National Institute of Standards and Technology

Technology Innovation Program White Paper

Critical National Need	Developing Strategies, Infrastructure & Knowledge for Equitable, Personalized Healthcare
Submitting Organization	Division of Personalized Nutrition and Medicine National Center for Toxicological Research Food and Drug Administration
Contributing Organizations	FDA/Center for Food Safety and Nutrition USDA – ARS Tufts University Boys, Girls, and Adults Community Development Center NuGO (European Nutrigenomics Organization) Department of Nutrition, University of Toronto Human Variome Project
Date	15 January 2009

Contacts

Jim Kaput
Director
Division of Personalized Nutrition and Medicine
FDA/NCTR
3900 NCTR Road
Jefferson, AR 72079
+1 870 543 7997
+1 870 543 7662 (F)
James.kaput@fda.hhs.gov

Key Words Personalized Healthcare, Nutritional, Physical Activity, Genetic Variation Assessments, Databases

Summary

The sequencing of the human genome, subsequent analyses of human genetic variation, and studies that associate gene variants with disease markers or other phenotypic alterations provides the knowledge for personalizing medical practice. Similarly, nutritional genomics – the study of gene and nutrient interactions [50, 52] – seeks information and knowledge for designing optimal diets that allow individuals to maintain health and prevent diseases. The goal of personalized nutrition and medicine, which form the basis of personalized healthcare, is to get the right nutrient or therapeutic to the right individual at the right time for the right outcome. In addition to the interest within the research community, the political establishment has also recognized the importance and need for improving healthcare: both the U.S. Senate and U.S. House are considering bills to expand and accelerate genomic research for improving diagnosis, increasing safety of drugs, and identifying novel treatments (S976 and H.R. 6498).

While the potential benefits of personalized healthcare are significant for individuals, public health, and the economy, research and applications face significant challenges: the genetic heterogeneity of humans, the complexity of foods, and the variable physiological mechanisms that produce health or disease states [48, 49]. A subtle but more important challenge is the nature of current biomedical research studies:

association studies, whether genetic, nutritional, or nutrigenomic, are based on population studies that yield the attributable fraction (AF) - “the proportional reduction in average disease risk over a specified time interval that would be achieved by eliminating the exposure of interest from the population” while other factors remain unchanged [93]. For genetic association studies, the population attributable factor is that proportion of cases in the population that would be avoided if nobody carried the risk allele [34]. AF is often misinterpreted as a risk factor [93] rather than the fractional change in number of cases within the population. Perhaps most importantly, the attributable fraction is usually calculated from population models and is not directly applicable to individuals [116] because individuals may differ genetically, physiologically, and nutritionally from the population averages.

Novel research strategies and experimental designs are necessary for creating individual nutritional recommendations or personalized healthcare. In order to be equitable, personalized healthcare research programs must include individuals from all socioeconomic groups and geographic regions. The complex nature of genotype – environment interactions that maintain health or produce disease require a multidisciplinary effort on the scale of the human genome and haplotype map projects.

The problems

Health and disease result from a complex interaction between an individual’s genetic makeup and the influence of environmental factors. However, many experimental designs focus solely on genetic factors or environmental factors that influence onset or progression of disease. In statistical terms, genetic and nutritional epidemiology experiments are typically designed to test for the statistical main effect of genes or nutrients (respectively) on some health or disease phenotype. While some genes must be regulated in the absence of environmental factors to make a human a human, a significant proportion of an organism’s genes must respond to environmental factors, the most important of which are nutrients, or the organism will die. This statement simply reiterates Darwin’s main hypothesis. Since all organisms, including humans, adapt to their local environments and may be selected over generations, it also holds that optimum diets for one individual will not be optimal for every individual. In statistical parlance, the main effect is the genotype – environment interaction, not just the genotype and not just the environment.

Although nutritional genomics and the related field of pharmacogenomics implicitly and explicitly accept these facts, experimental designs only recently have accounted for genetic and environmental variables. More importantly, research designs rely almost exclusively on determining the average response of a group or population.

Genetic studies. Variations (e.g., single nucleotide polymorphisms, copy number) in one or more genes are analyzed and those that are associated more frequently in cases than controls may contribute to the tested phenotype. Early association studies were criticized because they could not be replicated, did not appropriately match cases to controls, and were underpowered – that is, the sample size was small (e.g., [9, 36, 39, 65, 79, 92, 106]). More recently, genome wide association studies (GWAS) attempted to address those limitations by increasing the number of cases: in some experiments, DNA and physiology in more than 20,000 individuals were analyzed (e.g., [53]). Although novel genetic loci that contribute to disease were found, the new candidate genes identified collectively “explain” less than 10% of the population attributable fraction (PAF) for the phenotype tested (e.g., [53, 97]. Several limitations contributed to the number of genes identified and their small PAF. The various GWA studies used different criteria for including individuals in the case population. In some of the experimental designs, cases included newly diagnosed (not yet on medications) as well as those on some type of medication. Combining all the cases regardless of disease subtype, age, environment, reduced (or averaged) the contributions of causative genes in different pathways. In addition, regardless of matching the overall genomic architecture of individual cases to individual controls, epistatic interactions may still occur in seemingly homogeneous populations [8, 101] because distinct chromosomal regions (and not average total genomic structure) may differ among cases, among controls, and between cases and controls. New mapping strategies that are designed specifically for analyzing epistasis in association studies may provide a means to account for confounding gene – gene interactions (e.g., [12, 73, 111]).

Nutritional epidemiology. Nutrients consumed more frequently or in larger amounts in cases versus controls may contribute to the development of a measured phenotype. Nutritional epidemiology provides the most important information for developing food intake recommendations that are codified in the U.S. food pyramid (http://www.mypyramid.gov). The fundamental thesis of these recommendations is that all individuals will metabolize and respond similarly to the same nutrients. Although the new pyramid guide incorporated age and physical activity differences to determine diet, the recommendations did not address how individual genetic differences alters nutrient requirements, mainly because data for such recommendations is not available. Experimental evidence now exists that different environments exerted selective pressure on humans in different parts of the world. Published HapMap data analyzed by a novel algorithm identified chromosomal regions with a high Fst (Fixation index, a measure of population differentiation) between three ancestral populations (European, Chinese and African) [74]. These regions encoded genes involved in carbohydrate metabolism, skeletal development, and pigmentation. Such genetic differences may explain, for example, the differential effect in incidence of obesity and type 2 diabetes between European-Americans and Pima Indians in the similar physical and food environments [98]. However, other research has demonstrated that certain gene – diet interactions are not related to strictly to ancestral background. Individuals whose ancestors were exposed to either high starch or low starch environments differed in the number of amylase genes [82, 87]. Different populations (e.g., members of different tribes) in the same geographical region (e.g., central Africa) had different numbers of amylase genes. Differences in metabolism will also occur within an ancestral population as well as between populations. Genetic differences also have been shown to alter the requirements of micronutrients. For example, 14 nonsynonymous changes including 11 alleles with frequencies <1% along with the common alleles p.A222V, p.E429A, and p.R594Q were identified by resequencing the methylene tetrahydrofolate reductase gene [68] from 564 individuals of diverse genetic ancestry (Coriell Institute panels - http://ccr.coriell.org/Sections/BrowseCatalog/Populations.aspx?PgId=4). Increased levels of folate restored MTHFR activity to the normal range in 4 of the 5 biochemical enzyme variants. The sequence heterogeneity and remediation of enzyme activity by folate supports a greater emphasis on the ~600 cofactor dependent enzymes in the human proteome. Since many cofactors are derived from diet, such studies may identify individuals who require higher concentrations of vitamins for optimal health.

