Application of nutrigenomic concepts to Type 2 diabetes mellitus

Jim Kaput, Janelle Noble, Betul Hatipoglu, Kari Kohrs, Kevin Dawson, Amelia Bartholomew

Laboratory of Nutrigenomic Medicine, Department of Surgery, University of Illinois Chicago, 840 South Wood Street MC 958, Chicago, IL 60612, USA
Center of Excellence in Nutritional Genomics, University of California at Davis, One Shields Avenue, Davis, CA 95616, USA
NuGO (European Nutrigenomics Organisation)-http://www.nugo.org
Children’s Hospital of Oakland Research Institute (CHORI), 5700 Martin Luther King, Jr. Way, Oakland, CA 94609, USA
University of Illinois Medical Center, Nutrition and Wellness Center, Medical Center Outpatient Care Center, 1C, 1801 West Taylor Street (MC 531), Chicago, IL 60612, USA

Received 10 October 2006; received in revised form 27 November 2006; accepted 28 November 2006

Abstract The genetic makeup that individuals inherit from their ancestors is responsible for variation in responses to food and susceptibility to chronic diseases such as Type 2 diabetes mellitus (T2DM). Common variations in gene sequences, such as single nucleotide polymorphisms, produce differences in complex traits such as height or weight potential, food metabolism, food-gene interactions, and disease susceptibilities. Nutritional genomics, or nutrigenomics, is the study of how foods affect the expression of genetic information in an individual and how an individual’s genetic makeup affects the metabolism and response to nutrients and other bioactive components in food. Since both diet and genes alter one’s health and susceptibility to disease, identifying genes that are regulated by diet and that cause or contribute to chronic diseases could result in the development of diagnostic tools, individualized intervention, and eventually strategies for maintaining health.
Introduction

Nutritional genomics is based on concepts and data from disciplines which historically had been considered independent research fields [1]. The most publicized aspect of nutrigenomics is the study of gene-diet associations which uses molecular genetic epidemiological methods to find statistical associations among genes, foods, and biological outcomes, such as intermediate risk factors (elevated low density lipoprotein-cholesterol) or disease outcomes (incidence and severity of Type 2 diabetes). This approach is based upon classical epidemiology that associates diets and, in some cases, certain naturally-occurring food components to disease incidence or severity in populations. The results of these studies provide information and knowledge of the environmental influences on health and disease development. Although current studies typically focus on diagnostics for chronic disease, which is the focus of this article, the goal of much research will be to develop prognostic tests for promoting health through individualized nutrition and lifestyle.

Geneticists focus on the genetic contributions to disease processes by analyzing candidate genes and their variants, such as single nucleotide polymorphisms (SNPs), in populations or cases and controls in order to associate a gene variant (allele) with a biological response. For approximately 15 years, nutritional genomics researchers have been combining classical epidemiology and genetic association approaches to examine how nutrients affect one or more intermediate risk factors in individuals with different allelic variants of candidate genes [2–4]. The results of these studies demonstrate that diets have variable effects on individuals, depending on the genetic makeup of the individual. The genes examined in most reports are generally those that have been previously identified as genetically or biochemically involved in altering either an intermediate risk factor or the chronic disease itself. The classic example is the thymidine variant instead of cytosine at position 677 (C677T) in the methylenetetrahydrofolate reductase (MTHFR) gene, which is associated with neural tube defects in women with low intakes of dietary folate in certain populations [5]. Meta-analyses of many studies showed that the proportion of disease incidence that is attributable to the TT genotype (called the population attributable risk) is about 6% [5]. An infant with the TT allele is ~1.7 times more likely (i.e., odds-ratio) to have a neural tube defect compared to other genotypes [5]. These analyses indicate that other genes and environmental factors besides MTHFR are involved in the development of neural tube defects. Genetic testing of MTHFR variants therefore provides important but not complete information about disease risk or amount of folate needed. The low predictive nature of the current genetic tests can be explained, at least in part, by the fact that a significant number of biological traits result from the contributions of multiple genes and environmental factors, each contributing different amounts to the final phenotype.

Analyzing a single SNP associated with a dietary variable such as the type or amount of dietary fat [2,6] is, therefore, unlikely to have good predictive value for susceptibility to chronic disease or for determination of an optimal, individualized diet. Hence, although classical approaches are contributing significantly to our knowledge of disease risk factors, they are limited in their ability to identify the overall genetic makeup predisposed to disease or the optimum diets needed for an individual to maintain health and prevent disease.

