Published in final edited form as:

Mutat Res. 2007 September 1; 622(1-2): 19–32.

Complexity of Type 2 Diabetes Mellitus Data Sets Emerging from Nutrigenomic Research: A Case for Dimensionality Reduction?

Jim Kaput, PhD^{1,2,3} and Kevin Dawson¹

¹ Center of Excellence in Nutritional Genomics, University of California at Davis, Davis, CA 95616

² Laboratory of Nutrigenomic Medicine, Department of Surgery, University of Illinois Chicago, Chicago, IL 60612

³ NuGO (European Nutrigenomics Organisation) www.nugo.org

Abstract

Nutrigenomics promises personalized nutrition and an improvement in preventing, delaying, and reducing the symptoms of chronic diseases such as diabetes. Nutritional genomics 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. The path to those promises has significant challenges, from experimental designs that include analysis of genetic heterogeneity to the complexities of food and environmental factors. One of the more significant complications in developing the knowledge base and potential applications is how to analyze high-dimensional datasets of genetic, nutrient, metabolomic (clinical), and other variables influencing health and disease processes. Type 2 diabetes mellitus (T2DM) is used as an illustration of the challenges in studying complex phenotypes with nutrigenomics concepts and approaches.

Introduction

Type 2 diabetes (T2DM) is an example of a complex trait – that is, it results from the contribution of many genes [1], many environmental factors including diet [2], and the interactions among those genes and environmental factors. Various but differing combinations of these factors can produce the same clinical features. The key clinical feature characterizing T2DM, high blood glucose levels [3], may be caused by aberrations in one or more different molecular pathways. T2DM also presents with related physiological responses of hyperinsulinemia, insulin resistance, and other complications. Each of these physiological responses may result from gene – nutrient (and environment) interactions. Hence, the study of nutritional genomics applies to T2DM [4], and indeed almost all chronic diseases, because subsets of nutrient – gene interactions contribute to health or cause disease [5].

The overarching challenges for understanding these disease processes, and indeed, all biological processes including health, are the genetic heterogeneity of humans, the complications of overlapping and varied disease mechanisms, and the complexity of diet and other environmental variables. High-throughput technologies developed in the past 15 years now permit the analysis of hundreds of thousands of genes and their variants, hundreds to thousands of clinical markers such as metabolites, and, theoretically, large numbers of different nutrients and bioactives in foods. The ability to generate high-dimensional datasets however, presents another significant challenge: how to find patterns among genetic, environmental, and

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

clinical symptoms that define and explain complex biological processes. The focus of this review is on the concepts underlying the complexities of the gene – nutrient interactions and emerging approaches for analyzing interacting genetic, molecular, dietary, and clinical data.

The Characteristics of T2DM: Clinical Complexity

The mechanisms, etiology, epidemiology, and genetics of T2DM have been extensively reviewed elsewhere [6-16]. The key diagnostic indicator of T2DM is a fasting blood glucose level above 126 mg/dL (normal range: 70 to 100) on at least two occasions or random blood glucose level of more than 200 mg/dL with symptoms of polyuria and polydipsia (see [17] or [18,19]). Further classification of individuals with impaired fasting glucose levels is done with an oral glucose tolerance (OGT) test. Subjects consume a high-glucose drink (75 g of glucose) administered in the fasted state. Although there are gradations of responses to biological tests, individuals are nevertheless grouped into three classes: normal, impaired, and diabetic.

In addition to the abnormally high circulating glucose and the OGT response, individuals may also be obese, hypertensive, have dyslipidemia, insulin resistance, and/or hyperinsulinemia [6,11,20,21]. These abnormalities may have overlapping molecular and genetic causes to further complicate diagnosis and treatment options. During the course of the disease, many but not all patients develop co-morbidities of the disease including retinopathy, nephropathy, neuropathies, and cardiovascular disease [18]. The potential for these manifestations of the disease cannot be assessed during initial diagnosis, potentially leading to sub-optimal management of the disease, further complications, and increased healthcare costs.

While the varying complications of T2DM are well known, the majority of individuals with diabetic symptoms are treated similarly [18], first with lifestyle changes and then with drugs. Table 1 lists the 6 major drugs used to treat T2DM, 3 of which stimulate insulin production in the pancreas, and the others affect glucose production in the liver, glucose uptake in the intestine, and glucose utilization (through PPAR activity that alters insulin resistance) in the peripheral tissues. The final result is that only ~20% of patients control symptoms through lifestyle changes [22], about 50% of T2DM patients take oral medications only (Table 1), about 11% take combinations of oral agents with insulin, and 16% take insulin alone [22]. Drug responsiveness alone demonstrates the molecular and lifestyle heterogeneity of T2DM. Optimizing the medications for each patient can be a lengthy trial and error process, involving significant amounts of time and considerable expense.

Genetic Complexity of T2DM

The contributions of many genes [23] and the numerous environmental factors that alter genetic expression confound simple experimental designs for identifying gene, nutrient, or gene – nutrient interactions that cause (rather than affected by) T2DM. The experimental methods successfully used for identifying the genetic basis of monogenic (single gene) diseases {Jimenez-Sanchez, 2001 #1678}, such as Huntington disease or cystic fibrosis, can not be directly applied to identifying the genetic basis of complex traits such as T2DM [24].

However, a variation on association studies, quantitative trait locus (QTL) analysis, can identify regions of chromosomes that contribute to a complex trait [25,26]. QTLs are found by statistical analysis of how frequently a region of a chromosome is associated with a measurable phenotype, e.g., insulin levels or glucose response for T2DM. QTLs identify chromosomal regions containing one or more genes contribute different amounts to the phenotype. For example, one QTL may contribute 20% to the trait, while another may contribute only 1%. The contributions will vary depending upon gene – nutrient interactions for the gene responsible for the QTL and whether that gene interacts with other genes in the genome (epistasis – see below). The sum of the contributions from causative alleles in different QTLs

produces the specific trait or disease (rev. in [27]). 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

Seven QTLs with LOD (log of the odds, a measure of significance) greater than 3.6 [14] have been found to contribute to T2DM. These loci were found in specific and differing ancestral groups and are on chromosomes 1q25.3, 2q37.3, 3p24.1, 3q28, 10q26.13, 12q24.31, and 18p11.22 [14]. To illustrate the potential complexity of T2DM, assume that there are only 3 alleles at each QTL, with one providing protection against developing the disease, or being neutral, or contributing to the disease. With three possible alleles at 7 loci, the number of potential combinations is 2187. The actual number found in human populations is not as large 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. The genes, their quantitative contributions in different ancestral backgrounds, and the combination of alleles at the 7 loci that produce different age of onsets and severities of T2DM are not known. To add to the complexity, seventeen other “suggestive” QTLs found on regions of chromosomes 1, 2, 4, 5, 7, 8, 9, 10, 11, 12, 20, and X have also been identified. These QTLs have LOD scores between 2.0 and 3.6 [14], values below the accepted threshold of 3.6.

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 that alter expression of genetic information. Some examples of other nutrient and non-nutrient environmental factors affecting the T2DM phenotype are energy intake and balance (reviewed in [31]), sleep patterns ([32,33]), physical activity [34-38], liquid intake [39], and psychological factors, such as stress [40]. Aging and the changes induced by aging, produce health and chronic diseases that are not discrete, dichotomous states, but rather are a continuum. Populations have individuals at different points on this continuum from health to disease. For convenience and treatment, individuals with similar phenotypes are grouped regardless of the differing underlying molecular genetic causes of their condition. The inability to categorize individuals into classes or groups with similar genetic makeups and environmental exposures is a key limitation for identifying genes that cause from those that are affected by the disease.

Genes Associated with Type 2 Diabetes Mellitus

Genetic studies have identified over 50 genes (Table 2) as being involved in T2DM, primarily by association analyses (rev in [4]). While a full description of each of these genes is beyond the scope of this review, they participate in a variety of biochemical, regulatory, and signal transduction pathways, many of which are known to be involved in producing the phenotypes associated with T2DM.

The list in Table 2 is deceiving because many of the genes associated with T2DM in one population fail to be associated in other population, raising the question of whether these genes cause T2DM or are simply affected by disease process. In the absence of obvious flaws in study design, execution, or data analysis [41], lack of association of genes among populations may be due to (i) chronic diseases that 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 which confound statistical analyses, (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.

Gene-gene interaction (epistasis) may also affect gene–disease association studies. Epistatic interactions can occur through protein-protein, protein-gene, RNA-protein interactions, or RNA silencing [42-45]. Epistasis occurs because genes, RNA, proteins, and enzymes are usually part of a pathway or pathways, many of which are interconnected. The concept of biochemical buffering [46,47] or homeostasis, provides a plausible explanation: biological systems try to maintain balance. Inheriting one predisposing allele, therefore, may not alter the final phenotype (such as disease) because an allele of another gene may buffer the effect of the first allele. Gene – gene interactions can be allele-specific and may be additive, negative, or multiplicative. To illustrate epistasis, subjects carrying the adiponectin G allele and the PPAR γ 2 Ala12 allele seem to be more insulin sensitive than those that encode the adiponectin T allele. Interactions between adiponectin and PPAR γ 2 genotypes also contribute to fasting insulin concentrations, insulin concentrations in oral glucose tolerance tests, and insulin resistance index [48].

