaput J, Rodriguez R. Nutritional genomics. The next frontier in the post-genomic era. *Physiol Genomics* 2003(in press)
3. Jimenez-Sanchez G, Childs B, Valle D. Human disease genes. *Nature* 2001; 409(6822):853
4. Collins FS, Guyer MS, Charkravarti A. Variations on a theme: cataloging human DNA sequence variation. *Science* 1997;278(5343):1580
5. Lander ES. The new genomics. Global views of biology (see comments). *Science* 1996;274(5287):536
6. Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K. A comprehensive review of genetic association studies. *Genet Med* 2002;4:45

7. Lander ES, Botstein D. Mapping mendelian factors underlying quantitative traits using rflp linkage maps. *Genetics* 1989;121:185
8. Paterson AH, Hewitt JD, Zamir D, et al. Mendelian factors underlying quantitative traits in tomato: comparison across species, generations, and environments. *Genetics* 1991;127:181
9. Brockmann GA, Bevova MR. Using mouse models to dissect the genetics of obesity. *Trends Genet* 2002;18:367
10. Chagnon YC, Rankinen T, Snyder EE, Weisnagel SJ, Perusse L, Bouchard C. The human obesity gene map: the 2002 update. *Obes Res* 2003;11:313
11. Wright AF, Hastie ND. Complex genetic diseases. Controversy over the croesus code. *Genome Biol* 2001;2:COMMENT2007
12. Muller M, Kersten S. Opinion. Nutrigenomics: goals and strategies. *Nat Rev Genet* 2003;4:315
13. Jenkins DJA, Kendall CWC, Ransom TPP. Dietary fiber, the evolution of the human diet and coronary heart disease. *Nutr Res* 1998;18:633
14. Willett W. Isocaloric diets are of primary interest in experimental and epidemiological studies. *Int J Epidemiol* 2002;31:694
15. Jenkins DJ, Kendall CW, Augustin LS, et al. Glycemic index. Overview of implications in health and disease. *Am J Clin Nutr* 2002;76:266S
16. Krauss RM. Heterogeneity of plasma low-density lipoproteins and atherosclerosis risk. *Curr Opin Lipidol* 1994;5:339
17. Willett WC. Balancing life-style and genomics research for disease prevention. *Science* 2002;296(5568):695
18. Fairfield KM, Fletcher RH. Vitamins for chronic disease prevention in adults: scientific review. *JAMA* 2002;287:3116
19. Abumrad NA. The gene-nutrient-gene loop. *Curr Opin Clin Nutr Metab Care* 2001;4:407
20. Bonaiti-Pellie C. Methodological aspects of investigating gene-nutrient interactions. *Eur J Cancer Prev* 2002;11:69
21. Ordovas JM. Hdl genetics. Candidate genes, genome wide scans and gene-environment interactions. *Cardiovasc Drugs Ther* 2002;16:273
22. Cardon LR, Bell JI. Association study designs for complex diseases. *Nat Rev Genet* 2001;2:91
23. Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results (see comments). *Nat Genet* 1995;11:241
24. Risch N. Evolving methods in genetic epidemiology. II. Genetic linkage from an epidemiologic perspective. *Epidemiol Rev* 1997;19:24
25. Tabor HK, Risch NJ, Myers RM. Opinion. Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nat Rev Genet* 2002;3:391
26. Balmain A, Gray J, Ponder B. The genetics and genomics of cancer. *Nat Genet* 2003;33(suppl):238
27. Mischel PS, Shai R, Shi T, et al. Identification of molecular subtypes of glioblastoma by gene expression profiling. *Oncogene* 2003;22:2361
28. Devillard E, Bertucci F, Trempt P, et al. Gene expression profiling defines molecular subtypes of classical Hodgkin's disease. *Oncogene* 2002;21:3095
29. Wolff GL. Obesity as a pleiotropic effect of gene action. *J Nutr* 1997;127:1897S
30. Stoehr JP, Nadler ST, Schueler KL, et al. Genetic obesity unmasks nonlinear interactions between murine type 2 diabetes susceptibility loci. *Diabetes* 2000;49:1946
31. Lan H, Rabaglia ME, Stoehr JP, et al. Gene expression profiles of nondiabetic and diabetic obese mice suggest a role of hepatic lipogenic capacity in diabetes susceptibility. *Diabetes* 2003;52:688