The tasks of creating science-based information from nutritional genomics research for health care are many and varied. In this review article, those challenges are framed by discussing the principles and concepts of nutrigenomics as they apply to Type 2 diabetes mellitus (T2DM). Although the long term goal of nutrigenomics is to improve health of individuals and thereby prevent disease, the current research in nutritional genomics often links aberrant phenotypes (high density lipoprotein [HDL]-cholesterol, as an example) influenced by diet (intake of polyunsaturated fatty acids) to one or more genes (e.g., APOAI) [6]. The rationale for this approach is that clinical measurements are available for aberrant conditions: the biological measurements of health differ among individuals of a species, a fact called biochemical individuality [7]. Hence, it is also likely that the path from research to applications will proceed, for a short time at least, through clinical practice because...
of the availability of clinical measurements and access to blood samples for genetic testing.

The mechanisms, etiology, epidemiology, and genetics of T2DM have been extensively reviewed elsewhere [8–18]. The focus here will be to integrate information from the diverse disciplines of evolutionary history, genetics, molecular biology, epidemiology, nutrition, biochemistry, medicine, social sciences, and ethical implications, an approach that requires background information for interpreting or conducting nutrigenomics research. As a polygenic, multifactorial disease, Type 2 diabetes mellitus (T2DM) can serve as a model for cancer, obesity, cardiovascular disease, and other chronic diseases influenced by diet and environment.

The characteristics of T2DM: clinical complexity

A fasting glucose level above 126 mg/dL (normal range: 70 to 100) on at least two occasions or random glucose of more than 200 mg/dL with symptoms of polyuria and polydipsia are diagnostic indicators of T2DM (see refs. [19] or [20,21]). Individuals with impaired fasting glucose levels are often given an oral glucose tolerance test that is administered in the fasted state with consumption of a high glucose drink (75 g of glucose). Although there are gradations of responses to such tests, individuals are nevertheless grouped into three classes: normal, impaired, and diabetic.

Regardless of the grouping for diabetes status, individuals may also have obesity, dyslipidemia, hypertension, insulin resistance, and/or hyperinsulinemia, which further complicates simple classification schemes for diabetes [8,13,22,23]. These physiological abnormalities may have overlapping molecular and genetic causes to further complicate diagnosis and treatment options. Many but not all patients develop co-morbidities of the disease including retinopathy, nephropathy, neuropathies, and cardiovascular disease [20]. The potential for these unpredictable manifestations of the disease cannot be assessed during initial management, potentially leading to sub-optimal clinical care.

The varying complications of T2DM are well known, yet the majority of individuals with diabetic symptoms are treated similarly [20]. A common management scheme may not be optimal for disease with multifactorial causation. For T2DM, physicians usually recommend changes to diet and an increase in physical activity, but only ~20% of patients control symptoms through these interventions [24]. The patients not helped by diet and exercise alone, or those who present with severe symptoms, are treated with one or more of 6 classes of drugs (Table 1). These drugs target different pathways and organs: insulin secretion by the pancreas (sulfonylurea and meglitinides), glucose absorption by the intestines (α-glucosidase inhibitors), glucose production in the liver (biguanide = metformin), and insulin sensitivity in adipose and peripheral tissues (e.g., rosiglitazone and pioglitazone). A newly approved agonist of glucagon-like-peptide 1, exenatide, also acts in the pancreas to stimulate insulin production ([25], Nauck, 2005 #5180) only when glucose levels are high (http://www.diabetes.org/type2-diabetes/oral-medications.jsp). Approximately 50% of T2DM patients take oral medications only, about 11% take combinations of oral agents with insulin, and the remainder take no medications (20%) or insulin alone (16.4%) [24]. Thus, current medical management of T2DM can be a lengthy trial and error method, involving significant amounts of time and considerable expense.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Drug classes for the treatment of Type 2 diabetes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>Target tissue</td>
</tr>
<tr>
<td>Lifestyle</td>
<td>All</td>
</tr>
<tr>
<td>Sulfonylurea</td>
<td>Pancreas</td>
</tr>
<tr>
<td>Meglitinides</td>
<td>Pancreas</td>
</tr>
<tr>
<td>Exenatide</td>
<td>Pancreas</td>
</tr>
<tr>
<td>Biguanides</td>
<td>Liver</td>
</tr>
<tr>
<td>Alpha-glucosidase</td>
<td>Intestine</td>
</tr>
<tr>
<td>Thiazolidinediones</td>
<td>Adipose, muscle</td>
</tr>
</tbody>
</table>

The number of subtypes of T2DM can be estimated by the different drugs used to treat different clinical indications of Type 2 diabetes. PPG is postprandial glucose response. In addition to changes in diet and physical activity levels (lifestyle), there are 6 major classes of drugs, 3 targeting the pancreas, one the liver, other pathways in the intestine, and other classes of drugs the adipose and muscle. Some patients require multiple classes of drugs including insulin. Effectiveness is the percent of patients responding to treatment (from http://www.aafp.org/PreBuilt/monograph_diabetestreatment.pdf; Accessed 2 February 2006). See text for details.
Genetic complexity of T2DM

Identifying the genetic basis of diseases caused by single genes (monogenic diseases), such as Huntington disease or cystic fibrosis, is fairly straightforward: one analyzes how frequently a chromosomal region containing a mutated gene is found in individuals showing the disease versus the frequency in individuals who do not show symptoms. In many cases, monogenic diseases are studied in families or in populations where there is evidence of dominant inheritance (if you have the mutation, you develop the disease). Although these methods are powerful for monogenic diseases, many genetic association studies fail to identify single causative genes for chronic diseases like T2DM because multiple genes [26], and the influence of multiple environmental factors acting on these genes, make variable contributions to the complex trait.