Genetic Ancestry Matters

Epistatic interactions occur because of the gene variants that an individual inherits. The HapMap project has demonstrated that 85 – 90% [49,50] of human variation occurs within a population and only about 10 – 15% is unique to one of the geographical ancestral populations. However, the frequency of alleles differs among ancestral populations such that the probability of inheriting causative or interacting gene variants varies. Genetic drift, small populations existing for extended periods of time (population bottlenecks), or selective pressure all contribute to selection of various alleles. Even small differences in allele frequencies among populations, and therefore among individuals, 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 single nucleotide polymorphisms (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 [51]. Thus, the effect of a given allele on a trait or disease must be considered in the context of the other genes [52] in the individual and the environmental factors that may influence its expression and/or function [53,54].

Epigenesis and Chromosome Structure Affect Expression of Genetic Information

Epigenetic interactions also influence the statistical and real association of a SNP with a disease or response to diet. Epigenetics are heritable alterations in gene activity that occur without a change in the sequence of nuclear DNA. X-chromosome inactivation and gene silencing (imprinting) are examples of epigenesis [55]. The two primary mechanisms of epigenesis are DNA methylation and chromatin remodeling, which are separate but related pathways. DNA is methylated at (primarily [56]) CpG dinucleotides which often are clustered in CpG islands near promoters of genes [57]. Proteins which bind to methyl-CpG, (MCPs [58]) interact with other proteins, including chromatin protein modifying enzymes such as histone deacetylases (HDAC) [59]. HDACs in turn alter chromatin structure by changing the activators, co-activators, repressors, and co-repressors that bind to DNA. The net result of DNA methylation and chromatin remodeling is to change the accessibility of DNA to regulatory proteins and transcriptional complexes.

Nutrients alter epigenetic mechanisms through several distinct pathways [60,61]. Dietary deficiencies of choline, methionine, folate, vitamin B₁₂, vitamin B₆, and riboflavin [62,63] alter one carbon metabolism and the concentration of one of its final products, S-adenosylmethionine (S-AM). S-AM is the immediate precursor of DNA (and other) methylation reactions. Chromatin remodeling is a related mechanism that contributes to the control of gene expression. Chromatin structure is regulated in part by the energy balance in a cell: changes in calorie intake alters the NADH:NAD⁺ (reduced nicotinamide adenine dinucleotide:nicotinamide adenine nucleotide) ratio (reviewed in [64]) and the activity of *SIRT1* (sirtuin 1), an NAD⁺-dependent histone deacetylase. The soy – derived peptide, lunasin, may also influence chromatin remodeling by affecting the acetylation – deacetylation balance in chromatin [65]. It is not known whether other peptides can alter acetylation – deacetylation reactions. Changing transcriptional profiles by these methods would alter the phenotype independently of the presence of gene variants (SNPs or collections of SNPs).

Long-term exposure to diets that influence chromatin remodeling (e.g., calories and soybean protein) and DNA methylation (one carbon metabolites) 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 [21]). Developing experimental approaches for dissecting the environmental influences and the critical genes and pathways will be essential and challenging.

Genotype X Environment Interactions

The concept and definition of gene X environment interactions in nutrition were proposed by Young in 1979 [66]. The precise, statistical definition of gene X environment interaction is “a different effect of an environmental exposure on disease risk in persons with different genotypes,” or, alternatively, “a different effect of a genotype on disease risk in persons with different environmental exposures” [67]. In other words, nutrients affect expression of genetic information and genetic makeup affects how nutrients are metabolized.

Most studies examining candidate gene-disease associations (see Table 2) usually do not account for differences in nutrient intakes among study participants. Gene-diet-phenotype association studies have focused primarily on intermediate risk factors, particularly for cardiovascular disease [68]. 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., [69]) with the Pro12Ala variant of peroxisome proliferator activated receptor gamma 2 (PPAR- γ 2). Other gene – nutrient interactions, which are not necessarily involved in T2DM, are listed in Table 3. These nutrient – gene – phenotype associations are not always replicated in separate studies. Lack of replication may be attributed to small effect size, population stratification and/or too few study participants (rev. in [70,71]). Well-designed and highly-powered studies are needed to unravel the complexity of gene-nutrient interactions underlying T2DM and its precursor, the metabolic syndrome.

The Analytical Challenge: Finding the Patterns in High-Dimensional Datasets

A key challenge facing high throughput technologies is discovering the functional relationships among elements in genomics, transcriptomics, proteomics, and metabolomics data sets. How are the various levels of quantitative information (DNA, RNA, protein, and metabolite) that define biological processes related to each other and how does diet alter their levels, relationships, interactions, and affect on biological processes [72,73] High-dimensional “-

omics” datasets offer additional analytical challenges, for example, the need to control missing values and the high false discovery rate.

The fundamental statistical challenge occurs because each of these “-omics” datasets have high dimensionality – that is, quantitative measurements of the amount (or relative amount) of all (or many) individual components. In a typical genomic, transcriptomic, or proteomic dataset, the number of variables (dimensions) often exceeds the number of samples. For example, many genomic studies measure hundreds of thousands of single nucleotide polymorphism (SNPs). Typical gene expression studies measure tens of thousands of mRNA expression levels; and high-resolution matrix assisted laser desorption ionization time of flight (MALDI-TOF) mass spectrometry proteomic spectra may contain as much as 500,000 data points. In spite of the high-dimensional “-omics” datasets, the number of samples is limited by economic reasons. It is a common challenge in nutritional genomics, that practically each gene has some level of change of expression relative to a nutrient intake or treatment. Each gene – level measurement may have direct, indirect, or inverse correlative relationships to the levels of other gene transcripts, and the sum of these relationships define the transcriptional response to the nutrient or other perturbation, such as the presence or absence of disease or complications of disease. High-dimensional datasets may be represented in a lower dimensional coordinate system by creating a topological map of all gene – level values. The shape of low-dimensional models of any given high-dimensional “-omics” dataset will be influenced and functionally related to the other quantitative measurements of biological components (transcript – protein levels and/or metabolite levels).

A three dimensional topological map has both linear and nonlinear relationships among data elements. Traditional statistical methods are based on regression analysis, that is, they assume linear relationships. Standard statistical methods are often limited in analyzing complex patterns because they omit known interactions among genes and metabolites produced by interacting or competing intracellular pathways and physiological processes [74,75]. Rather, nonlinear analytical methods are increasingly being applied to the analyses of complex biological data (Table 4). One such approach is dimensionality reduction (DR). DR is the concept for analyzing complex data by transforming the original high-dimensional data set (n dimensions) into a lower dimensional data space (m dimensions, where $m < n$). Instead of graphing 20,000+ individual expression levels (for example), dimensionality reduction methods identify, group, and graph genes or samples with similar properties into 1, 2, 3, or more dimensions [74,75]. Dimensionality reduction algorithms can also discover low-dimensional structures in high-dimensional datasets. The discovered low-dimensional structures may identify relatedness of groups or clusters of genes, proteins, or metabolites that define disease processes, genetic regulatory networks, or disease subtypes. Since the structure of high-dimensional datasets can not be known *a priori*, the most appropriate method is determined as the one that can best explain most of the variance in the data.

The concept of dimensionality reduction differs from hypothesis – driven, or supervised, approaches. While hypothesis driven experiments are typically considered the gold standard for scientific research, they introduce subjective bias by selecting the possible outcome, do not consider or test for unexpected results, and often filter noise that is considered “unimportant” but which may have un-analyzed information. Factor analysis, classification analysis and regression trees (CART) methods, and artificial neuronal networks (ANN) are examples of supervised techniques (Table 4). Unsupervised analyses, or data-driven approaches, have the disadvantage of being descriptive, but these procedures avoid subjective bias of selecting analyses of a presumed result and they allow for class (or group) discovery. The disadvantage of unsupervised analyses, however, are that they high false discovery rate (type 1 error) and prior knowledge might not explain the results.

The known biological interactions among measured variables (e.g., known gene-gene interactions in gene networks) can be used to further refine results of the unsupervised analyses. Concentrations of one or more metabolites, genetic variations, and alterations in dietary intake may have specific patterns associated with physiological outcomes and responses to treatments. Examples of these interconnections among RNA, protein, and metabolite levels have been published [76]. The goal is to use one or more dimensionality reduction methods to discover interactions and patterns among these datasets that explain complex traits, including health and disease. That is, complexity can be considered an asset rather than an impediment [77], since biological systems are inherently complex, and reductionist approaches fail to account for that complexity. Multidimensional data for T2DM are quantitative measurements of serum or plasma levels of metabolites and proteins, single nucleotide polymorphisms in genes, ancestry-informative markers (AIMs for analyzing genetic background), and dietary variables.

Summary

Modern analytical methods have the ability to generate large datasets of DNA variations, RNA levels, protein amounts, and metabolite measurements. Aberrations in physiology, such as type 2 diabetes and other chronic diseases, introduce additional variables. Many of the standard analytical tools assumed that biological responses were based on linear relationships, an assumption which can not explain phenotypes based on gene – gene, gene – nutrient, or gene – nutrient interactions. While the number of reports describing the analyses of biological data with data reduction methods is increasing [74,75,78-86], it appears that no “rule-of-thumb” can be used to choose the most appropriate analytical method. Rather, the algorithm that best describes the data and relationships, usually as determined by the smallest variance, is the method of choice for that dataset. Discovery of novel patterns among nutrient intakes, genetic ancestry, candidate genes, and physiological measurements may allow us a greater understanding of disease pathologies, and a path to developing ways to maintain health and prevent or delay diseases.