The specific problems

1. Genetic variation and environmental exposures, and specifically nutrient intakes and physical activity, must be analyzed in the same experiment to generate a more complete description of the human system. These requirements are challenging because:

    a. Nutritional and physical activity assessment tools are difficult to use with accuracy and only a few are web-enabled
    b. Nutritional and physical activity databases are usually, if not always, laboratory based and inaccessible to other collaborators
    c. Genetic variation of populations has not been well studied although efforts are underway to increase representation of other ancestral populations.

2. New research strategies are needed that are based on individual responses and not the average response of a group or population.

3. The complexity of genotype – environment interactions requires a multidisciplinary collaborative effort on the scale of the human genome and haplotype map projects.

The challenges

Challenge 1: Human Genetic Diversity. The human genome project and haplotype diversity projects have identified over ~10 million single nucleotide polymorphisms in individuals from Yoruba in Ibadan, Nigeria, Japanese in Tokyo, Japan, Han Chinese in Beijing, China, and Utah residents with ancestry from northern and western Europe (http://www.hapmap.org/). The haplotype data showed that Africans have the greatest genetic variation while Asians have the least genetic diversity [35, 43]. Although 10 –

15% of genetic variation is unique to a given ancestral group, the distribution of these polymorphisms in genes involved in nutrient metabolism, regulation, or responsiveness to drugs has not been analyzed for most genes (see above and [74]). On a population level, the average efficacy and toxicity of certain drugs vary between different ancestral groups [7, 95, 119] indicating that individual differences will be important for proper treatment. Similarly nutrient – gene interactions may differ among individuals from different ethnic populations. Lactase persistence which allows individuals to consume lactose as adults is a prime example of a nutrient – gene interaction [50]. Lactase persistence in Europeans resulted from a C-13910T SNP polymorphism in promoter of the lactase phlorizin hydrolase gene (LCH) that arose as a mutation around 9,000 years ago [26]. This polymorphism is thought to alter regulatory protein - DNA interactions controlling expression of the gene [37]. Certain groups in Africa also exhibit lactase persistence even though the African and European populations have been separated for at least ~35,000 years. Sequence analyses showed that many lactose tolerant individuals in these populations have a G-14010C polymorphism 5' to the LCH gene that arose about ~7000 years ago [110]. The maintenance of the independently derived European and African variants conferred selective advantages that include improved nutrition, prevention of dehydration, and improved calcium absorption in environments that are nutrient poor [110]. Comprehensive studies analyzing macronutrients or micronutrients requirements for optimal health in different genetic groups have not been done. Geographic separation and evolution, epistatic (gene – gene interactions [12]), and epigenetic (reversible, heritable, but nongenetic) changes in gene regulation [15]) mechanisms alter nutrient utilization and susceptibility to complex diseases.

Epistatic interactions occur because proteins and enzymes are in linear or networked pathways which share metabolites or by direct physical interactions. The activity of an allele metabolizing a drug or nutrient may be influenced by a variant of an interacting protein [12, 111]. Since the probability of inheriting gene variants depends upon the allele frequencies in the population, epistatic interactions may be most obvious when comparing effects of a gene variant in different ancestral populations [57] or in admixed populations [104]. One example of this effect was shown by the effect of the HapK haplotype in the leukotriene 4 hydrolase (*LTH4A*) on cardiovascular disease (CVD) and myocardial infarction (MI). African-Americans who carry the European derived - HapK haploptype have an increased risk (~5 fold) for CVD and MI than European–Americans (1.35 fold) with the same LTHA4 haplotype. The difference in risk can be explained by at least three possible mechanisms: (i) LTH4A interacts differently with one or more gene variants in either African versus European chromosomal regions resulting in increased effect of LTH4 activity in African Americans, and/or (ii) different environmental factors alter the influence of LTH4A on myocardial infarction [31] or (iii) a combination of epistatic and gene – environment interactions [57]. LTH4A participates in leukotriene and prostaglandin metabolism which are linked to dietary fatty acid intake [54]. 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.

While epistasis may be most important for recently admixed populations (e.g., African Americans), genetic variation within a population such as European groups [8, 81, 101]} and Iceland [32] may also be affected by gene – gene interactions. The consequences of epistatic interactions are that (i) larger sample sizes are needed for gene – nutrient, gene – disease, and gene – nutrient - phenotype studies, (ii) ancestral markers may be needed to assess genetic background of each individual in the study, (iii) the results of genetic association studies may be specific only to that population [13, 57, 104, 108, 113, 114, 120] and (iv) genetic ancestry may alter the metabolism of dietary chemicals and drugs altering medical diagnosis and treatment options. A corollary of these consequences is that for purely scientific reasons, it will not be possible to completely describe nutrient – gene interactions (or drug – gene interactions) and their outcomes on physiology without comparing responses in different ancestral groups.

In addition to epistatic interactions, expression of genetic information can be altered by epigenetic mechanisms that change chromosome structure [22, 28, 41, 72]. Nutrients are involved in epigenetic mechanisms because certain dietary chemicals or their metabolites produce substrates for [15, 27] or otherwise alter [23] DNA methylation and the energy balance of a cell is connected to chromatin remodeling enzymes [88]. The study of these processes has spawned the field of nutritional epigenomics [29].

A growing body of evidence also suggests that nutrients may alter chromatin structure during developmental windows – key times during development where short term exposure to unbalanced environments or nutrients produce lasting changes to gene expression patterns (reviewed in [29, 70]). Many, if not most, human studies (for drugs, genetic association, or other) assume similar if not identical developmental paths to adulthood. However, genetic expression in individuals from impoverished environments, or in individuals who habitually consume nutritionally unbalanced diets, may differ significantly from the population average due to chromatin structure. Changing transcriptional regulation by chromatin remodeling may alter gene – phenotype or gene – nutrient – phenotype associations. Developing experimental approaches for dissecting the environmental influences and the critical genes and pathways will be essential and challenging.

Challenge 2: Complexity of Gene – Nutrient Interactions. A substantial percentage of dietary components can either directly or indirectly regulate cellular processes through metabolic pathways [50]. Well documented examples include genistein and hyperforin [99]. Genistein, present in soybean products and other plants, binds to the ligand (estradiol) binding site of estrogen receptor [89]. The binding affinity of genistein is only ~0.01% that of estradiol for estrogen receptor alpha (ERα) but ~3% relative to estradiol for ERβ [61]. Binding is physiologically relevant since this phytoestrogen activated transcription of estrogen regulated genes in uteri of rats [75]. Since ER subtypes [5, 61, 62] are distributed differently among organs, genistein is likely to produce specific effects on physiological processes in tissues where ERβ is expressed. Genistein is also known to alter signal transduction pathways through AKT, FAK, ERBB2 and BCL2 [60]. The many targets and pleiotropic effects of genisten preclude a simplistic explanation of how this isoflavonoid will alter physiology differently in different genotypes.

Dietary chemicals may also indirectly alter drug metabolism through the control of transcription or signaling pathways. Hyperforin, the active ingredient of St. John's wort, binds and activates the pregnane X receptor (PXR - [71, 90, 109, 117]). PXR is the key transcriptional factor regulating expression of CYP3A4 [42] which metabolizes ~40% of prescription drugs. Intake of St. John's wort is likely to alter drug metabolism through changes in the level of CYP3A4 mRNA and protein.

While these two compounds are examples of bioactives which directly affect expression of genetic information, numerous signaling and transcriptional activators are produced from dietary chemicals metabolized by intermediary metabolic pathways. For example, the peroxisome proliferator activated receptor gamma 2 (PPAR-γ), a target of thiazolidinediones (TZD – used for type 2 diabetes treatments), is activated by the dietary lipids linoleic, linolenic, arachidonic, and eicosapentaenoic acid [10, 80] and their metabolites. In addition to their role as transcriptional activators, dietary lipids and their metabolites alter regulation of many diverse cellular processes.

