A grand challenge for nutrigenomics

The low hanging fruit in genetics research has mostly been harvested, and now the work of studying genes as part of systems biology is well underway. Metabolism and nutrition seem to be an ideal complex system in which to apply the knowledge and methods of genetics and genomics. Diet is perhaps the most important environmental factor we are exposed to. Nutritional factors are thought to be the cause of 30–60% of cancers (similar in magnitude to smoking; Doll, 1992), diabetes is a nutritional/metabolic disorder, we are all familiar with the cholesterol–cardiovascular disease relationship, and obesity is a pressing nutritional problem, with a majority of Americans overweight and expanding rapidly (Ogden et al., 2007). If genetic scientists are looking for a complex system to step up to and address, what better choice? Hence, the developing discipline of nutrigenomics.

The first challenge for scientists entering this young discipline is to develop the fundamental knowledge base needed to start to address this complex system. Nutrition scientists are learning genetics, genetic scientists are mastering the complexities of metabolism, and bioinformatics scientists are learning the " languages" of both genetics and nutrition. This is the nature of any hybrid discipline in its early stages. Once such a cohort of multidisciplinary experts exists, nutrigenomics will be poised to address its grand challenge.

Before practical applications of nutrigenomics can arise (such as individualized nutrition recommendations or interventions designed to modify disease risks), the discipline must enhance understanding of bidirectional nutrient–gene interactions to the point where there are

overarching and integrated frameworks for thinking about how these interactions work. For example, advances in genetic science accelerated greatly once a catalog of the underlying basic mechanisms controlling gene expression was assembled. The grand challenge for nutrigenomics is to assemble such a catalog for nutrient–gene interactions, then use these data to develop an integrated framework for how the complex system works, and then test these hypotheses. The approach to answering this grand challenge is incremental.

Nutrigenomics is being applied to agriculture (enhanced plants and animal food sources) and to human health. For the purposes of this discussion, the application to individual differences in metabolism (individualized nutrition) is used to highlight how the field could proceed to address this challenge. The underlying mechanisms responsible for individual variation in metabolism and therefore, in the responses to and requirements for nutrients, are not yet fully known. The likely involvement of genetic variation and epigenetic mechanisms make these prime targets for study. In order to move forward and develop an overarching theory for predicting effects of genetic (epigenetic) variation on metabolism and nutrient requirements, a more complete data set describing nutrient–gene interactions is required. Without this information any practical application to humans will be flawed.

To date, there are thousands of genes for which there is no known function. Many of these genes will likely be important for metabolism. A systematic approach is needed to discover the function of these genes. For more than 9000 genes, the United States National Institutes of Health's Knockout Mouse Phenotyping Project as well as the EUMODIC consortium of 18

laboratories across Europe, are characterizing the functional effects of gene deletions in mice on phenotype, and hopefully they will comprehensively assess metabolic phenotypes.

The current catalog of genetic variations [single nucleotide polymorphisms (SNPs) and copy number variations (CNVs)] that cause metabolic inefficiencies is quite small, and most of the published literature considers SNPs as acting individually, rather than examining the systems effects of combinations of SNPs. There must be thousands of SNPs that alter metabolism, yet today, in the published literature there may be as many as 200 SNPs for which there are proven metabolic effects, and only a subset of these alter nutrient requirements in a significant portion of the population [e. g., the rs1801133 *MTHFR* SNP and folate requirement in 15–30% of the population (Solis et al., 2008) and the rs12325817 *PEMT* SNP and choline requirement in 20–45% of the population (da Costa et al., 2006)]. Some SNPs directly alter a metabolic response to a nutrient, rather than changing the requirement for it. For example, the rs3135506 SNP in *APOA5* modifies the effects of a high fat diet on blood pressure (Mattei et al., 2009). Just as the catalog of SNP-diet response relationships is incomplete, so, too is the collection of information regarding how diet may alter epigenetic marks in DNA and histones; published studies mainly focus on dietary methyl-donors (Waterland and Jirtle, 2003 ; Mehedint et al., 2010 and dietary fat Lomba et al., 2010 ; Vucetic et al., 2010). MicroRNA (miRNA) expression may be influenced by dietary manipulation (Alisi et al., 2010 ; Parra et al., 2010), but little data are available describing miRNA-level modulation of genes of metabolism. Finally, genetic variation influences eating behaviors (Dotson et

al., 2010 ; Fan et al., 2010) but these effects have not been systematically explored (this is an exciting potential collaboration between nutrigenomics and behavioral science). In part these catalogs are skimpy because nutrigenomic methods are relatively expensive when applied to large populations, and nutrition epidemiology studies often do not make genetic and metabolomic measures together.