Acknowledgements

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

1. Hansen L, Pedersen O. Genetics of type 2 diabetes mellitus: status and perspectives. *Diabetes Obes Metab* 2005;7:122–135. [PubMed: 15715885]
2. Schulze MB, Hu FB. Primary prevention of diabetes: what can be done and how much can be prevented? *Annu Rev Public Health* 2005;26:445–467. [PubMed: 15760297]
3. Stumvoll M, Goldstein BJ, van Haeften TW. Type 2 diabetes: principles of pathogenesis and therapy. *Lancet* 2005;365:1333–1346. [PubMed: 15823385]
4. Kaput J, Noble J, Hatipoglu B, Kohrs K, Dawson Kk, Bartholomew A. A Application of Nutrigenomic Concepts to Type 2 Diabetes Mellitus, Nutrition, Metabolism, and Cardiovascular. Disease. 2006in press
5. Fairweather-Tait SJ. Human nutrition and food research: opportunities and challenges in the post-genomic era. *Philos Trans R Soc Lond B Biol Sci* 2003;358:1709–1727. [PubMed: 14561328]
6. Curtis J, Wilson C. Preventing type 2 diabetes mellitus. *J Am Board Fam Pract* 2005;18:37–43. [PubMed: 15709062]
7. Pirola L, Johnston AM, Van Obberghen E. Modulation of insulin action. *Diabetologia* 2004;47:170–184. [PubMed: 14722654]
8. Patti ME. Gene expression in humans with diabetes and prediabetes: what have we learned about diabetes pathophysiology? *Curr Opin Clin Nutr Metab Care* 2004;7:383–390. [PubMed: 15192439]

9. Parikh H, Groop L. Candidate genes for type 2 diabetes. *Rev Endocr Metab Disord* 2004;5:151–176. [PubMed: 15041791]
10. Laakso M. Gene variants, insulin resistance, and dyslipidaemia. *Curr Opin Lipidol* 2004;15:115–120. [PubMed: 15017354]
11. Steinmetz A. Treatment of diabetic dyslipoproteinemia. *Exp Clin Endocrinol Diabetes* 2003;111:239–245. [PubMed: 12951627]
12. Hanson RL, Knowler WC. Quantitative trait linkage studies of diabetes-related traits. *Curr Diab Rep* 2003;3:176–183. [PubMed: 12728644]
13. Hansen L. Candidate genes and late-onset type 2 diabetes mellitus. Susceptibility genes or common polymorphisms? *Dan Med Bull* 2003;50:320–346. [PubMed: 14694850]
14. Florez JC, Hirschhorn J, Altshuler D. The inherited basis of diabetes mellitus: implications for the genetic analysis of complex traits. *Annu Rev Genomics Hum Genet* 2003;4:257–291. [PubMed: 14527304]
15. McCarthy MI, Froguel P. Genetic approaches to the molecular understanding of type 2 diabetes. *Am J Physiol Endocrinol Metab* 2002;283:E217–225. [PubMed: 12110525]
16. Freeman H, Cox RD. Type-2 diabetes: a cocktail of genetic discovery. *Hum Mol Genet* 2006;15:R202–209. [PubMed: 16987885]
17. EndocrineWeb. (Accessed 31 January 2006)
18. Nathan DM. Clinical practice. Initial management of glycemia in type 2 diabetes mellitus. *N Engl J Med* 2002;347:1342–1349. [PubMed: 12397193]
19. Ahmann AJ, Riddle MC. Current oral agents for type 2 diabetes. Many options, but which to choose when? *Postgrad Med* 2002;111:32–34, 37–40, 43–36. [PubMed: 12040862]
20. Cordain L, Eades MR, Eades MD. Hyperinsulinemic diseases of civilization: more than just Syndrome X. *Comp Biochem Physiol A Mol Integr Physiol* 2003;136:95–112. [PubMed: 14527633]
21. McMillen IC, Robinson JS. Developmental origins of the metabolic syndrome: prediction, plasticity, and programming. *Physiol Rev* 2005;85:571–633. [PubMed: 15788706]
22. Koro CE, Bowlin SJ, Bourgeois N, Fedder DO. Glycemic control from 1988 to 2000 among U.S. adults diagnosed with type 2 diabetes: a preliminary report. *Diabetes Care* 2004;27:17–20. [PubMed: 14693960]
23. Wolford JK, Vozarova de Courten B. Genetic basis of type 2 diabetes mellitus: implications for therapy. *Treat Endocrinol* 2004;3:257–267. [PubMed: 16026108]
24. Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 1995;11:241–247. [PubMed: 7581446]see comments
25. Belknap JK, Phillips TJ, O'Toole LA. Quantitative trait loci associated with brain weight in the BXD/Ty recombinant inbred mouse strains. *Brain Res Bull* 1992;29:337–344. [PubMed: 1393606]
26. Barton NH, Keightley PD. Understanding quantitative genetic variation. *Nat Rev Genet* 2002;3:11–21. [PubMed: 11823787]
27. Flint J, Valdar W, Shifman S, Mott R. Strategies for mapping and cloning quantitative trait genes in rodents. *Nat Rev Genet* 2005;6:271–286. [PubMed: 15803197]
28. Kaput J, Klein KG, Reyes EJ, Kibbe WA, Cooney CA, Jovanovic B, Visek WJ, Wolff GL. Identification of genes contributing to the obese yellow Avy phenotype: caloric restriction, genotype, diet x genotype interactions. *Physiol Genomics* 2004;18:316–324. [PubMed: 15306695]
29. Kaput J. Diet-disease gene interactions. *Nutrition* 2004;20:26–31. [PubMed: 14698010]
30. 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;124:1296S–1305S. [PubMed: 7914920]
31. Seeley RJ, Drazen DL, Clegg DJ. The critical role of the melanocortin system in the control of energy balance. *Annu Rev Nutr* 2004;24:133–149. [PubMed: 15189116]
32. Irwin M. Effects of sleep and sleep loss on immunity and cytokines. *Brain Behav Immun* 2002;16:503–512. [PubMed: 12401464]
33. Redwine L, Hauger RL, Gillin JC, Irwin M. Effects of sleep and sleep deprivation on interleukin-6, growth hormone, cortisol, and melatonin levels in humans. *J Clin Endocrinol Metab* 2000;85:3597–3603. [PubMed: 11061508]

34. Nieman DC, Davis JM, Henson DA, Walberg-Rankin J, Shute M, Dumke CL, Utter AC, Vinci DM, Carson JA, Brown A, Lee WJ, McAnulty SR, McAnulty LS. Carbohydrate ingestion influences skeletal muscle cytokine mRNA and plasma cytokine levels after a 3-h run. *J Appl Physiol* 2003;94:1917–1925. [PubMed: 12533503]
35. Nieman DC, Davis JM, Brown VA, Henson DA, Dumke CL, Utter AC, Vinci DM, Downs MF, Smith JC, Carson J, Brown A, McAnulty SR, McAnulty LS. Influence of carbohydrate ingestion on immune changes after 2 h of intensive resistance training. *J Appl Physiol* 2004;96:1292–1298. [PubMed: 14672962]
36. Nieman DC, Dumke CI, Henson DA, McAnulty SR, McAnulty LS, Lind RH, Morrow JD. Immune and oxidative changes during and following the Western States Endurance Run. *Int J Sports Med* 2003;24:541–547. [PubMed: 12968214]
37. Chakravarthy MV, Booth FW. Eating, exercise, and “thrifty” genotypes: connecting the dots toward an evolutionary understanding of modern chronic diseases. *J Appl Physiol* 2004;96:3–10. [PubMed: 14660491]
38. Gleeson M, Nieman DC, Pedersen BK. Exercise, nutrition and immune function. *J Sports Sci* 2004;22:115–125. [PubMed: 14971437]
39. Tomita M, Irwin KI, Xie ZJ, Santoro TJ. Tea pigments inhibit the production of type 1 (T(H1)) and type 2 (T(H2)) helper T cell cytokines in CD4(+) T cells. *Phytother Res* 2002;16:36–42. [PubMed: 11807963]
40. Irwin M, Clark C, Kennedy B, Christian Gillin J, Ziegler M. Nocturnal catecholamines and immune function in insomniacs, depressed patients, and control subjects. *Brain Behav Immun* 2003;17:365–372. [PubMed: 12946658]
41. Ioannidis JP. Why most published research findings are false. *PLoS Med* 2005;2:e124. [PubMed: 16060722]
42. He L, Hannon GJ. MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet* 2004;5:522–531. [PubMed: 15211354]
43. Nakahara K, Carthew RW. Expanding roles for miRNAs and siRNAs in cell regulation. *Curr Opin Cell Biol* 2004;16:127–133. [PubMed: 15196554]
44. Scherr M, Eder M. RNAi in functional genomics. *Curr Opin Mol Ther* 2004;6:129–135. [PubMed: 15195923]
45. Goto A, Blandin S, Royet J, Reichhart JM, Levashina EA. Silencing of Toll pathway components by direct injection of double-stranded RNA into *Drosophila* adult flies. *Nucleic Acids Res* 2003;31:6619–6623. [PubMed: 14602922]
46. Caporaso NE. Why have we failed to find the low penetrance genetic constituents of common cancers? *Cancer Epidemiol Biomarkers Prev* 2002;11:1544–1549. [PubMed: 12496041]
47. Hartman JL, Garvik B, Hartwell L. Principles for the buffering of genetic variation. *Science* 2001;291:1001–1004. [PubMed: 11232561]
48. Yang WS, Hsiung CA, Ho LT, Chen YT, He CT, Curb JD, Grove J, Quertermous T, Chen YD, Kuo SS, Chuang LM. Genetic epistasis of adiponectin and PPARgamma2 genotypes in modulation of insulin sensitivity: a family-based association study. *Diabetologia* 2003;46:977–983. [PubMed: 12827242]
49. Jorde LB, Wooding SP. Genetic variation classification and ‘race’. *Nat Genet* 2004;36:S28–33. [PubMed: 15508000]
50. Tishkoff SA, Kidd KK. Implications of biogeography of human populations for ‘race’ and medicine. *Nat Genet* 2004;36:S21–27. [PubMed: 15507999]
51. Helgadóttir A, Manolescu A, Helgason A, Thorleifsson G, Thorsteinsdóttir U, Gudbjartsson DF, Gretarsdóttir S, Magnusson KP, Gudmundsson G, Hicks A, Jonsson T, Grant SF, Sainz J, O’Brien S J, Sveinbjornsdóttir S, Valdimarsson EM, Matthiasson SE, Levey AI, Abramson JL, Reilly MP, Vaccarino V, Wolfe ML, Gudnason V, Quyyumi AA, Topol EJ, Rader DJ, Thorgeirsson G, Gulcher JR, Hakonarson H, Kong A, Stefansson K. A variant of the gene encoding leukotriene A4 hydrolase confers ethnicity-specific risk of myocardial infarction. *Nat Genet* 2006;38:68–74. [PubMed: 16282974]