Challenge 3: Clinical Heterogeneity of Physiology. All chronic diseases and many monogenic diseases are heterogeneous in underlying metabolism and clinical appearance. Type 2 diabetes mellitus (T2DM) can be used as an example of this heterogeneticy. The incidence and severity of T2DM is significantly associated with intake of excess calories, dietary fat, high glycemic index carbohydrates, and certain micronutrients [3, 4, 38, 77, 78]. The determination of optimum nutrient intakes to prevent, delay, or reduce the severity of the disease is complicated by the chemical complexity of food [44, 50]. T2DM begins at different ages and progresses differently among individual patients [103]. Rosenzweig defined a severity score for T2DM with 4 levels (low, moderate, high, and very high), each of which described the progression of 6 conditions: glycemic control (Hb1Ac, hypoglycemia, ketosis), cardiovascular risk factors, peripheral neuropathies, eye disease, renal disease, and autonomic neuropathies [94]. The classification scheme provides guidance for treatment strategies although individual patients have different responses to drugs and eventual disease outcome. Management of T2DM can be a lengthy trial and error method, involving significant amounts of time and considerable expenses for the patient and the healthcare system.

The initial recommendation and treatment for early stage cases of T2DM (level 1, glycemic control) provides an example of the varied responses to treatments among patients. The ideal first option for early stage T2DM is to modify diet and lifestyle, which is successful in only ~20% of patients [59]. Patients

refractory to lifestyle changes, or who present with more severe indices of the disease, are treated with one or more of 6 classes of drugs). 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, stimulates insulin production [64, 76] only when glucose levels are high (http://www.diabetes.org/type-2-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%) [59].

Identifying the genes that cause abnormal responses in these molecular pathways would allow the development of diagnostic tests for sorting individuals into treatment groups. However, knowledge of gene variants involved in absorption, distribution, metabolism, and excretion of drugs will be required to fully implement phramacogenomic testing [24, 40, 55, 56]. A further complication of pharmacogenomic, and eventually nutrigenomic testing, is that certain drugs or nutrients may affect not only a specific target, but other genes, proteins, or pathways [47]. The clinical heterogeneity at first diagnosis, responsiveness to treatments, differential toxicities of drugs in different individuals, and many different complications that arise as the disease progresses presents formidable challenges for designing phramacogenomic as well as nutrigenomic experiments to identify genes and environmental factors that cause chronic diseases. Studies of gene - environmental factor – chronic disease associations require more detailed phenotypic analyses, and the number of clinical and physiological variations within a population of T2DM patients makes that a challenging task [16, 17, 84].

Challenge 4: Experimental Designs. Biomedical research has made tremendous progress over the past 50 years using population based experimental designs. However, population attributable risks can not be applied to individuals because each differs in genotype and environmental exposure from the "average" individual in a study population [45]. The challenge for nutrigenomics and biomedical research is to account for the diversity and to develop a path for creating personalized nutrition and healthcare. The key aspect of any new approach (e.g., [45]) for biomedical research is to (i) assess physiology more completely (i.e., deep phenotyping [112]), (ii) measure candidate genes and origin of chromosomal segments (e.g., [96]), (iii) assess response to changes in lifestyle either over time or acutely (e.g., homeostatic challenges [115]). The Framingham Heart Study (http://www.framinghamheartstudy.org/) is one example of a study that allows for analyses of lifestyle, phenotype, and genetic makeup but has been acknowledged as being applicable to members of the ancestral population studied (essentially European-Americans [33].

Government support and a coherent multidisciplinary program

Although much media, political, and scientific attention has been focused on healthcare costs, the failing discovery pipeline of pharmaceuticals, the growing obesity epidemic, and the unbalanced nutrient contents of manufactured foods, little attention has been given to the fundamental basis of scientific and biomedical research. The U.S. scientific enterprise has produced many and varied successes over the past ~60 years of government investments. From 1950 through 2005, the NIH received ~ $400 billion (not adjusted for inflation) to support research, the majority of which was for extramural projects in principal investigator - managed laboratories. Approximately 80% of the 2008 budget of NIH supports 50,000 competitive grants to more than 325,000 researchers at over 3,000 universities, medical schools, and other research institutions in every state and around the world (http://www.nih.gov/about/budget.htm). Funding independent projects is the lifeblood of the American research enterprise. However, the result of these investments has contributed to over 18 million citations in the National Library of Medicine's PubMed service (http://www.ncbi.nlm.nih.gov/sites/entrez) as the healthcare crisis grows and obesity and chronic diseases cost the American society almost $1 trillion per year (http://www.sehn.org/tccchronicillnesscosts.html). The complex nature of biological systems requires a more systematic experimental paradigm involving multidisciplinary teams consisting of government, academic, and industry researchers: no one research group has the expertise or resources to analyze and interpret data from omic and lifestyle assessments.

We propose a 10 year program totaling $TBD to develop the infrastructure, strategies, and data for understanding environment – genotype interactions that are the basis of health and disease processes. Failure to coordinate a comprehensive biomedical research plan risks producing more of the same: small studies that are insufficiently powered focusing on an incomplete subset of the pathways or organs – the results of which do not provide information that completely describes the biological system, and produces little useful data to address the staggering healthcare crisis.

Potential technical solutions

The significant advances in understanding complex biological process relied on reductionistic approaches: hold all variables but one constant. While this strategy was successful for certain phenotypes, understanding complex systems requires analytical approaches that incorporate rather than avoid complexity. Genes interact with nutrients and nutrients alter genetic expression – analyzing one and ignoring the other results in incomplete analyses. The key challenge for personalizing health care then is not the complexity of the datasets, but acquiring those datasets in a manner to reduce noise and increase the true signals. This might best be accomplished by pre-selecting phenotypes based on quantitative data, or alternatively, pre-selecting genotypes that maximize differences in allele frequencies of candidate genes involved in nutrient metabolism or other physiological trait. The integrative whole-system analyses of the datasets and new visualization methods, such as shown with network analysis tools provide a path to not only perform these complex experiments, but also to develop biological insight into the outcomes. The development of nutrigenomics and genetics and the application of this knowledge will provide strategies for maintaining health and improving medical treatment of chronic diseases.

Several approaches are possible to identify gene - environment interactions involved in health maintenance.

1. **Identifying "metabolic groups" based on similar genotype – environment interactions** [45]. This concept is intermediate between population based and individualized nutrient – gene interaction data. Specifically, current omic (genomic, transcriptomic, proteomic, and metabolomio) technologies can be used to identify and group individuals with common phenotypes and analyze the genetic differences between them. Alternatively, individuals can be selected based on variations in genes (and not just variants used for genetic mapping) and phenotypes can be compared. The first groups tested would be those most different in phenotypes or genetic make-ups – that is, to determine the widest range of variation within the human population. Instead of selecting subjects of one ancestral population (Europeans or Asians or Africans) who are either healthy or have disease, analyses should be done between genetic or metabolic groups that are predetermined. Members of each group would have like genotypes or metabolic phenotypes, but between group differences would be large. The key aspect of this concept is that membership in the group is based on some quantitative measure of phenotype or genotype. Once maximum differences of differing phenotypes or genotypes are determined, groups between the extremes can be determined. While it is expected that most biological traits are continuous with no discrete breaks in the phenotypic or genetic continuum, such "binning" is a standard for medical practice which uses clinical measurements to group individuals into treatment options and for statistics which rely on tertiles, quartiles, quintiles, etc, to determine structure within experimental data. The difference from these standard approaches is that the binning is done prior to physiological analyses if the genetic variation is predetermined or prior to genetic analyses if different phenotypes are identified rather than after experimental data are acquired.