Geneticists have developed a method called quantitative trait locus (QTL) analysis to identify regions of chromosomes that contribute to a complex trait. QTLs are found by statistical analyses of how frequently a region of a chromosome is associated with a measurable phenotype, e.g., insulin levels or glucose response for T2DM. Each of the genes within QTLs may contribute different amounts to the phenotype. For example, one QTL may contribute 30% to the trait, while another may contribute 1% and the contribution may well be influenced by diet or other environmental variables. Gene variants (i.e., SNPs) may therefore be associated with small to large contributions to the complex trait. The sum of the contributions from causative alleles in different QTLs produces the specific trait or disease (rev. in ref. [27]). Almost all phenotypic traits (height potential, weight potential, fasting blood glucose, susceptibility to disease, etc.) are quantitative traits. The concept of multiple genes and multiple environmental influences contributing to a complex trait can be illustrated by examining what is currently known about the chromosomal regions containing genes that contribute to T2DM.

T2DM QTLs in humans

The approximate locations of seven QTLs that contribute to T2DM are shown in Fig. 1. These QTLs meet a minimum standard for significance of LOD (log of the odds, a measure of significance) greater than 3.6 [16]. Each of the 7 chromosomal regions is predicted to encode one or more genes that contribute to the development or severity of T2DM. Seventeen other QTLs (not shown) distributed on chromosomes 1, 2, 4, 5, 7, 8, 9, 10, 11, 12, 20, and X have also been identified but these have lower LOD (>2.0 but <3.6) scores [16]. Each of these regions, which may be as large as ~20 million base pairs, encodes one or more genes (more specifically a variant or allele of a gene) that contributes to the complex trait. Table 2 illustrates the potential complexity of T2DM by displaying how QTLs produce differing genetic susceptibilities to a chronic disease. If there are only 3 alleles at each QTL, with one contributing to T2DM (designated −), providing protection (+), or being neutral (o), the number of possible combinations for 7 loci is 2187. The actual number of combinations found in human populations is not so great because allele frequencies differ among ancestral groups. That is, chromosomes of European ancestry are likely to have a different proportion of the alleles at a given locus than do chromosomes of African ancestry. For the purpose of illustration, each of the 6 individuals "genotyped" in Table 2 inherits a different combination of protective, negative, or neutral alleles of the genes at the 7 QTLs. The genetic profiles at the extremes are individual A (all "protective" alleles and the least risk among the group) and individual F (all negative alleles and the greatest risk), with others (individuals B, C, D, and E) having intermediate risks. Note that these profiles describe only the genetic risk component and do not account for environmental influences.

Some of these genes are likely to be regulated by diet, since certain diets are risk factors for T2DM [28–30]. This means that the susceptibility to disease in each of these individuals will also vary depending upon nutrient intakes, physical activity, and other environmental factors. Some examples of other nutrient and non-nutrient environmental factors affecting the T2DM phenotype are:

- Overall sleep time and sleep continuity [31,32].
- Oxygen tension [33] which includes altitude or genetic conditions such as sickle cell disease.
- Over the counter drugs, e.g., non-steroidal anti-inflammatory drugs [34].
- Water intake relative to tea [35] and other beverages.
- Physical activity [36–40].
- Psychological factors, such as stress [41].
- Exposure to allergens and pollutants (e.g., ref. [42]).
- Circadian rhythm and seasonal changes [43].
- Balance between energy intake and expenditure (reviewed in ref. [44]).
Many genetic and environmental influences change during life and aging [45], with the net result that health and chronic diseases are not discrete, dichotomous states but rather are processes. Fig. 2 schematically shows theoretical paths of the 6 individuals (A through F in Table 2) during aging. The different heights of the initial condition (left axis) reflect the differences in genetic susceptibility (including epigenetic factors — see below). The width of each path was designed to suggest the influence of different environmental factors. Certain individuals (e.g., C and D) may be able to influence onset or severity of disease by altering lifestyle whereas others are destined for disease (e.g., F) or health (e.g., A) regardless of lifestyle. Clinical measurements are taken at discrete time points along this curve. Therefore, these measures can be thought of as only a single frame of a movie, and they may not accurately reflect past physiological processes or predict future outcomes. Although they are important, these snapshot diagnostics need to be supplemented with genetic analyses of susceptibility genes (which eventually would be assessed at birth), along with a greater understanding of their interactions with diet and the environment over the life of the individual. The recent, sudden, and dramatic increase in obesity and T2DM throughout the world [46,47] would suggest that a majority of individuals have a genetic susceptibility that can be influenced significantly by diet and lifestyle.