52. Hamon SC, Kardia SL, Boerwinkle E, Liu K, Klos KL, Clark AG, Sing CF. Evidence for consistent intragenic and intergenic interactions between SNP effects in the APOA1/C3/A4/A5 gene cluster. *Hum Hered* 2006;61:87–96. [PubMed: 16710093]
53. Corella D, Lai CQ, Demissie S, Cupples LA, Manning AK, Tucker KL, Ordovas JM. APOA5 gene variation modulates the effects of dietary fat intake on body mass index and obesity risk in the Framingham Heart Study. *J Mol Med*. 2007
54. Corella D, Ordovas JM. Single Nucleotide Polymorphisms that Influence Lipid Metabolism: Interaction with Dietary Factors. *Annu Rev Nutr* 2005;25:341–390. [PubMed: 16011471]
55. Delaval K, Feil R. Epigenetic regulation of mammalian genomic imprinting. *Curr Opin Genet Dev* 2004;14:188–195. [PubMed: 15196466]
56. Sneider TW. Methylation of mammalian deoxyribonucleic acid. II. The distribution of 5-methylcytosine in pyrimidine deoxyribonucleotide clusters in Novikoff hepatoma cell deoxyribonucleic acid. *J Biol Chem* 1971;246:4774–4783. [PubMed: 4327327]
57. Esteller M. CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. *Oncogene* 2002;21:5427–5440. [PubMed: 12154405]
58. Prokhorchouk E, Hendrich B. Methyl-CpG binding proteins and cancer: are MeCpGs more important than MBDs? *Oncogene* 2002;21:5394–5399. [PubMed: 12154402]
59. Cho KS, Elizondo LI, Boerkoel CF. Advances in chromatin remodeling and human disease. *Curr Opin Genet Dev* 2004;14:308–315. [PubMed: 15172675]
60. Go VL, Butrum RR, Wong DA. Diet, nutrition, and cancer prevention: the postgenomic era. *J Nutr* 2003;133:3830S–3836S. [PubMed: 14608122]
61. Cooney CA, Dave AA, Wolff GL. Maternal methyl supplements in mice affect epigenetic variation and DNA methylation of offspring. *J Nutr* 2002;132:2393S–2400S. [PubMed: 12163699]
62. Mason JB. Biomarkers of nutrient exposure and status in one-carbon (methyl) metabolism. *J Nutr* 2003;133:941S–947S. [PubMed: 12612180]
63. Stover PJ, Garza C. Bringing individuality to public health recommendations. *J Nutr* 2002;132:2476S–2480S. [PubMed: 12163715]
64. Blander G, Guarente L. The Sir2 family of protein deacetylases. *Annu Rev Biochem* 2004;73:417–435. [PubMed: 15189148]
65. Magbanua, MJM.; Dawson, K.; Huang, L.; Malyj, W.; Gregg, J.; Galvez, A.; Rodriguez, RL. Nutrient - Gene Interactions involving soy Peptide and Chemopreventive Genes in Prostate Epithelial Cells. In: Kaput, J.; Rodriguez, RL., editors. *Nutritional Genomics Discovering the Path to Personalized Nutrition*. John Wiley and Sons; Hoboken, NJ: 2006. p. 255-276.
66. Young VR, Scrimshaw NS. Genetic and biological variability in human nutrient requirements. *Am J Clin Nutr* 1979;32:486–500. [PubMed: 420139]
67. Ottman R. Gene-environment interaction: definitions and study designs. *Prev Med* 1996;25:764–770. [PubMed: 8936580]
68. Corella D, Ordovas JM. The metabolic syndrome: a crossroad for genotype-phenotype associations in atherosclerosis. *Curr Atheroscler Rep* 2004;6:186–196. [PubMed: 15068743]
69. Memisoglu A, Hu FB, Hankinson SE, Manson JE, De Vivo I, Willett WC, Hunter DJ. Interaction between a peroxisome proliferator-activated receptor gamma gene polymorphism and dietary fat intake in relation to body mass. *Hum Mol Genet* 2003;12:2923–2929. [PubMed: 14506127]
70. Roche HM, Phillips C, Gibney MJ. The metabolic syndrome: the crossroads of diet and genetics. *Proc Nutr Soc* 2005;64:371–377. [PubMed: 16048671]
71. Roche HM. Fatty acids and the metabolic syndrome. *Proc Nutr Soc* 2005;64:23–29. [PubMed: 15877919]
72. Dawson, K.; Rodriguez, RL.; Hawkes, WC.; Malyj, W. Biocomputation and the Analysis of Complex Data Sets in Nutritional Genomics. In: Kaput, J.; Rodriguez, RL., editors. *Nutritional Genomics Discovering the Path to Personalized Nutrition*. John Wiley and Sons; Hoboken, NJ: 2006. p. 375-402.
73. Sebastiani P, Gussoni E, Kohan IS, Ramoni MF. Statistical Challenges in Functional Genomics. *Statistical Science* 2003;18:33–70.

74. Motsinger AA, Ritchie MD. Multifactor dimensionality reduction: an analysis strategy for modelling and detecting gene-gene interactions in human genetics and pharmacogenomics studies. *Hum Genomics* 2006;2:318–328. [PubMed: 16595076]
75. Ritchie MD. Bioinformatics approaches for detecting gene-gene and gene-environment interactions in studies of human disease. *Neurosurg Focus* 2005;19:E2. [PubMed: 16241104]
76. van Ommen B. Nutrigenomics: exploiting systems biology in the nutrition and health arenas. *Nutrition* 2004;20:4–8. [PubMed: 14698007]
77. West M, Ginsburg GS, Huang AT, Nevins JR. Embracing the complexity of genomic data for personalized medicine. *Genome Res* 2006;16:559–566. [PubMed: 16651662]
78. Feitosa MF, Province MA, Heiss G, Arnett DK, Myers RH, Pankow JS, Hopkins PN, Borecki IB. Evidence of QTL on 15q21 for high-density lipoprotein cholesterol: The National Heart, Lung, and Blood Institute Family Heart Study (NHLBI FHS). *Atherosclerosis* 2007;190:232–237. [PubMed: 16529751]
79. Wilke RA, Reif DM, Moore JH. Combinatorial Pharmacogenetics. *Nature Reviews Drug Discovery* 2005;4:911–918.
80. Dawson K, Rodriguez RL, Malyj W. Sample phenotype clusters in high-density oligonucleotide microarray data sets are revealed using Isomap, a nonlinear algorithm. *BMC Bioinformatics* 2005;6:195. [PubMed: 16076401]
81. Williams SM, Ritchie MD, Phillips JA 3rd, Dawson E, Prince M, Dzhura E, Willis A, Semenza A, Summar M, White BC, Addy JH, Kpodonu J, Wong LJ, Felder RA, Jose PA, Moore JH. Multilocus analysis of hypertension: a hierarchical approach. *Hum Hered* 2004;57:28–38. [PubMed: 15133310]
82. Moore JH. Computational analysis of gene-gene interactions using multifactor dimensionality reduction. *Expert Rev Mol Diagn* 2004;4:795–803. [PubMed: 15525222]
83. Cho YM, Ritchie MD, Moore JH, Park JY, Lee KU, Shin HD, Lee HK, Park KS. Multifactor-dimensionality reduction shows a two-locus interaction associated with Type 2 diabetes mellitus. *Diabetologia* 2004;47:549–554. [PubMed: 14730379]
84. Bastone L, Reilly M, Rader DJ, Foulkes AS. MDR and PRP: a comparison of methods for high-order genotype-phenotype associations. *Hum Hered* 2004;58:82–92. [PubMed: 15711088]
85. Lan H, Stoehr JP, Nadler ST, Schueler KL, Yandell BS, Attie AD. Dimension reduction for mapping mRNA abundance as quantitative traits. *Genetics* 2003;164:1607–1614. [PubMed: 12930764]
86. Wang J, Delabie J, Aasheim H, Smeland E, Myklebost O. Clustering of the SOM easily reveals distinct gene expression patterns: results of a reanalysis of lymphoma study. *BMC Bioinformatics* 2002;3:36. [PubMed: 12445336]
87. Barroso I, Luan J, Middelberg RP, Harding AH, Franks PW, Jakes RW, Clayton D, Schafer AJ, O'Rahilly S, Wareham NJ. Candidate gene association study in type 2 diabetes indicates a role for genes involved in beta-cell function as well as insulin action. *PLoS Biol* 2003;1:E20. [PubMed: 14551916]
88. Hani EH, Clement K, Velho G, Vionnet N, Hager J, Philippi A, Dina C, Inoue H, Permutt MA, Basdevant A, North M, Demenais F, Guy-Grand B, Froguel P. Genetic studies of the sulfonylurea receptor gene locus in NIDDM and in morbid obesity among French Caucasians. *Diabetes* 1997;46:688–694. [PubMed: 9075812]
89. Lucarini N, Antonacci E, Bottini N, Borgiani P, Faggioni G, Gloria-Bottini F. Phosphotyrosine-protein-phosphatase and diabetic disorders. Further studies on the relationship between low molecular weight acid phosphatase genotype and degree of glycemic control. *Dis Markers* 1998;14:121–125. [PubMed: 9868599]
90. Bottini E, Gloria-Bottini F. Adenosine deaminase and body mass index in non-insulin-dependent diabetes mellitus. *Metabolism* 1999;48:949–951. [PubMed: 10459555]
91. Gonzalez Sanchez JL, Proenza AM, Martinez Larrad MT, Ramis JM, Fernandez Perez C, Palou A, Serrano Rios M. The glutamine 27 glutamic acid polymorphism of the beta2-adrenoceptor gene is associated with abdominal obesity and greater risk of impaired glucose tolerance in men but not in women: a population-based study in Spain. *Clin Endocrinol (Oxf)* 2003;59:476–481. [PubMed: 14510910]
92. Ellsworth DL, Coady SA, Chen W, Srinivasan SR, Elkasabany A, Gustat J, Boerwinkle E, Berenson GS. Influence of the beta2-adrenergic receptor Arg16Gly polymorphism on longitudinal changes in