   This comparative phenotype/comparative genotype concept is based on similar strategies used in laboratory animals studies: phenotypic differences among inbred strains (i.e., different genotypes) are utilized for genetic mapping [6, 58, 63, 111], but also for analyzing differences caused by gene – nutrient interactions [2, 6, 11, 14, 51, 85, 86, 102, 107]. We [48, 49] and others [67] proposed that patients be grouped based on clinical measurements (see clinical heterogeneity section above). A successful application of this idea was pre-selecting women with the phenotype of early onset (< 35 years of age) vs late onset of breast cancer that allowed the identification of 17q21 [30] and subsequent identification of BRCA1 gene [105]. An international group of

researchers from 22 countries are harmonizing a protocol to analyze gene – nutrient interactions in newly diagnosed diabetics. Two workshops (Melbourne, Australia, May 2008 and Potsdam, Germany, September 2008) were held to discuss the challenges and solutions for this project.

2. **Homeostatic challenges.** The "healthy" state may be just as metabolically diverse among individuals as disease states. The challenge homeostasis model [25, 115] notes that an individual's physiological response to an acute challenge, such as the oral glucose tolerance test (OGTT), may provide a means to sort individuals into metabolic groups and predict long term health outcomes. While there are three existing "bins" for the OGTT (normal, impaired glucose tolerance, diabetic - [118]), additional groups could be identified by assessing other metabolic systems (e.g., amino acid level differences) which are altered by the glucose bolus. Such changes occur because amino acid metabolism is linked to glucose metabolism. Additional nutritional challenges for lipid intake, micronutrient response, or other nutritional variables may be developed and the various omic technologies now permit these more detailed analyses at ever-reducing costs. Complete nutrient intake assessments would, of course, be required to completely analyze baseline phenotypic data as well as data derived from homeostatic challenges.

   A subset of homeostatic challenges may consist of nutritional interventions – that is, how do individuals respond to increased (or decreased) levels of a nutrient. A group of nutrigenomics researchers is organizing an international effort that will analyze responses of individuals of differing ancestral backgrounds to increased levels of micronutrients (vitamins and minerals). The first workshop on this approach was held in Potsdam, Germany in September 2008, and a second workshop to organize a steering committee will be held in Vancouver, Canada, on February 16,17, 2009 (http://www.nutrigenomics.ca/IMGP/Research/).

Regardless of the experimental approach – or approaches – that are used, the infrastructure for genetic and environmental assessments must be improved. Specifically, funds are needed to

1. Re-sequence genes involved in or regulated by environmental factors in individuals from diverse ancestral backgrounds. The Human Genome, Haplotype Map, and newly organized 1000 genomes project (http://www.1000genomes.org and [66, 100]) will analyze less than 2000 individuals. More importantly, these sequences will not be linked to phenotype. The Human Variome Project (HVP) is organizing an international effort to re-sequence genes involved in monogenic and polygenic diseases and curate locus specific databases for linking phenotype to gene sequence (http://www.humanvariome.org and [1, 18-21, 46, 83, 91]). The HVP is developing the strategies, software tools, and databases which can be incorporated into the U.S. efforts for understanding gene – environment interactions.
2. Development of nutritional and physical activity assessment tools for research and public use. Scientists in the the omics' communities (genomic, proteomic, metabolomic, transcriptomic, phenotypic) have developed software tools and databases that can be accessed freely and used by researchers to improve experimental designs and for interpretation of data. The nutritional community has not developed a similar suite of tools for research or for easy use by the public. Hence, analyzing environmental exposures, including nutrient intakes and physical activity, are challenging and are often stand-alone efforts that can not be linked among researchers. Building an infrastructure for environmental exposures linked to omic analyses is a high priority for this initiative. The USDA – ARS, FDA (NCTR and CFSAN), and NIH are coordinating a workshop to be held in Beltsville, MD, February 24,25, 2009, to discuss the needs and strategies to develop environmental assessment tools and databases.
3. The NIH has embarked on an expanded research program to translate biomedical research to the clinic as part of the NIH Roadmap for Medical Research (http://nihroadmap.nih.gov/clinicalresearch/overview-translational.asp). However, these centers are located at major medical universities, typically in large cities. A great need exists to develop research programs in small cities and rural areas. Primary care research and

community based participatory research (e.g., [69]) may provide a path to involve individuals of different socioeconomic classes in small cities and rural areas of the U.S. Gene – environment interactions may differ among these populations and their inclusion improves the science and more importantly, the equitable distribution of resources and benefits that may develop from this research initiative.

Stakeholders

Academic, government, and industry scientists will be interested in this initiative. However, a key component will be the creation of a consortium of scientists from a wide diversity of social and biomedical research fields. The partners listed on this proposal are only the nucleus of that team, and contacts within industry will be approached to join this effort. A concerted effort will be made to develop collaborations through community based participatory research with community centers in rural and underserved areas of the U.S. One example of this type of program is the Boys, Girls, and Adults Community Development Center of Marvell, AR (http://www.bgacdc.org - [69])

International efforts will be integrated into this initiative since a complete understanding of gene – nutrient interactions will require comparison of responses in different genetic ancestries and cultures (e.g., [45]). The European Nutrigenomics Organization (NuG – http://www.nugo.org), the Human Variome Project, and the international efforts for diabetes research and micronutrient genetics (International Micronutrient Genetics Project). Linking our efforts with international programs will improve the science and ensure the equitable distribution of benefits to all populations.

References

1. What is the human variome project? *Nat Genet* 2007; 39(4): p. 423.
2. Akhi M, Kose H, and Matsumoto K. Fine mapping of the hyperglycemic and obesity QTL by congenic strains suggests multiple loci on rat chromosome 14. *J Med Invest* 2005; 52(1-2): p. 109-13.
3. Biesalski HK. Diabetes preventive components in the Mediterranean diet. *Eur J Nutr* 2004; 43 Suppl 1: p. I/26-30.
4. Bonnefont-Rousselot D. The role of antioxidant micronutrients in the prevention of diabetic complications. *Treat Endocrinol* 2004; 3(1): p. 41-52.
5. Brandenberger AW, Tee MK, Lee JY, Chao V, and Jaffe RB. Tissue distribution of estrogen receptors alpha (ER-alpha) and beta (ER- beta) mRNA in the midgestational human fetus. *J Clin Endocrinol Metab* 1997; 82(10): p. 3509-12.
6. Brown AC, Olver WI, Donnelly CJ, May ME, Naggert JK, Shaffer DJ, and Roopenian DC. Searching QTL by gene expression: analysis of diabesity. *BMC Genet* 2005; 6(1): p. 12.
7. Burchard EG, Avila PC, Nazario S, Casal J, Torres A, Rodriguez-Santana JR, Toscano M, Sylvia JS, Alioto M, Salazar M, Gomez I, Fagan JK, Salas J, Lilly C, Matallana H, Ziv E, Castro R, Selman M, Chapela R, Sheppard D, Weiss ST, Ford JG, Boushey HA, Rodriguez-Cintron W, Drazen JM, and Silverman EK. Lower bronchodilator responsiveness in Puerto Rican than in Mexican subjects with asthma. *Am J Respir Crit Care Med* 2004; 169(3): p. 386-92.
8. Campbell CD, Ogburn EL, Lunetta KL, Lyon HN, Freedman ML, Groop LC, Altshuler D, Ardlie KG, and Hirschhorn JN. Demonstrating stratification in a European American population. *Nat Genet* 2005; 37(8): p. 868-72.
9. Cardon LR and Bell JI. Association study designs for complex diseases. *Nat Rev Genet* 2001; 2(2): p. 91-9.
10. Chambrier C, Bastard JP, Rieusset J, Chevillotte E, Bonnefont-Rousselot D, Therond P, Hainque B, Riou JP, Laville M, and Vidal H. Eicosapentaenoic acid induces mRNA expression of peroxisome proliferator-activated receptor gamma. *Obes Res* 2002; 10(6): p. 518-25.
11. Cheverud JM, Ehrich TH, Hrbek T, Kenney JP, Pletscher LS, and Semenkovich CF. Quantitative trait loci for obesity- and diabetes-related traits and their dietary responses to high-fat feeding in LGXSM recombinant inbred mouse strains. *Diabetes* 2004; 53(12): p. 3328-36.
12. Chiu S, Diament, A.L., Fisler, J.S., and Warden, C.H., *Gene-Gene Epistasis and Gene Environment Interactions Influence Diabetes and Obesity*, in *Nutritional Genomics. Discovering*