Genes associated with Type 2 diabetes mellitus

Identifying the causative genes within these QTLs has proven challenging, probably in part because of epistatic (gene-gene) interactions and gene — environment interactions (see below). Researchers have therefore used data from cell culture systems, laboratory animals, human physiology experiments, and candidate gene association studies to identify potential candidate genes involved in T2DM [1]. For example, Table 3 lists candidate genes that have been analyzed in gene variant-disease or gene variant-intermediate risk factor studies. Fifty-two genes in a variety of biochemical, regulatory, and signal transduction pathways have good or suggestive evidence of contribution to T2DM. A complete analysis of these genes, the biochemical pathways they are involved in, and the potential effects of each on T2DM is beyond the scope of this review. Nevertheless, polymorphisms in these genes are associated with subphenotypes of T2DM, at least in some populations (see below). It is likely that other T2DM genes have yet to be identified.

The list of genes is deceiving because many of the genes associated with T2DM in one population fail to be associated in other populations, raising the question of whether the genes listed in Table 3 cause T2DM or are simply affected by disease processes. In the absence of obvious flaws in study
Table 2  Hypothetical genotypes of 6 individuals at 7 disease loci

<table>
<thead>
<tr>
<th>QTL</th>
<th>Individual</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>+</td>
<td>o</td>
<td>o</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>o</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>o</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 2  Hypothetical genotypes of 6 individuals at 7 disease loci

Disease incidence and/or severity

Theoretical allele distribution of genes within 7 quantitative trait loci (QTL 1 through 7) for 6 individuals (A–F). + indicates an allele that is protective against T2DM, that is, an allele whose gene product contributes to optimum metabolism for preventing symptoms. o indicates a neutral allele having neither detrimental nor protective effects. – indicates an allele that alters metabolism in such a manner that symptoms of the complex trait are negatively affected. The lower portion of the figure models the effect of increasing the number of detrimental alleles on genetic susceptibility. This figure implies a linear relationship. Evidence exists that gene-gene interactions among QTLs can be additive, multiplicative, or inhibitory and certain QTLs and their interactions will be altered by gene X environment interactions. Hence, it is unlikely that simple linear relationships between the number of detrimental alleles and some complex trait will be found in nature. The linear model was used for its simplicity in explaining the concepts of QTL and genetic susceptibility. Adapted from ref. [84].

Figure 2  Genetic susceptibility, environment, aging. Individuals A through F (from Table 2) with different genetic susceptibilities are influenced differently during aging since some genes within quantitative trait loci (QTLs) are regulated by environmental influences. The symptoms of T2DM during life will differ depending upon genetic susceptibility (genetic makeup) and the influences of the environment. These influences may change during aging depending upon the genes inherited. See text for details. Adapted from ref. [84].