- obesity from childhood through young adulthood in a biracial cohort: the Bogalusa Heart Study. *Int J Obes Relat Metab Disord* 2002;26:928–937. [PubMed: 12080445]
93. Barroso I. Genetics of Type 2 diabetes. *Diabet Med* 2005;22:517–535. [PubMed: 15842505]
 94. Marks DL, Boucher N, Lanouette CM, Perusse L, Brookhart G, Comuzzie AG, Chagnon YC, Cone RD. Ala67Thr polymorphism in the Agouti-related peptide gene is associated with inherited leanness in humans. *Am J Med Genet A* 2004;126:267–271. [PubMed: 15054840]
 95. Fumeron F, Aubert R, Siddiq A, Betoulle D, Pean F, Hadjadj S, Tichet J, Wilpart E, Chesnier MC, Balkau B, Froguel P, Marre M. Adiponectin gene polymorphisms and adiponectin levels are independently associated with the development of hyperglycemia during a 3-year period: the epidemiologic data on the insulin resistance syndrome prospective study. *Diabetes* 2004;53:1150–1157. [PubMed: 15047634]
 96. Gu HF, Abulaiti A, Ostenson CG, Humphreys K, Wahlestedt C, Brookes AJ, Efendic S. Single nucleotide polymorphisms in the proximal promoter region of the adiponectin (APM1) gene are associated with type 2 diabetes in Swedish caucasians. *Diabetes* 2004;53:S31–35. [PubMed: 14749263]
 97. Leipold H, Knofler M, Gruber C, Haslinger P, Bancher-Todesca D, Worda C. Calpain-10 haplotype combination and association with gestational diabetes mellitus. *Obstet Gynecol* 2004;103:1235–1240. [PubMed: 15172858]
 98. Carlsson E, Fredriksson J, Groop L, Ridderstrale M. Variation in the calpain-10 gene is associated with elevated triglyceride levels and reduced adipose tissue messenger ribonucleic acid expression in obese Swedish subjects. *J Clin Endocrinol Metab* 2004;89:3601–3605. [PubMed: 15240652]
 99. Hamaguchi K, Terao H, Kusuda Y, Yamashita T, Hazoury Bahles JA, Cruz LM, Brugal VL, Jongchong WB, Yoshimatsu H, Sakata T. The PC-1 Q121 allele is exceptionally prevalent in the Dominican Republic and is associated with type 2 diabetes. *J Clin Endocrinol Metab* 2004;89:1359–1364. [PubMed: 15001634]
 100. Weiss EP, Brown MD, Shuldiner AR, Hagberg JM. Fatty acid binding protein-2 gene variants and insulin resistance: gene and gene-environment interaction effects. *Physiol Genomics* 2002;10:145–157. [PubMed: 12209017]
 101. Gertow K, Bellanda M, Eriksson P, Boquist S, Hamsten A, Sunnerhagen M, Fisher RM. Genetic and structural evaluation of fatty acid transport protein-4 in relation to markers of the insulin resistance syndrome. *J Clin Endocrinol Metab* 2004;89:392–399. [PubMed: 14715877]
 102. Carlsson E, Groop L, Ridderstrale M. Role of the FOXC2 -512C>T polymorphism in type 2 diabetes: possible association with the dysmetabolic syndrome. *Int J Obes Relat Metab Disord* 2005;29:268–274.
 103. Holmkvist J, Almgren P, Parikh H, Zucchelli M, Kere J, Groop L, Lindgren CM. Haplotype construction of the FRDA gene and evaluation of its role in type II diabetes. *Eur J Hum Genet* 2005;13:849–855. [PubMed: 15827563]
 104. Hirai M, Suzuki S, Hinokio Y, Hirai A, Chiba M, Akai H, Suzuki C, Toyota T. Variations in vitamin D-binding protein (group-specific component protein) are associated with fasting plasma insulin levels in Japanese with normal glucose tolerance. *J Clin Endocrinol Metab* 2000;85:1951–1953. [PubMed: 10843180]
 105. Marz W, Nauck M, Hoffmann MM, Nagel D, Boehm BO, Koenig W, Rothenbacher D, Winkelmann BR. G(-30)A polymorphism in the pancreatic promoter of the glucokinase gene associated with angiographic coronary artery disease and type 2 diabetes mellitus. *Circulation* 2004;109:2844–2849. [PubMed: 15173029]
 106. Gidh-Jain M, Takeda J, Xu LZ, Lange AJ, Vionnet N, Stoffel M, Froguel P, Velho G, Sun F, Cohen D. Glucokinase mutations associated with non-insulin-dependent (type 2) diabetes mellitus have decreased enzymatic activity: implications for structure/function relationships. *Proc Natl Acad Sci U S A* 1993;90:1932–1936. [PubMed: 8446612]
 107. Takeda J, Gidh-Jain M, Xu LZ, Froguel P, Velho G, Vaxillaire M, Cohen D, Shimada F, Makino H, Nishi S, et al. Structure/function studies of human beta-cell glucokinase. Enzymatic properties of a sequence polymorphism, mutations associated with diabetes, and other site-directed mutants. *J Biol Chem* 1993;268:15200–15204. [PubMed: 8325892]