Current methods in genomics and epigenomics are evolving rapidly. A new generation of techniques for high throughput gene sequencing is just appearing, and these methods could offer capacity and cost structures that make it practical to study genetic variation and epigenetic marks in large numbers of people. The challenge will be to design practical approaches for studying many people's samples in one sequencing batch, thereby amortizing the relatively high cost of gene sequencing. Barcoding techniques (Matsumura et al., 2010) (in which an identifying sequence of nucleotides is attached to each individual's DNA, which is then mixed together with many subjects' DNA, sequenced, and then sorted using the barcodes so that an individuals' gene sequences can be assembled) offer the promise for cost-efficient use of sequencing in nutrigenomic studies. Also, barcoding will make high throughput methylated gene sequencing possible (Varley and Mitra, 2010). The challenge will be to adapt these methods so that they are validated for use in nutrigenomic studies.

The methods for assessing gene variation and epigenetic marks are far more mature than are the high throughput methods for studying metabolism. Traditionally, many studies of genes and metabolism used a targeted approach, measuring a small set of metabolites in pathways likely to be

related to a change in a gene. Metabolomic profiling (also called metabonomic profiling) has matured sufficiently that it can be used to simultaneously measure many untargeted small molecules that are the product of many pathways in metabolism. Mass spectroscopic (Lawton et al., 2008) and nuclear magnetic resonance spectroscopic (Dumas et al., 2006) techniques are used in metabolomic profiling, each having distinct advantages and disadvantages. These methods offer an opportunity to more precisely measure the effects of genes on metabolism.

Unfortunately, measuring the effects of diet on genes faces a major challenge: the methods for measuring dietary intake are much more imprecise than are genetic or biochemical measurements. There is a dearth of novel methodology for measuring dietary exposures and there is a real need for an improved tool chest for those investigating how diet interacts with genotype to determine phenotype (Penn et al., 2010). Most nutrition scientists recognize that current diet assessment methods are not likely to provide more than an approximation (perhaps off by 30% or more) for individual dietary intake (Penn et al., 2010). Though the development of each of these methodologies are the focus of different disciplines, nutrigenomic scientists face the significant challenge of improving and adapting these methods in combination with genomic methods so that a more complete catalog of nutrient–gene interactions can be developed.

Interpretation of measurements of tens of thousands of genes, with millions of genetic variations, some of which have effects on hundreds or thousands of the small molecules of metabolism (and perhaps on thousands of peptides or proteins involved in metabolism) requires new bioinformatic and statistical

approaches. The bioinformatic tool chest for genomics is far more mature than the metabolomic profiling tool chest in this regard, and there are few integrated bioinformatic approaches available where changes in genes are superimposed on changes in metabolic pathways. It is difficult at this time to visualize the combined data, and it is even more difficult to analyze, in an integrated manner, how multiple genes and multiple nutrients interact simultaneously. The NUGO project in Europe is beginning to perform such complex analyses (Coort et al., 2008), but there is a long way to go before we answer this challenge.

The grand challenge – develop an overarching and integrated framework for thinking about how gene–nutrient interactions influence metabolism. To address this challenge, nutrigenomics has to build a cohort of multidisciplinary scientists who can talk each other's " language," develop better methods, and establish a comprehensive data set from which the great thinkers among us can develop the theory. These tasks should keep us busy for a decade or more.

Acknowledgment

References
Alisi, A., Da Sacco, L., Bruscalupi, G., Piemonte, F., Panera, N., De Vito, R., Leoni, S., Bottazzo, G. F., Masotti, A., and Nobili, V. (2010). Mirnome analysis reveals novel molecular determinants in the pathogenesis of diet-induced nonalcoholic fatty liver disease. *Lab Invest.* doi: 10. 1038/labinvest. 2010. 166. [Epub ahead of print].

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Coort, S. L., van Iersel, M. P., van Erk, M., Kooistra, T., Kleemann, R., and Evelo, C. T. (2008). Bioinformatics for the NuGO proof of principle study: analysis of gene expression in muscle of ApoE3*Leiden mice on a high-fat diet using PathVisio. *Genes Nutr* . 3, 185–191.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

da Costa, K. A., Kozyreva, O. G., Song, J., Galanko, J. A., Fischer, L. M., and Zeisel, S. H. (2006). Common genetic polymorphisms affect the human requirement for the nutrient choline. *FASEB J.* 20, 1336–1344.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Doll, R. (1992). The lessons of life: keynote address to the nutrition and cancer conference. *Cancer Res.* 52(Suppl.), 2024s–2029s.