J. Kaput et al.
<table>
<thead>
<tr>
<th>Genetic</th>
<th>Common name</th>
<th>Function</th>
<th>Chromosome</th>
<th>References</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCC8</td>
<td>Sulfonylurea receptor</td>
<td>Potassium channel</td>
<td>11p15.1</td>
<td>[16, 85–87]</td>
<td>R</td>
</tr>
<tr>
<td>ACP1</td>
<td>Acid Phosphatase1, soluble</td>
<td>Phosphatase</td>
<td>2p25</td>
<td>[88]</td>
<td></td>
</tr>
<tr>
<td>ADA</td>
<td>Adenosine deaminase</td>
<td>Enzyme, Purine catabolic pathway</td>
<td>20q13.11*</td>
<td>[89]</td>
<td></td>
</tr>
<tr>
<td>ADRB2</td>
<td>2-Adrenergic Receptor</td>
<td>Receptor linked to catecholamine, obesity regulation</td>
<td>5q32–q34*</td>
<td>[90–92]</td>
<td>R</td>
</tr>
<tr>
<td>AGRP</td>
<td>Agouti related protein (homolog of mouse agouti)</td>
<td>Signaling, melanocortin antagonist</td>
<td>16q22</td>
<td>[93]</td>
<td></td>
</tr>
<tr>
<td>APM1</td>
<td>Adiponectin (ACDC)</td>
<td>Adipocyte hormone</td>
<td>3q27b</td>
<td>[11, 94, 95]</td>
<td></td>
</tr>
<tr>
<td>CAPN10</td>
<td>Calpain 10</td>
<td>Cysteine protease</td>
<td>2q37.3b</td>
<td>[96, 97]</td>
<td>R</td>
</tr>
<tr>
<td>ENPP1</td>
<td>Glycoprotein PC-1</td>
<td>Inhibits insulin signaling</td>
<td>6q22–q23</td>
<td>[92, 98]</td>
<td>R</td>
</tr>
<tr>
<td>FABP2</td>
<td>Liver fatty acid binding protein</td>
<td>Long chain fatty acid transport protein</td>
<td>4q28–q31</td>
<td>[99]</td>
<td></td>
</tr>
<tr>
<td>FATP4</td>
<td>Fatty acid transporter, SLCL27A4</td>
<td>Long chain fatty acid transport protein (RBC)</td>
<td>Chr. 9</td>
<td>[100]</td>
<td></td>
</tr>
<tr>
<td>FOXc2</td>
<td>Transcription Factor</td>
<td>Regulator of adipocyte metabolism</td>
<td>16q24.3</td>
<td>[11, 101]</td>
<td></td>
</tr>
<tr>
<td>FRDA</td>
<td>Frataxin</td>
<td>Mitochondrial ion metabolism</td>
<td>9q13</td>
<td>[102]</td>
<td></td>
</tr>
<tr>
<td>GC</td>
<td>Group specific component, Vitamin D binding protein</td>
<td>Vit D involved in regulating insulin levels</td>
<td>4q12</td>
<td>[92, 103]</td>
<td></td>
</tr>
<tr>
<td>GCGR</td>
<td>Glucagon receptor</td>
<td>Glucose homeostasis</td>
<td>17q25a</td>
<td>[16]</td>
<td></td>
</tr>
<tr>
<td>GCK</td>
<td>Glucokinase, liver</td>
<td>Enzyme, first step in glycolysis</td>
<td>7p15–p13*</td>
<td>[104–106]</td>
<td>E, M</td>
</tr>
<tr>
<td>GFPT2</td>
<td>Glutamine:fructose 6-phosphate amidotransferase 2</td>
<td>Hexosamine biosynthesis</td>
<td>5q34–q35a</td>
<td>[107]</td>
<td>E</td>
</tr>
<tr>
<td>GNB3</td>
<td>Guanine nucleotide binding protein 3</td>
<td>Signaling, obesity</td>
<td>12p13</td>
<td>[110]</td>
<td>E</td>
</tr>
<tr>
<td>GYS1</td>
<td>Glycogen synthase</td>
<td>Enzyme, impaired glycogen synthesis</td>
<td>19q13.3a</td>
<td>[11]</td>
<td></td>
</tr>
<tr>
<td>HNF1</td>
<td>Hepatic nuclear factor 1</td>
<td>Transcription factor, cholesterol homeostasis</td>
<td>12q24.2a</td>
<td>[111]</td>
<td>E, M</td>
</tr>
<tr>
<td>HNF4A</td>
<td>Hepatic nuclear factor 4</td>
<td>Transcription factor, hepatic glycogen stores</td>
<td>20q12–q13.1</td>
<td>[86]</td>
<td>M</td>
</tr>
<tr>
<td>IAPP</td>
<td>Insulin amyloid protein, Amylin</td>
<td>Hormone, glucose uptake pancreas</td>
<td>12p12.3–p12.1</td>
<td>[112,113]</td>
<td></td>
</tr>
<tr>
<td>IGF1</td>
<td>Insulin growth factor 1</td>
<td>Hormone, growth</td>
<td>12q22–q24.1</td>
<td>[114]</td>
<td></td>
</tr>
<tr>
<td>IL6</td>
<td>Interleukin 6</td>
<td>Cytokine</td>
<td>7p21</td>
<td>[115–118]</td>
<td>R, E</td>
</tr>
<tr>
<td>INS</td>
<td>Variable number tandem repeat in the insulin gene</td>
<td>Glucose regulation</td>
<td>11p15.5</td>
<td>[119–122]</td>
<td>R, E</td>
</tr>
<tr>
<td>INSR</td>
<td>Insulin receptor</td>
<td>Receptor</td>
<td>19p13.2</td>
<td>[86]</td>
<td></td>
</tr>
<tr>
<td>IPF1</td>
<td>Insulin promoter factor 1</td>
<td>Binds to promoters</td>
<td>12q12.1</td>
<td>[123–125]</td>
<td>M</td>
</tr>
<tr>
<td>IRS1</td>
<td>Insulin receptor substrate 1</td>
<td>Signal transduction</td>
<td>2q36b</td>
<td>[119]</td>
<td></td>
</tr>
<tr>
<td>IRS2</td>
<td>Insulin receptor substrate 2</td>
<td>Signal transduction</td>
<td>13q34</td>
<td>[126]</td>
<td></td>
</tr>
<tr>
<td>KCNJ11</td>
<td>Potassium inward rectifier channel Kir6.2</td>
<td>Potassium channel</td>
<td>11p15.1</td>
<td>[127]</td>
<td>R</td>
</tr>
</tbody>
</table>

(continued on next page)
design, execution, or data analysis, lack of association of genes among populations may be due to:
(i) chronic diseases are caused by contributions of several genes that may differ among individuals of different ancestral background; (ii) different individuals may have one or more complications such as dyslipidemia, insulin resistance, or obesity; (iii) many cases in case-control studies are molecularly heterogeneous—that is, the same phenotype can result from alterations in different genes and pathways; and (iv) the environmental variables of diet and physical activity were not analyzed.