the Path to Personalized Nutrition, Kaput J, Rodriguez, R.L., Editor. 2006, John Wiley and Sons: Hoboken, NJ. p. 135 - 152.
13. Choudhry S, Coyle NE, Tang H, Salari K, Lind D, Clark SL, Tsai HJ, Naqvi M, Phong A, Ung N, Matallana H, Avila PC, Casal J, Torres A, Nazario S, Castro R, Battle NC, Perez-Stable EJ, Kwok PY, Sheppard D, Shriver MD, Rodriguez-Cintron W, Risch N, Ziv E, and Burchard EG. Population stratification confounds genetic association studies among Latinos. *Hum Genet* 2006; 118(5): p. 652-64.
14. Clee SM, Yandell BS, Schueler KM, Rabaglia ME, Richards OC, Raines SM, Kabara EA, Klass DM, Mui ET, Stapleton DS, Gray-Keller MP, Young MB, Stoehr JP, Lan H, Boronenkov I, Raess PW, Flowers MT, and Attie AD. Positional cloning of Sorcs1, a type 2 diabetes quantitative trait locus. *Nat Genet* 2006; 38(6): p. 688-93.
15. Cooney CA, *Maternal Nutrition: Nutrients and Control of Expression*, in *Nutritional Genomics. Discovering the Path to Personalized Nutrition*, Kaput J, Rodriguez, RL, Editor. 2006, John Wiley and Sons: Hoboken, NJ. p. 219 - 254.
16. Corella D and Ordovas JM. Integration of environment and disease into 'omics' analysis. *Curr Opin Mol Ther* 2005; 7(6): p. 569-76.
17. Corella D and Ordovas JM. Single Nucleotide Polymorphisms that Influence Lipid Metabolism: Interaction with Dietary Factors. *Annu Rev Nutr* 2005; 25: p. 341-390.
18. Cotton RG, Appelbe W, Auerbach AD, Becker K, Bodmer W, Boone DJ, Boulyjenkov V, Brahmachari S, Brody L, Brookes A, Brown AF, Byers P, Cantu JM, Cassiman JJ, Claustres M, Concannon P, den Dunnen JT, Flicek P, Gibbs R, Hall J, Hasler J, Katz M, Kwok PY, Laradi S, Lindblom A, Maglott D, Marsh S, Masimirembwa CM, Minoshima S, de Ramirez AM, Pagon R, Ramesar R, Ravine D, Richards S, Rimoin D, Ring HZ, Scriver CR, Sherry S, Shimizu N, Stein L, Tadmouri GO, Taylor G, and Watson M. Recommendations of the 2006 Human Variome Project meeting. *Nat Genet* 2007; 39(4): p. 433-6.
19. Cotton RG, Auerbach AD, Axton M, Barash CI, Berkovic SF, Brookes AJ, Burn J, Cutting G, den Dunnen JT, Flicek P, Freimer N, Greenblatt MS, Howard HJ, Katz M, Macrae FA, Maglott D, Moslein G, Povey S, Ramesar RS, Richards CS, Seminara D, Smith TD, Sobrido MJ, Solbakk JH, Tanzi RE, Tavtigian SV, Taylor GR, Utsunomiya J, and Watson M. GENETICS. The Human Variome Project. *Science* 2008; 322(5903): p. 861-2.
20. Cotton RG, Auerbach AD, Brown AF, Carrera P, Christodoulou J, Claustres M, Compton J, Cox DW, De Baere E, den Dunnen JT, Greenblatt M, Fujiwara M, Hilbert P, Jani A, Lehvaslaiho H, Nebert DW, Verma I, and Vihinen M. A structured simple form for ordering genetic tests is needed to ensure coupling of clinical detail (phenotype) with DNA variants (genotype) to ensure utility in publication and databases. *Hum Mutat* 2007; 28(10): p. 931-2.
21. Cotton RGH, Auerbach AD, Axton M, Barash CI, Berkovic SF, Brookes AJ, Burn J, Cutting G, den Dunnen JT, Flicek P, Freimer N, Greenblatt MS, Howard HJ, Katz M, Macrae FA, Maglott D, Moslein G, Povey S, Ramesar RS, Richards CS, Seminara D, Smith TD, Sobrido M-J, Solbakk JH, Tanzi RE, Tavtigian SV, Taylor GR, Utsunomiya J, and Watson M. GENETICS: The Human Variome Project. *Science* 2008; 322(5903): p. 861-862.
22. Delaval K and Feil R. Epigenetic regulation of mammalian genomic imprinting. *Curr Opin Genet Dev* 2004; 14(2): p. 188-95.
23. Dolinoy DC, Weidman JR, Waterland RA, and Jirtle RL. Maternal genistein alters coat color and protects Avy mouse offspring from obesity by modifying the fetal epigenome. *Environ Health Perspect* 2006; 114(4): p. 567-72.
24. Ekins S, Nikolsky Y, and Nikolskaya T. Techniques: Application of systems biology to absorption, distribution, metabolism, excretion and toxicity. *Trends in Pharmacological Sciences* 2005; 26(4): p. 202-209.
25. Elliott R, Pico C, Dommels Y, Wybranska I, Hesketh J, and Keijer J. Nutrigenomic approaches for benefit-risk analysis of foods and food components: defining markers of health. *Br J Nutr* 2007; 98(6): p. 1-6.
26. Enattah NS, Sahi T, Savilahti E, Terwilliger JD, Peltonen L, and Jarvela I. Identification of a variant associated with adult-type hypolactasia. *Nat Genet* 2002; 30(2): p. 233-7.
27. Fenech M. The Genome Health Clinic and Genome Health Nutrigenomics concepts: diagnosis and nutritional treatment of genome and epigenome damage on an individual basis. *Mutagenesis* 2005; 20(4): p. 255-69.