108. Zhang H, Jia Y, Cooper JJ, Hale T, Zhang Z, Elbein SC. Common variants in glutamine:fructose-6-phosphate amidotransferase 2 (GFPT2) gene are associated with type 2 diabetes, diabetic nephropathy, and increased GFPT2 mRNA levels. *J Clin Endocrinol Metab* 2004;89:748–755. [PubMed: 14764791]
109. Miraglia del Giudice E, Santoro N, Cirillo G, Raimondo P, Grandone A, D'Aniello A, Di Nardo M, Perrone L. Molecular screening of the ghrelin gene in Italian obese children: the Leu72Met variant is associated with an earlier onset of obesity. *Int J Obes Relat Metab Disord* 2004;28:447–450. [PubMed: 14724664]
110. Korbonits M, Gueorguiev M, O'Grady E, Lecoeur C, Swan DC, Mein CA, Weill J, Grossman AB, Froguel P. A variation in the ghrelin gene increases weight and decreases insulin secretion in tall, obese children. *J Clin Endocrinol Metab* 2002;87:4005–4008. [PubMed: 12161552]
111. Poston WS, Haddock CK, Spertus J, Catanese DM, Pavlik VN, Hyman DJ, Hanis CL, Forevt JP. Physical activity does not mitigate G-protein-related genetic risk for obesity in individuals of African descent. *Eat Weight Disord* 2002;7:68–71. [PubMed: 11930987]
112. Boutin P, Gresh L, Cisse A, Hara M, Bell G, Babu S, Eisenbarth G, Froguel P. Missense mutation Gly574Ser in the transcription factor HNF-1alpha is a marker of atypical diabetes mellitus in African-American children. *Diabetologia* 1999;42:380–381. [PubMed: 10096793]
113. Gasa R, Gomis R, Casamitjana R, Rivera F, Novials A. Glucose regulation of islet amyloid polypeptide gene expression in rat pancreatic islets. *Am J Physiol* 1997;272:E543–549. [PubMed: 9142872]
114. Ritzel RA, Butler PC. Replication increases beta-cell vulnerability to human islet amyloid polypeptide-induced apoptosis. *Diabetes* 2003;52:1701–1708. [PubMed: 12829636]
115. Arends N, Johnston L, Hokken-Koelega A, van Duijn C, de Ridder M, Savage M, Clark A. Polymorphism in the IGF-I gene: clinical relevance for short children born small for gestational age (SGA). *J Clin Endocrinol Metab* 2002;87:2720. [PubMed: 12050240]
116. Mohlig M, Boeing H, Spranger J, Osterhoff M, Kroke A, Fisher E, Bergmann MM, Ristow M, Hoffmann K, Pfeiffer AF. Body mass index and C-174G interleukin-6 promoter polymorphism interact in predicting type 2 diabetes. *J Clin Endocrinol Metab* 2004;89:1885–1890. [PubMed: 15070960]
117. McKenzie JA, Weiss EP, Ghiu IA, Kulaputana O, Phares DA, Ferrell RE, Hagberg JM. Influence of the interleukin-6 -174 G/C gene polymorphism on exercise training-induced changes in glucose tolerance indexes. *J Appl Physiol* 2004;97:1338–1342. [PubMed: 15180970]
118. Kubaszek A, Pihlajamaki J, Komarovski V, Lindi V, Lindstrom J, Eriksson J, Valle TT, Hamalainen H, Ilanne-Parikka P, Keinanen-Kiukaanniemi S, Tuomilehto J, Uusitupa M, Laakso M. Promoter polymorphisms of the TNF-alpha (G-308A) and IL-6 (C-174G) genes predict the conversion from impaired glucose tolerance to type 2 diabetes: the Finnish Diabetes Prevention Study. *Diabetes* 2003;52:1872–1876. [PubMed: 12829659]
119. Hassan MI, Aschner Y, Manning CH, Xu J, Aschner JL. Racial differences in selected cytokine allelic and genotypic frequencies among healthy, pregnant women in North Carolina. *Cytokine* 2003;21:10–16. [PubMed: 12668154]
120. Sanchez-Corona J, Flores-Martinez SE, Machorro-Lazo MV, Galaviz-Hernandez C, Moran-Moguel MC, Perea FJ, Mujica-Lopez KI, Vargas-Ancona L, Laviada-Molina HA, Fernandez V, Pardio J, Arroyo P, Barrera H, Hanson RL. Polymorphisms in candidate genes for type 2 diabetes mellitus in a Mexican population with metabolic syndrome findings. *Diabetes Res Clin Pract* 2004;63:47–55. [PubMed: 14693412]
121. Undlien DE, Bennett ST, Todd JA, Akselsen HE, Ikaheimo I, Reijonen H, Knip M, Thorsby E, Ronningen KS. Insulin gene region-encoded susceptibility to IDDM maps upstream of the insulin gene. *Diabetes* 1995;44:620–625. [PubMed: 7789624]
122. Matejkova-Behanova M, Vankova M, Hill M, Kucera P, Cinek O, Andel M, Bendlova B. Polymorphism of INS VNTR is associated with glutamic acid decarboxylase antibodies and postprandial C-peptide in patients with onset of diabetes after 35 years of age. *Physiol Res* 2004;53:187–190. [PubMed: 15046555]
123. Heude B, Dubois S, Charles MA, Deweirder M, Dina C, Borys JM, Ducimetiere P, Froguel P. VNTR polymorphism of the insulin gene and childhood overweight in a general population. *Obes Res* 2004;12:499–504. [PubMed: 15044667]

124. Macfarlane WM, Frayling TM, Ellard S, Evans JC, Allen LI, Bulman MP, Ayres S, Shepherd M, Clark P, Millward A, Demaine A, Wilkin T, Docherty K, Hattersley AT. Missense mutations in the insulin promoter factor-1 gene predispose to type 2 diabetes. *J Clin Invest* 1999;104:R33–39. [PubMed: 10545530]
125. Hani EH, Stoffers DA, Chevre JC, Durand E, Stanojevic V, Dina C, Habener JF, Froguel P. Defective mutations in the insulin promoter factor-1 (IPF-1) gene in late-onset type 2 diabetes mellitus. *J Clin Invest* 1999;104:R41–48. [PubMed: 10545531]
126. Yamada K, Yuan X, Ishiyama S, Ichikawa F, Kohno S, Shoji S, Hayashi H, Nonaka K. Identification of a single nucleotide insertion polymorphism in the upstream region of the insulin promoter factor-1 gene: an association study with diabetes mellitus. *Diabetologia* 1998;41:603–605. [PubMed: 9628281]
127. Mammarella S, Romano F, Di Valerio A, Creati B, Esposito DL, Palmirotta R, Capani F, Vitullo P, Volpe G, Battista P, Della Loggia F, Mariani-Costantini R, Cama A. Interaction between the G1057D variant of IRS-2 and overweight in the pathogenesis of type 2 diabetes. *Hum Mol Genet* 2000;9:2517–2521. [PubMed: 11030756]
128. Nielsen EM, Hansen L, Carstensen B, Echwald SM, Drivsholm T, Glumer C, Thorsteinsson B, Borch-Johnsen K, Hansen T, Pedersen O. The E23K variant of Kir6.2 associates with impaired post-OGTT serum insulin response and increased risk of type 2 diabetes. *Diabetes* 2003;52:573–577. [PubMed: 12540638]
129. Todorova B, Kubaszek A, Pihlajamaki J, Lindstrom J, Eriksson J, Valle TT, Hamalainen H, Ilanne-Parikka P, Keinanen-Kiukkaanniemi S, Tuomilehto J, Uusitupa M, Laakso M. The G-250A promoter polymorphism of the hepatic lipase gene predicts the conversion from impaired glucose tolerance to type 2 diabetes mellitus: the Finnish Diabetes Prevention Study. *J Clin Endocrinol Metab* 2004;89:2019–2023. [PubMed: 15126514]
130. Yabu Y, Noma K, Nakatani K, Nishioka J, Suematsu M, Katsuki A, Hori Y, Yano Y, Sumida Y, Wada H, Nobori T. C-514T polymorphism in hepatic lipase gene promoter is associated with elevated triglyceride levels and decreasing insulin sensitivity in nondiabetic Japanese subjects. *Int J Mol Med* 2005;16:421–425. [PubMed: 16077949]
131. Ma YQ, Thomas GN, Ng MC, Critchley JA, Chan JC, Tomlinson B. The lipoprotein lipase gene HindIII polymorphism is associated with lipid levels in early-onset type 2 diabetic patients. *Metabolism* 2003;52:338–343. [PubMed: 12647273]
132. Malecki MT, Jhala US, Antonellis A, Fields L, Doria A, Orban T, Saad M, Warram JH, Montminy M, Krolewski AS. Mutations in NEUROD1 are associated with the development of type 2 diabetes mellitus. *Nat Genet* 1999;23:323–328. [PubMed: 10545951]
133. Ruiz-Quezada S, Vazquez-Del Mercado M, Parra-Rojas I, Rangel-Villalobos H, Best-Aguilera C, Sanchez-Orozco LV, Munoz-Valle JF. Genotype and allele frequency of PAI-1 promoter polymorphism in healthy subjects from the west of Mexico. Association with biochemical and hematological parameters. *Ann Genet* 2004;47:155–162. [PubMed: 15183748]
134. Lopes C, Dina C, Durand E, Froguel P. PAI-1 polymorphisms modulate phenotypes associated with the metabolic syndrome in obese and diabetic Caucasian population. *Diabetologia* 2003;46:1284–1290. [PubMed: 12856128]
135. Hoffstedt J, Andersson IL, Persson L, Isaksson B, Arner P. The common -675 4G/5G polymorphism in the plasminogen activator inhibitor -1 gene is strongly associated with obesity. *Diabetologia* 2002;45:584–587. [PubMed: 12032637]
136. Cao H, van der Veer E, Ban MR, Hanley AJ, Zinman B, Harris SB, Young TK, Pickering JG, Hegele RA. Promoter polymorphism in PCK1 (phosphoenolpyruvate carboxykinase gene) associated with type 2 diabetes mellitus. *J Clin Endocrinol Metab* 2004;89:898–903. [PubMed: 14764811]
137. Muller YL, Bogardus C, Pedersen O, Baier L. A Gly482Ser missense mutation in the peroxisome proliferator-activated receptor gamma coactivator-1 is associated with altered lipid oxidation and early insulin secretion in Pima Indians. *Diabetes* 2003;52:895–898. [PubMed: 12606537]
138. Hara K, Tobe K, Okada T, Kadowaki H, Akanuma Y, Ito C, Kimura S, Kadowaki T. A genetic variation in the PGC-1 gene could confer insulin resistance and susceptibility to Type II diabetes. *Diabetologia* 2002;45:740–743. [PubMed: 12107756]
139. Ek J, Andersen G, Urhammer SA, Gaede PH, Drivsholm T, Borch-Johnsen K, Hansen T, Pedersen O. Mutation analysis of peroxisome proliferator-activated receptor-gamma coactivator-1 (PGC-1)