Two additional molecular mechanisms, epistatic (gene-gene) interaction and epigenetic, also affect gene-disease association studies. Epistatic interactions can occur through protein-protein, protein-gene, RNA-protein interactions, or RNA silencing [48–51]. Proteins or enzymes produced by a gene or its variant do not act alone, but are usually part of a pathway, and many pathways are interconnected.

### Table 3 (continued)

<table>
<thead>
<tr>
<th>Genetic</th>
<th>Common name</th>
<th>Function</th>
<th>Chromosome</th>
<th>References</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>LPL</td>
<td>Lipoprotein lipase</td>
<td>Enzyme; chylomicrons and triglyceride</td>
<td>8p22</td>
<td>[130]</td>
<td>E</td>
</tr>
<tr>
<td>MAPK8IP1</td>
<td>Mitogen activated protein kinase 8—interacting protein 1</td>
<td>Signal transduction</td>
<td>11p12p11.2</td>
<td>OMIM 604641</td>
<td></td>
</tr>
<tr>
<td>NeuroD1</td>
<td>NeuroD/BETA2</td>
<td>Transcription factor, development</td>
<td>2q32b</td>
<td>[131]</td>
<td></td>
</tr>
<tr>
<td>PC1</td>
<td>Pachyonychia congenita 1</td>
<td>Inhibits insulin signaling</td>
<td>5q15–q21</td>
<td>[98]</td>
<td>R</td>
</tr>
<tr>
<td>PCK1</td>
<td>Phosphoenolpyruvate carboxykinase 1</td>
<td>Enzyme, regulation of gluconeogenesis</td>
<td>20q13.31b</td>
<td>[135]</td>
<td></td>
</tr>
<tr>
<td>PIK3R1</td>
<td>Phosphoinositide-3-kinase regulatory subunit p85</td>
<td>Glucose clearance</td>
<td>5q13</td>
<td>[139]</td>
<td></td>
</tr>
<tr>
<td>PON2</td>
<td>Paraoxonase 2</td>
<td>High fasting glucose</td>
<td>7q21.3</td>
<td>[140]</td>
<td></td>
</tr>
<tr>
<td>PPP1R3A</td>
<td>Protein phosphatase 1, regulatory (inhibitor) subunit 3A</td>
<td>Glycogen metabolism</td>
<td>7q11.23–q21.11</td>
<td>[15,144,145]</td>
<td></td>
</tr>
<tr>
<td>RORC</td>
<td>RAR-related orphan receptor C</td>
<td>Nuclear hormone, immune response</td>
<td>1q21</td>
<td>[146]</td>
<td>E</td>
</tr>
<tr>
<td>RRAD</td>
<td>Ras-related associated with diabetes</td>
<td>Insulin sensitivity</td>
<td>16q22</td>
<td>[92,147]</td>
<td></td>
</tr>
<tr>
<td>SLC2A2</td>
<td>GLUT2 glucose transporter</td>
<td>Glucose transporter</td>
<td>3q26.1–q26.3b</td>
<td>[86]</td>
<td></td>
</tr>
<tr>
<td>SLC2A4</td>
<td>GLUT4 glucose transporter</td>
<td>Guanine nucleotide exchange factor</td>
<td>17p13b</td>
<td>OMIM 138190</td>
<td>R</td>
</tr>
<tr>
<td>SOS1</td>
<td>Son of sevenless homolog</td>
<td>Blood glucose homeostasis</td>
<td>2p22–p21</td>
<td>[86]</td>
<td></td>
</tr>
<tr>
<td>TCF7L2</td>
<td>Transcription factor</td>
<td>T of rs7903146 &amp; T of rs12255372</td>
<td>[148]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNF</td>
<td>Tumor necrosis factor</td>
<td>Proinflammatory cytokine</td>
<td>6p21.3</td>
<td>[117]</td>
<td>[149–151]</td>
</tr>
<tr>
<td>UCP2</td>
<td>Uncoupling protein 2</td>
<td>Mitochondrial transporter</td>
<td>11q13</td>
<td>[117]</td>
<td></td>
</tr>
</tbody>
</table>

Incomplete list of candidate genes identified in literature searches for gene—T2DM genetic association studies. The class indicates the type or number of studies: R = replicated in at least 4 studies, E = Ethnic study, M = Mature Onset Diabetes of the Young (MODY). Some of the candidate genes map to regions that overlap QTL that have been associated with T2DM:

- Map position overlaps QTL shown in Fig. 1.
- Map position overlaps QTL with near suggestive LOD score [16].
As one example, a G to A (guanine-to-adenine) polymorphism (IVS6 + G82A) in the tyrosine phosphatase 1B (PTP1B) gene interacts statistically with a polymorphism (Gln223Arg) in the leptin receptor (LEPR) gene in a study of T2DM [52]. PTP1B and LEPR may not interact directly but may be in the same signal transduction pathway such that variants in one affect the activity of the other. 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 individual steps in a pathway to maintain the overall balance within the system is called “buffering” [53,54]. Hence, one may inherit a predisposing allele of one gene that may interact with an allele of another gene to buffer the predicted outcome. The interaction could be allele-specific and might be additive, negative, or multiplicative. These interactions are difficult to analyze in human studies because of the genetic variation in the human population that results from random matings.