Pubmed Abstract | Pubmed Full Text

Dotson, C. D., Shaw, H. L., Mitchell, B. D., Munger, S. D., and Steinle, N. I. (2010). Variation in the gene TAS2R38 is associated with the eating behavior disinhibition in old order Amish women. *Appetite* 54, 93–99.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Dumas, M. E., Barton, R. H., Toye, A., Cloarec, O., Blancher, C., Rothwell, A., Fearnside, J., Tatoud, R., Blanc, V., Lindon, J. C., Mitchell, S. C., Holmes, E., McCarthy, M. I., Scott, J., Gauguier, D., and Nicholson, J. K. (2006). Metabolic

profiling reveals a contribution of gut microbiota to fatty liver phenotype in insulin-resistant mice. *Proc. Natl. Acad. Sci. U. S. A.* 103, 12511–12516.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Fan, B., Lkhagvadorj, S., Cai, W., Young, J., Smith, R. M., Dekkers, J. C., Huff-Lonergan, E., Lonergan, S. M., and Rothschild, M. F. (2010). Identification of genetic markers associated with residual feed intake and meat quality traits in the pig. *Meat Sci.* 84, 645–650.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Lawton, K. A., Berger, A., Mitchell, M., Milgram, K. E., Evans, A. M., Guo, L., Hanson, R. W., Kalhan, S. C., Ryals, J. A., and Milburn, M. V. (2008). Analysis of the adult human plasma metabolome. *Pharmacogenomics* 9, 383–397.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Lomba, A., Martinez, J. A., Garcia-Diaz, D. F., Paternain, L., Marti, A., Campion, J., and Milagro, F. I. (2010). Weight gain induced by an isocaloric pair-fed high fat diet: a nutriepigenetic study on FASN and NDUFB6 gene promoters. *Mol. Genet. Metab.* 101, 273–278.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Matsumura, H., Yoshida, K., Luo, S., Kimura, E., Fujibe, T., Albertyn, Z., Barrero, R. A., Kruger, D. H., Kahl, G., Schroth, G. P., and Terauchi, R. (2010). High-throughput SuperSAGE for digital gene expression analysis of multiple samples using next generation sequencing. *PLoS ONE* 5, e12010. doi: 10. 1371/journal. pone. 0012010

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Mattei, J., Demissie, S., Tucker, K. L., and Ordovas, J. M. (2009). Apolipoprotein A5 polymorphisms interact with total dietary fat intake in association with markers of metabolic syndrome in Puerto Rican older adults. *J. Nutr.* 139, 2301–2308.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Mehedint, M. G., Niculescu, M. D., Craciunescu, C. N., and Zeisel, S. H. (2010). Choline deficiency alters global histone methylation and epigenetic marking at the Re1 site of the calbindin 1 gene. *FASEB J.* 24, 184–195.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Ogden, C. L., Yanovski, S. Z., Carroll, M. D., and Flegal, K. M. (2007). The epidemiology of obesity. *Gastroenterology* 132, 2087–2102.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Parra, P., Serra, F., and Palou, A. (2010). Expression of adipose microRNAs is sensitive to dietary conjugated linoleic acid treatment in mice. *PLoS ONE* 5, e13005. doi: e13005 [pii] 10. 1371/journal. pone. 0013005

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text

Penn, L., Boeing, H., Boushey, C. J., Dragsted, L. O., Kaput, J., Scalbert, A., Welch, A. A., and Mathers, J. C. (2010). Assessment of dietary intake: NuGO symposium report. *Genes Nutr* . 5, 205–213.

CrossRef Full Text

Solis, C., Veenema, K., Ivanov, A. A., Tran, S., Li, R., Wang, W., Moriarty, D. J., Maletz, C. V., and Caudill, M. A. (2008). Folate intake at RDA levels is inadequate for Mexican American men with the methylenetetrahydrofolate reductase 677TT genotype. *J. Nutr.* 138, 67–72.

Varley, K. E., and Mitra, R. D. (2010). Bisulfite patch PCR enables multiplexed sequencing of promoter methylation across cancer samples. *Genome Res.* 20, 1279–1287.

Pubmed Abstract | Pubmed Full Text

Vucetic, Z., Kimmel, J., Totoki, K., Hollenbeck, E., and Reyes, T. M. (2010). Maternal high-fat diet alters methylation and gene expression of dopamine and opioid-related genes. *Endocrinology* 151, 4756–4764.

Pubmed Abstract | Pubmed Full Text

Waterland, R. A., and Jirtle, R. L. (2003). Transposable elements: targets for early nutritional effects on epigenetic gene regulation. *Mol. Cell. Biol.* 23, 5293–5300.

Pubmed Abstract | Pubmed Full Text | CrossRef Full Text