28. Fowler AM and Alarid ET. Dynamic control of nuclear receptor transcription. *Sci STKE* 2004; 2004(256): p. pe51.
29. Gallou-Kabani C and Junien C. Nutritional epigenomics of metabolic syndrome: new perspective against the epidemic. *Diabetes* 2005; 54(7): p. 1899-906.
30. Hall JM, Lee MK, Newman B, Morrow JE, Anderson LA, Huey B, and King MC. Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* 1990; 250(4988): p. 1684-9.
31. Helgadottir A, Manolescu A, Helgason A, Thorleifsson G, Thorsteinsdottir U, Gudbjartsson DF, Gretarsdottir S, Magnusson KP, Gudmundsson G, Hicks A, Jonsson T, Grant SF, Sainz J, O'Brien S J, Sveinbjornsdottir 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, and Stefansson K. A variant of the gene encoding leukotriene A4 hydrolase confers ethnicity-specific risk of myocardial infarction. *Nat Genet* 2006; 38(1): p. 68-74.
32. Helgason A, Yngvadottir B, Hrafnkelsson B, Gulcher J, and Stefansson K. An Icelandic example of the impact of population structure on association studies. *Nat Genet* 2005; 37(1): p. 90-5.
33. Hemann BA, Bimson WF, and Taylor AJ. The Framingham Risk Score: an appraisal of its benefits and limitations. *Am Heart Hosp J* 2007; 5(2): p. 91-6.
34. Hemminki K and Bermejo JL. Constraints for genetic association studies imposed by attributable fraction and familial risk. *Carcinogenesis* 2007; 28(3): p. 648-56.
35. Hinds DA, Stuve LL, Nilsen GB, Halperin E, Eskin E, Ballinger DG, Frazer KA, and Cox DR. Whole-genome patterns of common DNA variation in three human populations. *Science* 2005; 307(5712): p. 1072-9.
36. Hirschhorn JN, Lohmueller K, Byrne E, and Hirschhorn K. A comprehensive review of genetic association studies. *Genet Med* 2002; 4(2): p. 45-61.
37. Hollox EJ, Poulter M, Wang Y, Krause A, and Swallow DM. Common polymorphism in a highly variable region upstream of the human lactase gene affects DNA-protein interactions. *Eur J Hum Genet* 1999; 7(7): p. 791-800.
38. Hung T, Sievenpiper JL, Marchie A, Kendall CW, and Jenkins DJ. Fat versus carbohydrate in insulin resistance, obesity, diabetes and cardiovascular disease. *Curr Opin Clin Nutr Metab Care* 2003; 6(2): p. 165-76.
39. Ioannidis JP. Why most published research findings are false. *PLoS Med* 2005; 2(8): p. e124.
40. Jain KK. Applications of AmpliChip CYP450. *Mol Diagn* 2005; 9(3): p. 119-27.
41. Jiang YH, Bressler J, and Beaudet AL. Epigenetics and human disease. *Annu Rev Genomics Hum Genet* 2004; 5: p. 479-510.
42. Jones SA, Moore LB, Shenk JL, Wisely GB, Hamilton GA, McKee DD, Tomkinson NCO, LeCluyse EL, Lambert MH, Willson TM, Kliewer SA, and Moore JT, *The Pregnane X Receptor: A Promiscuous Xenobiotic Receptor That Has Diverged during Evolution*. 2000. p. 27-39.
43. Jorde LB and Wooding SP. Genetic variation, classification and 'race'. *Nat Genet* 2004; 36 Suppl 1: p. S28-33.
44. Kaput J. Diet-disease gene interactions. *Nutrition* 2004; 20(1): p. 26-31.
45. Kaput J. Nutrigenomics research for personalized nutrition and medicine. *Curr Opin Biotechnol* 2008; 19(2): p. 110-20.
46. Kaput J, Cotton R, Hardman L, Aqeel A, Al-Aama J, Al-Mulla F, Aretz S, Auerbach A, Axton M, Bapat B, Bernstein I, Bhak J, Bleoo S, Blöcker H, Brenner S, Burn J, Bustamante M, Calzone R, Cambon-Thomsen A, Cargill M, Carrera P, Cavedon L, Cho Y, Chung Y-J, Claustres M, Cutting G, Dalgleish R, Dunnen J, Díaz C, Dobrowolski S, dos Santos R, Ekong R, Flanagan S, Flicek P, Furukawa Y, Genuardi M, Ghang H, Golubenko M, Greenblatt M, Hamosh A, Hancock J, Hardison R, Harrison T, Hoffmann R, Horaitis R, Howard H, Isaacson-Barash C, Izagirre N, Jung J, Kojima T, Laradi S, Lee Y-S, Lee J-Y, Gil-da-Silva-Lopes V, Macrae F, Maglott D, Marafie M, Marsh S, Matsubara Y, Messiaen L, Möslein G, Netea M, Norton M, Oefner P, Oetting W, O'Leary J, Oller de Ramirez A, Paalman M, Parboosingh J, Patrinos G, Perozzi G, Phillips I, Povey S, Prasad S, Qi M, Quin D, Ramesar R, Richards C, Savige J, Scheible D, Scott R, Seminara D, Shephard E, Sijmons R, Smith T, Sobrido M-J, Tanaka T, Tavtigian S, Taylor G, Teague J, Töpel T, Ullman-Cullere M, Utsunomiya J, van Kranen H, Vihinen M, Watson M, Webb E, Weber T, Yeager M, Yeom Y, Yim S-H, Yoo H-S and Meeting oboctt HVPP. Planning the Human Variome Project: The Spain Report. *Human Mutatation* 2009; in press.

47. Kaput J and Dawson K. Complexity of type 2 diabetes mellitus data sets emerging from nutrigenomic research: a case for dimensionality reduction? *Mutat Res* 2007; 622(1-2): p. 19-32.
48. Kaput J, Noble J, Hatipoglu B, Kohrs K, Dawson K, and Bartholomew A. Application of nutrigenomic concepts to Type 2 diabetes mellitus*. Nutr Metab Cardiovasc Dis* 2007; 17(2): p. 89-103.
49. Kaput J, Perlina A, Hatipoglu B, Bartholomew A, and Nikolsky Y. Nutrigenomics: concepts and applications to pharmacogenomics and clinical medicine*. Pharmacogenomics* 2007; 8(4): p. 369-390.
50. Kaput J and Rodriguez RL. Nutritional genomics: the next frontier in the postgenomic era*. Physiol Genomics* 2004; 16(2): p. 166-77.
51. Kaput J, Swartz D, Paisley E, Mangian H, Daniel WL, and Visek WJ. Diet-disease interactions at the molecular level: an experimental paradigm*. J Nutr* 1994; 124(8 Suppl): p. 1296S-1305S.
52. Kaput JaR, RL, *Nutritional Genomics. Discovering the Path to Personalized Nutrition*. 2006, Hoboken, NJ: John Wiley and Sons. 469.
53. Kathiresan S, Willer CJ, Peloso GM, Demissie S, Musunuru K, Schadt EE, Kaplan L, Bennett D, Li Y, Tanaka T, Voight BF, Bonnycastle LL, Jackson AU, Crawford G, Surti A, Guiducci C, Burtt NP, Parish S, Clarke R, Zelenika D, Kubalanza KA, Morken MA, Scott LJ, Stringham HM, Galan P, Swift AJ, Kuusisto J, Bergman RN, Sundvall J, Laakso M, Ferrucci L, Scheet P, Sanna S, Uda M, Yang Q, Lunetta KL, Dupuis J, de Bakker PI, O'Donnell CJ, Chambers JC, Kooner JS, Hercberg S, Meneton P, Lakatta EG, Scuteri A, Schlessinger D, Tuomilehto J, Collins FS, Groop L, Altshuler D, Collins R, Lathrop GM, Melander O, Salomaa V, Peltonen L, Orho-Melander M, Ordovas JM, Boehnke M, Abecasis GR, Mohlke KL, and Cupples LA. Common variants at 30 loci contribute to polygenic dyslipidemia*. Nat Genet* 2009; 41(1): p. 56-65.
54. Kelley DS. Modulation of human immune and inflammatory responses by dietary fatty acids*. Nutrition* 2001; 17(7-8): p. 669-73.
55. Kirchheiner J, Roots I, Goldammer M, Rosenkranz B, and Brockmoller J. Effect of genetic polymorphisms in cytochrome p450 (CYP) 2C9 and CYP2C8 on the pharmacokinetics of oral antidiabetic drugs: clinical relevance*. Clin Pharmacokinet* 2005; 44(12): p. 1209-25.
56. Kirchheiner J, Thomas S, Bauer S, Tomalik-Scharte D, Hering U, Doroshyenko O, Jetter A, Stehle S, Tsahuridu M, Meineke I, Brockmoller J, and Fuhr U. Pharmacokinetics and pharmacodynamics of rosiglitazone in relation to CYP2C8 genotype*. Clin Pharmacol Ther* 2006; 80(6): p. 657-67.
57. Klos KLE, Kardia SLR, Hixson JE, Turner ST, Hanis C, Boerwinkle E, and Sing CF. Linkage Analysis of Plasma ApoE in Three Ethnic Groups: Multiple Genes with Context-Dependent Effects*. Annals of Human Genetics* 2005; 69(2): p. 157-167.
58. Kobayashi M, Io F, Kawai T, Kumazawa M, Ikegami H, Nishimura M, Ohno T, and Horio F. Major quantitative trait locus on chromosome 2 for glucose tolerance in diabetic SMXA-5 mouse established from non-diabetic SM/J and A/J strains*. Diabetologia* 2006; 49(3): p. 486-95.
59. Koro CE, Bowlin SJ, Bourgeois N, and Fedder DO. Glycemic control from 1988 to 2000 among U.S. adults diagnosed with type 2 diabetes: a preliminary report*. Diabetes Care* 2004; 27(1): p. 17-20.
60. Kousidou O, Tzanakakis GN, and Karamanos NK. Effects of the natural isoflavonoid genistein on growth, signaling pathways and gene expression of matrix macromolecules by breast cancer cells*. Mini Rev Med Chem* 2006; 6(3): p. 331-7.
61. Kuiper GG, Carlsson B, Grandien K, Enmark E, Haggblad J, Nilsson S, and Gustafsson JA. Comparison of the ligand binding specificity and transcript tissue distribution of estrogen receptors alpha and beta*. Endocrinology* 1997; 138(3): p. 863-70.
62. Kuiper GG, Shughrue PJ, Merchenthaler I, and Gustafsson JA. The estrogen receptor beta subtype: a novel mediator of estrogen action in neuroendocrine systems*. Front Neuroendocrinol* 1998; 19(4): p. 253-86.
63. Kumar KG, Poole AC, York B, Volaufova J, Zuberi A, and Richards BK. Quantitative trait loci for carbohydrate and total energy intake on mouse chromosome 17: congenic strain confirmation and candidate gene analyses (Glo1, Glp1r)*. Am J Physiol Regul Integr Comp Physiol* 2007; 292(1): p. R207-16.