Genetic ancestry matters

The probability of inheriting causative or interacting gene variants varies among populations. Variation in allele frequencies among populations may be attributed to a number of causes, including random changes (that is, genetic drift), small populations existing as small populations for extended periods of time (population bottlenecks), or selective pressure. These genetic mechanisms acted on humans during migration from east Africa that resulted in the peopling of 6 continents [55] and subsequent inbreeding within these geographically isolated populations. SNP and simple tandem repeat (STR) analyses have yielded more detailed information about human relatedness: on average, most genetic variation (estimated range of 86–88%) occurs within a geographic population (Asia, for example) [56] and only 12–14% is different between geographically distinct populations, for example, between Asia and Europe [56]. Even these small differences in allele frequencies will lead to differences in biological responses, which include responses to diet.

A specific example that illustrates this point has recently been published. The HapK haplotype (a collection of SNPs within a chromosomal region in the leukotriene 4 hydrolase (LTH4A) gene) is a greater risk factor for myocardial infarction (MI) in African Americans than in European Americans. This is presumably caused by: (i) LTH4A interacting differently with one or more gene variants in either African versus European chromosomal regions; and/or (ii) different environmental factors altering the influence of LTH4A on myocardial infarction [57]. Thus, the effect of a given allele on a trait or disease must be considered in the context of the other genes in the individual and the environmental factors that may influence its expression and/or function.

Epigenesis and chromosome structure affect expression of genetic information

Another variable that influences the statistical and real association of a SNP with a disease or response to diet is epigenetic interaction. Epigenesis is the study of heritable changes in gene function that occur without a change in the sequence of nuclear deoxyribonucleic acid (DNA). X-chromosome inactivation and gene silencing (imprinting) are examples of epigenesis [58]. Epigenetic mechanisms of altering gene regulation are DNA methylation and chromatin remodeling. Both mechanisms change the accessibility of DNA to regulatory proteins and complexes that affect transcriptional regulation and may thereby alter the expression of genetic information. If these processes result in different chromatin structure or accessibility, they can confound standard genetic analyses.

Nutrient intake affects DNA methylation status because DNA methyltransferases catalyze the transfer of a methyl group from S-adenosylmethionine to specific sites in DNA [59]. The products of the reaction are DNA methylated at (usually) CpG residues (which often occur in “cytosine-guanine islands” in DNA sequences near genes) and S-adenosylhomocysteine (S-hcy). S-adenosylmethionine is generated by the one carbon metabolic pathway, a network of interconnected biochemical reactions that transfer one-carbon groups from one metabolite to another [60]. Dietary deficiencies of choline, methionine, folate, vitamin B-12, vitamin B-6, and riboflavin affect one carbon metabolism, impair DNA methylation, and increase the risk of neural tube defects, cancer, and cardiovascular diseases [61].

Chromatin remodeling, another epigenetic mechanism that alters accessibility of DNA for transcription, is regulated in part by the energy balance in a cell, since changing calorie intake has been shown to alter chromatin remodeling, changing the NADH:NAD÷ (reduced nicotinamide adenine dinucleotide:nicotinamide adenine nucleotide) ratio (reviewed in ref. [62]) and the activity of SIRT1 (sirtuin 1), an NAD÷-dependent histone deacetylase. Chromatin remodeling occurs through a series of enzymatic reactions and protein — protein interactions that ultimately affect expression
of genetic information. Alteration of the level of the proteins, enzymes, and RNAi (interfering RNA [63]) involved in chromatin remodeling, by diet or other environmental factors, is another control point for regulating gene expression.

Long term exposure to diets that influence chromatin remodeling and DNA methylation could induce permanent epigenetic changes in the genome. Such changes might explain why certain individuals can more easily control symptoms of chronic diseases by changing lifestyle but many seem to pass an irreversible threshold. Epigenetic changes may also explain "developmental windows" — key times during development, such as in utero, where short-term environmental influences may produce long-lasting changes in gene expression and metabolic potential (reviewed in ref. [23]). Developing experimental approaches for dissecting the environmental influences and the critical genes and pathways will be essential and challenging.