32. Young VR, Scrimshaw NS. Genetic and biological variability in human nutrient requirements. *Am J Clin Nutr* 1979;32:486
33. Ottman R. Gene-environment interaction. Definitions and study designs. *Prev Med* 1996;25:764
34. Ranheim T, Dumke C, Schueler KL, Cartee GD, Attie AD. Interaction between *btrb* and *c57bl/6j* genomes produces an insulin resistance syndrome in (*btrb* × *c57bl/6j*) *f1* mice. *Arterioscler Thromb Vasc Biol* 1997;17:3286
35. Kobayashi M, Ohno T, Tsuji A, Nishimura M, Horio F. Combinations of nondiabetic parental genomes elicit impaired glucose tolerance in mouse *smxa* recombinant inbred strains. *Diabetes* 2003;52:180
36. Caporaso NE. Why have we failed to find the low penetrance genetic constituents of common cancers? *Cancer Epidemiol Biomarkers Prev* 2002;11:1544
37. Hartman JLT, Garvik B, Hartwell L. Principles for the buffering of genetic variation. *Science* 2001;291(5506):1001
38. Dixon RA, Ferreira D. Genistein. *Phytochemistry* 2002;60:205
39. Francis GA, Fayard E, Picard F, Auwerx J. Nuclear receptors and the control of metabolism. *Annu Rev Physiol* 2002;65:261
40. Nobel S, Abrahamson L, Oppermann U. Metabolic conversion as a pre-receptor control mechanism for lipophilic hormones. *Eur J Biochem* 2001;268:4113
41. Riddihough G, Pennis E. The evolution of epigenetics. *Science* 2001;293(5532):1063
42. Cooney CA, Dave AA, Wolff GL. Maternal methyl supplements in mice affect epigenetic variation and DNA methylation of offspring. *J Nutr* 2002;132(suppl):2393S
43. Masson LF, McNeill G, Avenell A. Genetic variation and the lipid response to dietary intervention: a systematic review. *Am J Clin Nutr* 2003;77:1098
44. Vincent S, Planells R, Defoort C, et al. Genetic polymorphisms and lipoprotein responses to diets. *Proc Nutr Soc* 2002;61:427
45. Ye SQ, Kwieterovich PO, Jr. Influence of genetic polymorphisms on responsiveness to dietary fat and cholesterol. *Am J Clin Nutr* 2000;72(suppl):1275S
46. Krauss RM. Dietary and genetic effects on LDL heterogeneity. *World Review of Nutrition and Dietetics* 2001;89:12
47. Dreon DM, Fernstrom HA, Williams PT, Krauss RM. A very-low-fat diet is not associated with improved lipoprotein profiles in men with a predominance of large, low-density lipoproteins. *Am J Clin Nutr* 1999;69:411
48. Nishina PM, Wang J, Toyofuku W, Kuypers FA, Ishida BY, Paigen B. Atherosclerosis, and plasma, and liver lipids in nine inbred strains of mice. *Lipids* 1993;28:599
49. Dueland S, France D, Wang SL, Trawick JD, Davis RA. Cholesterol 7 $\alpha$ -hydroxylase influences the expression of hepatic *apoA-I* in two inbred mouse strains displaying different susceptibilities to atherosclerosis and in hepatoma cells. *J Lipid Res* 1997;38:1445
50. Elliott TS, Swartz DA, Paisley EA, Mangian HJ, Visek WJ, Kaput J. F1f0-*atpase* subunit *c* gene isolated in a screen for diet regulated genes. *Biochem Biophys Res Commun* 1993;190:167
51. Kaput J, Swartz D, Paisley E, Mangian H, Daniel WL, Visek WJ. Diet-disease interactions at the molecular level: an experimental paradigm. *J Nutr* 1994;124(suppl):1296S
52. Paisley EA, Park EI, Swartz DA, Mangian HJ, Visek WJ, Kaput J. Temporal-regulation of serum lipids and stearoyl CoA desaturase and lipoprotein lipase mRNA in *balb/c* mice. *J Nutr* 1996;126:2730
53. Park EI, Paisley EA, Mangian HJ, et al. Lipid level and type alter stearoyl CoA desaturase mRNA abundance differently in mice with distinct susceptibilities to diet-influenced diseases. *J Nutr* 1997;127:566
54. Swartz DA, Park EI, Visek WJ, Kaput J. The *c* subunit gene of murine *f1f0-*atp** synthase. Genomic sequence, chromosomal mapping, and diet regulation. *J Biol Chem* 1996;271:20942
55. Schadt EE, Monks SA, Drake TA, et al. Genetics of gene expression surveyed in maize, mouse and man. *Nature* 2003;422(6929):297