64. Kwon G, Marshall CA, Pappan KL, Remedi MS, and McDaniel ML. Signaling elements involved in the metabolic regulation of mTOR by nutrients, incretins, and growth factors in islets. *Diabetes* 2004; 53 Suppl 3: p. S225-32.
65. Lander E and Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results [see comments]. *Nat Genet* 1995; 11(3): p. 241-7.
66. Lang L. Three sequencing companies join the 1000 genomes project. *Gastroenterology* 2008; 135(2): p. 336-7.
67. Lohmueller KE, Wong LJ, Mauney MM, Jiang L, Felder RA, Jose PA, and Williams SM. Patterns of genetic variation in the hypertension candidate gene GRK4: ethnic variation and haplotype structure. *Ann Hum Genet* 2006; 70(Pt 1): p. 27-41.
68. Marini NJ, Gin J, Ziegle J, Keho KH, Ginzinger D, Gilbert DA, and Rine J. The prevalence of folate-remedial MTHFR enzyme variants in humans. *Proc Natl Acad Sci U S A* 2008; 105(23): p. 8055-60.
69. McCabe-Sellers B, Lovera D, Nuss H, Wise C, Green B, Teitel C, Ning B, Clark B, Bogle M, and Kaput J. Community Based Participatory Research and Omics Technologies. *Omics, A Journal of Integrative Biology* 2008: p. in press.
70. McMillen IC and Robinson JS. Developmental origins of the metabolic syndrome: prediction, plasticity, and programming. *Physiol Rev* 2005; 85(2): p. 571-633.
71. Moore LB, Goodwin B, Jones SA, Wisely GB, Serabjit-Singh CJ, Willson TM, Collins JL, and Kliewer SA. St. John's wort induces hepatic drug metabolism through activation of the pregnane X receptor. *Proc Natl Acad Sci U S A* 2000; 97(13): p. 7500-2.
72. Morgan HD, Santos F, Green K, Dean W, and Reik W. Epigenetic reprogramming in mammals. *Hum Mol Genet* 2005; 14 Spec No 1: p. R47-58.
73. Motsinger AA and 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(5): p. 318-28.
74. Myles S, Tang K, Somel M, Green RE, Kelso J, and Stoneking M. Identification and analysis of genomic regions with large between-population differentiation in humans. *Ann Hum Genet* 2008; 72(Pt 1): p. 99-110.
75. Naciff JM, Jump ML, Torontali SM, Carr GJ, Tiesman JP, Overmann GJ, and Daston GP. Gene expression profile induced by 17alpha-ethynyl estradiol, bisphenol A, and genistein in the developing female reproductive system of the rat. *Toxicol Sci* 2002; 68(1): p. 184-99.
76. Nauck MA and Meier JJ. Glucagon-like peptide 1 and its derivatives in the treatment of diabetes. *Regul Pept* 2005; 128(2): p. 135-48.
77. Neff LM. Evidence-based dietary recommendations for patients with type 2 diabetes mellitus. *Nutr Clin Care* 2003; 6(2): p. 51-61.
78. Nestel P. Nutritional aspects in the causation and management of the metabolic syndrome. *Endocrinol Metab Clin North Am* 2004; 33(3): p. 483-92, v.
79. Newton-Cheh C and Hirschhorn JN. Genetic association studies of complex traits: design and analysis issues. *Mutat Res* 2005; 573(1-2): p. 54-69.
80. Nosjean O and Boutin JA. Natural ligands of PPARgamma: are prostaglandin J(2) derivatives really playing the part? *Cell Signal* 2002; 14(7): p. 573-83.
81. Novembre J, Johnson T, Bryc K, Kutalik Z, Boyko AR, Auton A, Indap A, King KS, Bergmann S, Nelson MR, Stephens M, and Bustamante CD. Genes mirror geography within Europe. *Nature* 2008.
82. Novembre J, Pritchard JK, and Coop G. Adaptive drool in the gene pool. *Nat Genet* 2007; 39(10): p. 1188-90.
83. Oetting WS. The 2006 Human Genome Variation Society scientific meeting. *Hum Mutat* 2007; 28(5): p. 517-21.
84. Ordovas JM and Corella D, *Gene-Environment Interactions: Defining the Playfield*, in *Nutritional Genomics. Discovering the Path to Personalized Nutrition*, Kaput J and Rodriguez RL, Editors. 2006, John Wiley and Sons: Hoboken, NJ. p. 57 - 76.
85. Paigen B, Ishida BY, Verstuyft J, Winters RB, and Albee D. Atherosclerosis susceptibility differences among progenitors of recombinant inbred strains of mice. *Arteriosclerosis* 1990; 10(2): p. 316-23.