**Genotype X environment interactions**

Genes that cause chronic diseases must be regulated directly or indirectly by calorie intake and/or by specific chemicals in the diet because diet alters the transcription of genetic information. Alteration of the level of the proteins, enzymes, and RNAi (interfering RNA [63]) involved in chromatin remodeling, by diet or other environmental factors, is another control point for regulating gene expression.

Long term exposure to diets that influence chromatin remodeling and DNA methylation could induce permanent epigenetic changes in the genome. Such changes might explain why certain individuals can more easily control symptoms of chronic diseases by changing lifestyle but many seem to pass an irreversible threshold. Epigenetic changes may also explain "developmental windows" — key times during development, such as in utero, where short-term environmental influences may produce long-lasting changes in gene expression and metabolic potential (reviewed in ref. [23]). Developing experimental approaches for dissecting the environmental influences and the critical genes and pathways will be essential and challenging.

Many studies examining candidate gene-disease associations (see Table 3) usually do not account for differences in diet. Gene-diet-phenotype association studies have focused primarily on intermediate risk factors, particularly for cardiovascular disease [3]. Fewer such studies have been conducted for T2DM or the metabolic syndrome. The primary exception has been the association of total and saturated dietary fats (e.g., ref. [66]) with the Pro12Ala variant of peroxisome proliferator activated receptor gamma 2 (PPAR-γ2). Other genes listed in Table 3 that have been associated with nutrient intakes are:

- Hepatic lipase and fat [73];
- INS and glucose intake [74];
- PON and alcohol [75];
- PPAR-γ and fat [66];
- UCP and energy and body weight [76] and chronic overfeeding [77].

However, conflicting results have been obtained that may be attributable to population stratification and/or too few study participants (rev. in refs. [78,79]). Well-designed and highly-powered studies are needed to unravel the complexity of gene-nutrient interactions underlying TZDM and its precursor, the metabolic syndrome.

**Converting science into practice**

Until our understanding of diet-gene interactions for a particular disorder is analyzed in greater detail and depth, the knowledge cannot be transformed into useful applications for societal benefit, and nutrigenomics remains more promise than practice. Diagnostics, preventive lifestyle guidelines, more efficacious dietary recommendations, health-promoting food supplements, and drugs are some of the anticipated end-products of nutrigenomics research.

Genetic and metabolomic diagnostics will be critical for developing treatment options for disease. Components in food often influence pathways involved in disease development (e.g., ref. [29,30]), which means that nutrigenomics testing is complementary to and overlaps pharmacogenomics testing (genetic testing for drug efficacy and safety in an individual). As one example, the natural ligands that activate peroxisome proliferator activated receptor gamma 2 (PPAR-γ) are eicosapentaenoic acid [80] and its derivatives (e.g., 15-deoxy-D12,14-prostaglandin J2 (PGJ2) [81,82]). Rosiglitazone, a member of the thiazolidinedione (TZD) class of drugs for T2DM, also binds and activates PPAR-γ. Hence, some components of the diet affect the same pathways that drugs affect, demonstrating the overlap between nutrigenomics and pharmacogenomics.

For consumers, the initial introduction to the practical applications of nutrigenomics will be through clinical diagnostics for "subphenotypes" such as insulin levels, glucose tolerance, or some defined and well-accepted "intermediate" biomarkers of a disease. This type of genetic testing is no different than the analyses of other diagnostic biomarkers such as cholesterol levels. However, analyzing variants in one gene/protein in the complex pathway of absorption, transport, metabolism, or utilization of a dietary component is not likely to provide information for dietary advice.
Personalized nutrition and its application to personalized medicine based on genetic testing require additional scientific research for determining the makeup of accurate genetic tests. Nevertheless, the use of genotype-phenotype-biomarker diagnostics is beginning and will be of fundamental importance in the translation of nutrigenomics beyond the laboratory to the consumer market.

Many consumers state that they make food choices with the intent of benefiting their or their families’ health. The complexity of analyzing multiple genes in multiple metabolic pathways will be difficult to interpret for the average consumer. Hence, a key need for translating nutrigenomic tests from academic knowledge to real world utility will be nutritional genetic (nutrigenomics) practitioners capable of interpreting genetic tests and linking that knowledge to specific nutrients and diets.

Summary

Nutrigenomics holds great potential for improving personal and public health through prognostic testing of genes associated with specific classes of nutrients. The tasks facing the development of science-based gene-diet interactions, while challenging, may be overcome by adopting best practices for gene-diet-disease association studies and by strategic international alliances [83] for testing gene-nutrient associations in different ethnic populations. Nutrition professionals will contribute significantly to the conversion of research to practical knowledge and will be the key purveyors of this knowledge for optimizing health of individuals.

Acknowledgment

We thank Ruth DeBusk for reading, commenting, and editing this manuscript during its preparation. Supported in part by National Center for Minority Health and Health Disparities Center of Excellence in Nutritional Genomics (MD00222) and from the European Union, EU FP6 NoE Grant, Contract No. CT2004-505944.

References