- and relationships of identified amino acid polymorphisms to Type II diabetes mellitus. *Diabetologia* 2001;44:2220–2226. [PubMed: 11793024]
140. Hansen L, Zethelius B, Berglund L, Reneland R, Hansen T, Berne C, Lithell H, Hemmings BA, Pedersen O. In vitro and in vivo studies of a naturally occurring variant of the human p85alpha regulatory subunit of the phosphoinositide 3-kinase: inhibition of protein kinase B and relationships with type 2 diabetes, insulin secretion, glucose disappearance constant, and insulin sensitivity. *Diabetes* 2001;50:690–693. [PubMed: 11246893]
 141. Fanella S, Harris SB, Young TK, Hanley AJ, Zinman B, Connelly PW, Hegele RA. Association between PON1 L/M55 polymorphism and plasma lipoproteins in two Canadian aboriginal populations. *Clin Chem Lab Med* 2000;38:413–420. [PubMed: 10952224]
 142. Kao WH, Coresh J, Shuldiner AR, Boerwinkle E, Bray MS, Brancati FL. Pro12Ala of the peroxisome proliferator-activated receptor-gamma2 gene is associated with lower serum insulin levels in nonobese African Americans: the Atherosclerosis Risk in Communities Study. *Diabetes* 2003;52:1568–1572. [PubMed: 12765972]
 143. Doney A, Fischer B, Frew D, Cumming A, Flavell DM, World M, Montgomery HE, Boyle D, Morris A, Palmer CN. Haplotype analysis of the PPARgamma Pro12Ala and C1431T variants reveals opposing associations with body weight. *BMC Genet* 2002;3:21. [PubMed: 12429071]
 144. Doney AS, Fischer B, Cecil JE, Boylan K, McGuigan FE, Ralston SH, Morris AD, Palmer CN. Association of the Pro12Ala and C1431T variants of PPARG and their haplotypes with susceptibility to Type 2 diabetes. *Diabetologia* 2004;47:555–558. [PubMed: 14730381]
 145. Permana PA, Mott DM. Genetic analysis of human type 1 protein phosphatase inhibitor 2 in insulin-resistant Pima Indians. *Genomics* 1997;41:110–114. [PubMed: 9126490]
 146. Xia J, Scherer SW, Cohen PT, Majer M, Xi T, Norman RA, Knowler WC, Bogardus C, Prochazka M. A common variant in PPP1R3 associated with insulin resistance and type 2 diabetes. *Diabetes* 1998;47:1519–1524. [PubMed: 9726244]
 147. Wang H, Chu W, Das SK, Zheng Z, Hasstedt SJ, Elbein SC. Molecular screening and association studies of retinoid-related orphan receptor gamma (RORC): a positional and functional candidate for type 2 diabetes. *Mol Genet Metab* 2003;79:176–182. [PubMed: 12855222]
 148. Garvey WT, Maianu L, Kennedy A, Wallace P, Ganaway E, Hamacher LL, Yarnall DP, Lenhard JM, Burns DK. Muscle Rad expression and human metabolism: potential role of the novel Ras-related GTPase in energy expenditure and body composition. *Diabetes* 1997;46:444–450. [PubMed: 9032101]
 149. Grant SF, Thorleifsson G, Reynisdottir I, Benediktsson R, Manolescu A, Sainz J, Helgason A, Stefansson H, Emilsson V, Helgadottir A, Styrkarsdottir U, Magnusson KP, Walters GB, Palsdottir E, Jonsdottir T, Gudmundsdottir T, Gylfason A, Saemundsdottir J, Wilensky RL, Reilly MP, Rader DJ, Bagger Y, Christiansen C, Gudnason V, Sigurdsson G, Thorsteinsdottir U, Gulcher JR, Kong A, Stefansson K. Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet.* 2006
 150. Sesti G, Cardellini M, Marini MA, Frontoni S, D'Adamo M, Del Guerra S, Lauro D, De Nicolais P, Sbraccia P, Del Prato S, Gambardella S, Federici M, Marchetti P, Lauro R. A Common Polymorphism in the Promoter of UCP2 Contributes to the Variation in Insulin Secretion in Glucose-Tolerant Subjects. *Diabetes* 2003;52:1280–1283. [PubMed: 12716765]
 151. Le Fur S, Le Stunff C, Dos Santos C, Bougneres P. The common -866 G/A polymorphism in the promoter of uncoupling protein 2 is associated with increased carbohydrate and decreased lipid oxidation in juvenile obesity. *Diabetes* 2004;53:235–239. [PubMed: 14693721]
 152. D'Adamo M, Perego L, Cardellini M, Marini MA, Frontoni S, Andreozzi F, Sciacqua A, Lauro D, Sbraccia P, Federici M, Paganelli M, Pontiroli AE, Lauro R, Perticone F, Folli F, Sesti G. The -866A/A genotype in the promoter of the human uncoupling protein 2 gene is associated with insulin resistance and increased risk of type 2 diabetes. *Diabetes* 2004;53:1905–1910. [PubMed: 15220218]
 153. Mantzoros CS, Williams CJ, Manson JE, Meigs JB, Hu FB. Adherence to the Mediterranean dietary pattern is positively associated with plasma adiponectin concentrations in diabetic women. *Am J Clin Nutr* 2006;84:328–335. [PubMed: 16895879]
 154. Cardillo S, Seshadri P, Iqbal N. The effects of a low-carbohydrate versus low-fat diet on adipocytokines in severely obese adults: three-year follow-up of a randomized trial. *Eur Rev Med Pharmacol Sci* 2006;10:99–106. [PubMed: 16875041]

155. Butte NF, Comuzzie AG, Cai G, Cole SA, Mehta NR, Bacino CA. Genetic and environmental factors influencing fasting serum adiponectin in Hispanic children. *J Clin Endocrinol Metab* 2005;90:4170–4176. [PubMed: 15827100]
156. Eisenach JH, Schroeder DR, Pike TL, Johnson CP, Schrage WG, Snyder EM, Johnson BD, Garovic VD, Turner ST, Joyner MJ. Dietary sodium restriction and beta2-adrenergic receptor polymorphism modulate cardiovascular function in humans. *J Physiol* 2006;574:955–965. [PubMed: 16740612]
157. Loos RJ, Rankinen T, Rice T, Rao DC, Leon AS, Skinner JS, Bouchard C, Argyropoulos G. Two ethnic-specific polymorphisms in the human Agouti-related protein gene are associated with macronutrient intake. *Am J Clin Nutr* 2005;82:1097–1101. [PubMed: 16280444]
158. Tozawa Y. G protein beta3 subunit variant: tendency of increasing susceptibility to hypertension in Japanese. *Blood Press* 2001;10:131–134. [PubMed: 11688759]
159. Bos G, Dekker JM, Feskens EJ, Ocke MC, Nijpels G, Stehouwer CD, Bouter LM, Heine RJ, Jansen H. Interactions of dietary fat intake and the hepatic lipase -480C-->T polymorphism in determining hepatic lipase activity: the Hoorn Study. *Am J Clin Nutr* 2005;81:911–915. [PubMed: 15817871]
160. Dos Santos C, Fallin D, Le Stunff C, LeFur S, Bougneres P. INS VNTR is a QTL for the insulin response to oral glucose in obese children. *Physiol Genomics* 2004;16:309–313. [PubMed: 14657411]
161. Sierksma A, van der Gaag MS, van Tol A, James RW, Hendriks HF. Kinetics of HDL cholesterol and paraoxonase activity in moderate alcohol consumers. *Alcohol Clin Exp Res* 2002;26:1430–1435. [PubMed: 12351939]
162. Cha MH, Shin HD, Kim KS, Lee BH, Yoon Y. The effects of uncoupling protein 3 haplotypes on obesity phenotypes and very low-energy diet-induced changes among overweight Korean female subjects. *Metabolism* 2006;55:578–586. [PubMed: 16631432]
163. Ukkola O, Tremblay A, Sun G, Chagnon YC, Bouchard C. Genetic variation at the uncoupling protein 1, 2 and 3 loci and the response to long-term overfeeding. *Eur J Clin Nutr* 2001;55:1008–1015. [PubMed: 11641751]
164. Kustra R, Shioda R, Zhu M. A factor analysis model for functional genomics. *BMC Bioinformatics* 2006;7:216. [PubMed: 16630343]
165. Taguchi YH, Oono Y. Relational patterns of gene expression via non-metric multidimensional scaling analysis. *Bioinformatics* 2005;21:730–740. [PubMed: 15509613]
166. Tan Y, Shi L, Tong W, Wang C. Multi-class cancer classification by total principal component regression (TPCR) using microarray gene expression data. *Nucleic Acids Res* 2005;33:56–65. [PubMed: 15640445]
167. Liu L, Hawkins DM, Ghosh S, Young SS. Robust singular value decomposition analysis of microarray data. *Proc Natl Acad Sci U S A* 2003;100:13167–13172. [PubMed: 14581611]
168. Niyogi MBP. Laplacian Eigenmaps for Dimensionality Reduction and Data Representation. *Neural Computation* 2003;15:1373–1396.
169. Roweis ST, Saul LK. Nonlinear dimensionality reduction by locally linear embedding. *Science* 2000;290:2323–2326. [PubMed: 11125150]
170. Tenenbaum JB, de Silva V, Langford JC. A global geometric framework for nonlinear dimensionality reduction. *Science* 2000;290:2319–2323. [PubMed: 11125149]
171. Donoho DL, Grimes C. Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data. *Proc Natl Acad Sci U S A* 2003;100:5591–5596. [PubMed: 16576753]
172. Mendez MA, Hodar C, Vulpe C, Gonzalez M, Cambiazo V. Discriminant analysis to evaluate clustering of gene expression data. *FEBS Lett* 2002;522:24–28. [PubMed: 12095613]
173. Papana A, Ishwaran H. CART variance stabilization and regularization for high-throughput genomic data. *Bioinformatics* 2006;22:2254–2261. [PubMed: 16844707]
174. Beiko RG, Charlebois RL. GANN: genetic algorithm neural networks for the detection of conserved combinations of features in DNA. *BMC Bioinformatics* 2005;6:36. [PubMed: 15725347]
175. Yang ZR. Biological applications of support vector machines. *Brief Bioinform* 2004;5:328–338. [PubMed: 15606969]

Table 1**Drug Classes for the Treatment of Type 2 Diabetes**

The number of subtypes of T2DM can be estimated by the different drugs used to treat different clinical indications of type 2 diabetes. 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.