86. Park EI, Paisley EA, Mangian HJ, Swartz DA, Wu MX, O'Morchoe PJ, Behr SR, Visek WJ, and Kaput J. Lipid level and type alter stearoyl CoA desaturase mRNA abundance differently in mice with distinct susceptibilities to diet-influenced diseases. *J Nutr* 1997; 127(4): p. 566-73.
87. Perry GH, Dominy NJ, Claw KG, Lee AS, Fiegler H, Redon R, Werner J, Villanea FA, Mountain JL, Misra R, Carter NP, Lee C, and Stone AC. Diet and the evolution of human amylase gene copy number variation. *Nat Genet* 2007; 39(10): p. 1256-60.
88. Picard F, Kurtev M, Chung N, Topark-Ngarm A, Senawong T, Machado De Oliveira R, Leid M, McBurney MW, and Guarente L. Sirt1 promotes fat mobilization in white adipocytes by repressing PPAR-gamma. *Nature* 2004; 429(6993): p. 771-6.
89. Pike AC, Brzozowski AM, Hubbard RE, Bonn T, Thorsell AG, Engstrom O, Ljunggren J, Gustafsson JA, and Carlquist M. Structure of the ligand-binding domain of oestrogen receptor beta in the presence of a partial agonist and a full antagonist. *Embo J* 1999; 18(17): p. 4608-18.
90. Rebbeck TR, Spitz M, and Wu X. Assessing the function of genetic variants in candidate gene association studies. *Nat Rev Genet* 2004; 5(8): p. 589-97.
91. Ring HZ, Kwok PY, and Cotton RG. Human Variome Project: an international collaboration to catalogue human genetic variation. *Pharmacogenomics* 2006; 7(7): p. 969-72.
92. Risch N. Evolving methods in genetic epidemiology. II. Genetic linkage from an epidemiologic perspective. *Epidemiol Rev* 1997; 19(1): p. 24-32.
93. Rockhill B, Newman B, and Weinberg C. Use and misuse of population attributable fractions. *Am J Public Health* 1998; 88(1): p. 15-9.
94. Rosenzweig JL, Weinger K, Poirier-Solomon L, and Rushton M. Use of a disease severity index for evaluation of healthcare costs and management of comorbidities of patients with diabetes mellitus. *Am J Manag Care* 2002; 8(11): p. 950-8.
95. Rosskopf D, Manthey I, and Siffert W. Identification and ethnic distribution of major haplotypes in the gene GNB3 encoding the G-protein beta3 subunit. *Pharmacogenetics* 2002; 12(3): p. 209-20.
96. Sankararaman S, Sridhar S, Kimmel G, and Halperin E. Estimating local ancestry in admixed populations. *Am J Hum Genet* 2008; 82(2): p. 290-303.
97. Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H, Roix JJ, Kathiresan S, Hirschhorn JN, Daly MJ, Hughes TE, Groop L, Altshuler D, Almgren P, Florez JC, Meyer J, Ardlie K, Bengtsson Bostrom K, Isomaa B, Lettre G, Lindblad U, Lyon HN, Melander O, Newton-Cheh C, Nilsson P, Orho-Melander M, Rastam L, Speliotes EK, Taskinen MR, Tuomi T, Guiducci C, Berglund A, Carlson J, Gianniny L, Hackett R, Hall L, Holmkvist J, Laurila E, Sjogren M, Sterner M, Surti A, Svensson M, Tewhey R, Blumenstiel B, Parkin M, Defelice M, Barry R, Brodeur W, Camarata J, Chia N, Fava M, Gibbons J, Handsaker B, Healy C, Nguyen K, Gates C, Sougnez C, Gage D, Nizzari M, Gabriel SB, Chirn GW, Ma Q, Parikh H, Richardson D, Ricke D, and Purcell S. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 2007; 316(5829): p. 1331-6.
98. Schulz LO, Bennett PH, Ravussin E, Kidd JR, Kidd KK, Esparza J, and Valencia ME. Effects of traditional and western environments on prevalence of type 2 diabetes in Pima Indians in Mexico and the U.S. *Diabetes Care* 2006; 29(8): p. 1866-71.
99. Schuster GU, *Nutrients and Gene Expression*, in *Nutritional Genomics. Discovering the Path to Personalized Nutrition*, Kaput J, Rodriguez, R.L., Editor. 2006, John Wiley and Sons: Hoboken, NJ. p. 153 - 176.
100. Siva N. 1000 Genomes project. *Nat Biotechnol* 2008; 26(3): p. 256.
101. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, Boutin P, Vincent D, Belisle A, Hadjadj S, Balkau B, Heude B, Charpentier G, Hudson TJ, Montpetit A, Pshezhetsky AV, Prentki M, Posner BI, Balding DJ, Meyre D, Polychronakos C, and Froguel P. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 2007; 445(7130): p. 881-5.
102. Smith Richards BK, Belton BN, Poole AC, Mancuso JJ, Churchill GA, Li R, Volaufova J, Zuberi A, and York B. QTL analysis of self-selected macronutrient diet intake: fat, carbohydrate, and total kilocalories. *Physiol Genomics* 2002; 11(3): p. 205-17.
103. Stumvoll M, Goldstein BJ, and van Haeften TW. Type 2 diabetes: principles of pathogenesis and therapy. *Lancet* 2005; 365(9467): p. 1333-46.
104. Suarez-Kurtz G and Pena SD. Pharmacogenomics in the Americas: the impact of genetic admixture. *Curr Drug Targets* 2006; 7(12): p. 1649-58.

105. Szabo CI and King MC. Inherited breast and ovarian cancer. *Hum Mol Genet* 1995; 4 Spec No: p. 1811-7.
106. Tabor HK, Risch NJ, and Myers RM. OPINION: Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nat Rev Genet* 2002; 3(5): p. 391-7.
107. Takeshita S, Moritani M, Kunika K, Inoue H, and Itakura M. Diabetic modifier QTLs identified in F2 intercrosses between Akita and A/J mice. *Mamm Genome* 2006; 17(9): p. 927-40.
108. Tang H, Coram M, Wang P, Zhu X, and Risch N. Reconstructing genetic ancestry blocks in admixed individuals. *Am J Hum Genet* 2006; 79(1): p. 1-12.
109. Tirona RG, Leake BF, Podust LM, and Kim RB. Identification of amino acids in rat pregnane X receptor that determine species-specific activation. *Mol Pharmacol* 2004; 65(1): p. 36-44.
110. Tishkoff SA, Reed FA, Ranciaro A, Voight BF, Babbitt CC, Silverman JS, Powell K, Mortensen HM, Hirbo JB, Osman M, Ibrahim M, Omar SA, Lema G, Nyambo TB, Ghori J, Bumpstead S, Pritchard JK, Wray GA, and Deloukas P. Convergent adaptation of human lactase persistence in Africa and Europe. *Nat Genet* 2007; 39(1): p. 31-40.
111. Togawa K, Moritani M, Yaguchi H, and Itakura M. Multidimensional genome scans identify the combinations of genetic loci linked to diabetes-related phenotypes in mice. *Hum Mol Genet* 2006; 15(1): p. 113-28.
112. Tracy RP. 'Deep phenotyping': characterizing populations in the era of genomics and systems biology. *Curr Opin Lipidol* 2008; 19(2): p. 151-7.
113. Tsai HJ, Choudhry S, Naqvi M, Rodriguez-Cintron W, Burchard EG, and Ziv E. Comparison of three methods to estimate genetic ancestry and control for stratification in genetic association studies among admixed populations. *Hum Genet* 2005; 118(3-4): p. 424-33.
114. Tsai HJ, Kho JY, Shaikh N, Choudhry S, Naqvi M, Navarro D, Matallana H, Castro R, Lilly CM, Watson HG, Meade K, Lenoir M, Thyne S, Ziv E, and Burchard EG. Admixture-matched case-control study: a practical approach for genetic association studies in admixed populations. *Hum Genet* 2006; 118(5): p. 626-39.
115. van Ommen B, Keijer, J, Heil, SG, Kaput, J. Challenging homeostasis to define biomarkers for individual nutrition related health. *Molecular Nutrition and Food Research, in press* 2009.
116. Vineis P and Kriebel D. Causal models in epidemiology: past inheritance and genetic future. *Environ Health* 2006; 5: p. 21.
117. Watkins RE, Maglich JM, Moore LB, Wisely GB, Noble SM, Davis-Searles PR, Lambert MH, Kliewer SA, and Redinbo MR. 2.1 A crystal structure of human PXR in complex with the St. John's wort compound hyperforin. *Biochemistry* 2003; 42(6): p. 1430-8.
118. Waugh N, Scotland G, McNamee P, Gillett M, Brennan A, Goyder E, Williams R, and John A. Screening for type 2 diabetes: literature review and economic modelling. *Health Technol Assess* 2007; 11(17): p. 1-144.
119. Xie HG, Kim RB, Wood AJ, and Stein CM. Molecular basis of ethnic differences in drug disposition and response. *Annu Rev Pharmacol Toxicol* 2001; 41: p. 815-50.
120. Yang N, Li H, Criswell LA, Gregersen PK, Alarcon-Riquelme ME, Kittles R, Shigeta R, Silva G, Patel PI, Belmont JW, and Seldin MF. Examination of ancestry and ethnic affiliation using highly informative diallelic DNA markers: application to diverse and admixed populations and implications for clinical epidemiology and forensic medicine. *Hum Genet* 2005; 118(3-4): p. 382-92.