Treatment	Target Tissue	Indications	Effectiveness
Lifestyle	All	All	15%
Sulfonylurea	Pancreas	T2DM < 5 yr	~50%
Meglitinides	Pancreas	T2DM < 5 yr & ↑ PPG ²	?
Exenatide	Pancreas	T2DM	2 nd line
Biguanides	Liver	Obese, insulin resistant	~75%
alpha-glucosidase	Intestine	↑ PPG	2 nd line
Thiazolidinediones	Adipose, muscle	Obese, insulin resistant	2 nd line

Candidate Type 2 Diabetes Mellitus Genes

Incomplete list of candidate genes identified in literature searches for gene – T2DM genetic association studies.

Table 2

Genetic	Common Name	Function	Chromosome	References
ABCC8	Sulfonylurea receptor	Potassium channel	11p15.1	[1,14,87,88]
ACPI	Acid Phosphatase 1, soluble	Phosphatase	2p25	[89]
ADA	Adenosine deaminase	Enzyme, Purine catabolic pathway	20q13.11 ^a	[90]
ADRB2	2-Adrenergic Receptor	Receptor linked to catecholamine, obesity	5q32-q34 ^a	[91-93]
ADRB3	3-Adrenergic Receptor	Receptor, lipolysis regulation	8p12-p11.2	[9]
AGRP	Agouti related protein (homolog of mouse agouti)	Signaling, melanocortin antagonist	16q22	[94]
APM1	Adiponectin (ACDC)	Adipocyte hormone	3q27 ^b	[9,95,96]
CAPN10	Calpain 10	Cysteine protease	2q37.3 ^b	[97,98]
ENPP1	Glycoprotein PC-1	Inhibits insulin signaling	6q22-q23	[99,99]
FABP2	Liver fatty acid binding protein	Long chain fatty acid transport protein	4q28-q31	[100]
FATP4	Fatty acid transporter, SLC27A4	Long chain fatty acid transport protein (RBC)	Chr. 9	[101]
FOXc2	Transcription Factor	Regulator of adipocyte metabolism	16q24.3	[9,102]
FRDA	Frataxin	Mitochondrial iron metabolism	9q13	[103]
GC	Group specific component, Vitamin D binding protein	Vit D involved in regulating insulin levels	4q12	[93,104]
GCGR	Glucagon receptor	Glucose homeostasis	17q25 ^a	[14]
GCK	Glucokinase, liver	Enzyme, first step in glycolysis	7p15-p13 ^a	[105-107]
GFTPT2	Glutamine:fructose 6-phosphate amidotransferase 2	Hexosamine biosynthesis	5q34-q35 ^a	[108]
GHRL	Ghrelin	Hormone, Energy homeostasis and feeding	3p26-p25	[109,110]
GNB3	Guanine nucleotide binding protein 3	Signaling, obesity	12p13	[111]
GYS1	Glycogen synthase	Enzyme, impaired glycogen synthesis	19q13.3 ^a	[9]
HNFI	Hepatic nuclear factor 1	Transcription factor, cholesterol homeostasis	12q24.2 ^a	[112]
HNFI4A	Hepatic nuclear factor 4	Transcription factor, hepatic glycogen stores	20q12-q13.1 ^c	[87]
IAPP	Insulin amyloid protein, Amylin	Hormone, glucose uptake pancreas	12p12.3-p12.1	[113,114]
IGF1	Insulin growth factor 1	Hormone, growth	12q22-q24.1 ^a	[115]
IL6	Interleukin 6	Cytokine	7p21	[116-119]
INS	Variable number tandem repeat in the insulin gene	Glucose regulation	11p15.5	[120-123]
INSR	Insulin receptor	Receptor	19p13.2	[87]
IPF1	Insulin promoter factor 1	Binds to promoters	12q12.1	[124-126]
IRS1	Insulin receptor substrate 1	Signal transduction	2q36 ^b	[120]
IRS2	Insulin receptor substrate 2	Signal transduction	13q34	[127]
KCNJ11	Potassium inward rectifier channel Kir6.2	Potassium channel	11p15.1	[128]
LIPC	Hepatic lipase	Lipid, lipoprotein regulation	15q21-q23	[129,130]
LIPE	Hormone sensitive lipase	Mobilization of fatty acids	19q13.1-q13.2	[9]
LPL	Lipoprotein lipase	Enzyme; chylomicrons and triglyceride	8p22	[131]
MAPK8IP1	Mitogen activated protein kinase 8 – interacting protein	Signal transduction	11p12p11.2 ^a	OMIM 604641
NeuroD1	NeuroD/BETA2	Transcription factor, development	2q32 ^b	[132]
PAI1	Plasminogen activating inhibitor	Control point in coagulation	7q21.2-q22	[133-135]
PC1	Pachonychia congenita 1	Inhibits insulin signaling	5q15-q21	[99]
PCK1	Phosphoenolpyruvate carboxykinase 1	Enzyme, regulation of gluconeogenesis	20q13.31 ^a	[136]
PGC1	Peroxisome proliferators-activated receptor coactivator - 1	Transcriptional coactivator	4p15.1	[137-139]
PIK3R1	Phosphoinositide-3-kinase regulatory subunit p85	Glucose clearance	5q13	[140]
PON2	Paraoxonase 2	High fasting glucose	7q21.3	[141]
PPARG	Peroxisome proliferator-activated receptor- gamma 2	Lipid and glucose regulation	3p25 ^b	[142-144]
PP1R3A	Protein phosphatase 1, regulatory (inhibitor) subunit 3A	Glycogen metabolism	7q11.23-q21.11	[13,145,146]
RORC	RAR-related orphan receptor C	Nuclear hormone, immune response	1q21	[147]
RRAD	Ras-related associated with diabetes	Insulin sensitivity	16q22	[93,148]
SLC2A2	GLUT2 glucose transporter	Glucose transporter	3q26.1-q26.3 ^b	[87]
SLC2A4	GLUT4 glucose transporter	Glucose transporter	17p13 ^a	OMIM 138190
SOS1	Son of sevenless homolog	Guanine nucleotide exchange factor	2p22-p21	[87]
TCF7L2	Transcription factor	Blood glucose homeostasis	T of rs7903146 & T of rs12255372	[149]
TNF	Tumor necrosis factor	Proinflammatory cytokine	6p21.3	[118]
UCP2	Uncoupling protein 2	Mitochondrial transporter	11q13	[150-152]

^aMap position overlaps QTL from diabetes search at <http://www.ncbi.nlm.nih.gov/mapview>

^bMap position overlaps QTL shown in Figure 1

^cMap position overlaps QTL with near suggestive LOD score [14].

Adapted from [4]

Table 3**Examples of Gene – Nutrient Interactions**

Examples of nutrient – gene interactions.

Gene	Nutrient Variable	Ref
Adiponectin	Mediterranean diet	[153]
	low fat diets	[154]
	macronutrient intake	[155]
Adrenergic receptors	sodium intake	[156]
AGRP	macronutrient intake	[157]
GNB3	sodium intake	[158]
Hepatic lipase	fat	[159]
INS	glucose intake	[160]
PON	alcohol	[161]
PPAR γ	fat	[69]
UCP	energy and body weight	[162]
	chronic overfeeding	[163]

Table 4

Selected Methods of Complex, High Dimensional Data

See text for details.

Unsupervised Learning (structure detection)	
Hierarchical Clustering (HC)	Clusters items based on a similarity metric and organizes them into leaves along branches of a dendrogram
Factor analysis	A multivariate technique that aims to summarize a large number of variables with a small number of factors
Multi-dimensional Scaling (MDS)	MDS is a dimensionality reduction method that tries to preserve the interpoint distance between points in the low-dimensional space
Principal Component Analysis (PCA)	Find the n dimensions that explain most of the variance.
Singular Value Decomposition (SVD)	
Laplacian Eigenmaps	A computationally effective algorithm that has a natural connection to clustering
Isomap	Nonlinear embedding; uses geodesic distances
Locally Linear Embedding (LLE)	Similar to ISOMAP but differs in focus on analyzing distances to nearest neighbors and mapping them to a smooth nonlinear manifold of lower dimensionality
Hessian Eigenmaps (HLE)	Refinement of the original Isomap algorithm
Kohonen Self-organizing Maps (SOM)	A single-layer feed-forward ANN used for visualizing high-dimensional data
Supervised Learning (structure confirmation)	
Linear Discriminant Analysis (LDA)	LDA aims to identify the best linear combination of features that can separate classes of samples
Support Vector Machines (SVM)	SVMs maximize the geometric margin while minimizing the classification error (maximum margin classifiers)
Classification analysis and regression tree (CART)	Analytic procedure for predicting the values of a continuous response variable or categorical response variable from continuous or categorical predictors
Artificial Neural Networks (ANN)	A form of multiprocessor computer system, with simple processing elements, a high degree of interconnection, simple scalar messages, and adaptive interaction between elements.