Published in final edited form as:

Obesity (Silver Spring). 2008 December ; 16(Suppl 3): S79–S81. doi:10.1038/oby.2008.523.

Obesity Genes and Gene–Environment–Behavior Interactions: Recommendations for a Way Forward

Alan R. Shuldiner¹

¹*Division of Endocrinology, Diabetes and Nutrition, University of Maryland School of Medicine, Baltimore, Maryland, USA*

Abstract

Obesity is a classical complex trait, influenced by both genetic and lifestyle factors. The number of obesity gene variants is currently unknown but, based on sound evolutionary principles, likely to be many, each with a modest effect on the phenotype. Recent advances in our knowledge of variation in the human genome and high throughput genotyping technologies have made possible genome-wide association (GWA) analysis and the identification of *bona fide* susceptibility genes for many complex diseases and phenotypes, including obesity and its comorbid conditions. GWA analysis in even larger numbers of individuals through collaborative efforts of many investigators will likely identify those polygenes of moderate and modest effect size that manifest in our typical environment. Once the subset of real-world-relevant obesity susceptibility variants is identified, follow-up studies, including detailed molecular analysis of the loci, stratified analyses, prospective and interventional studies in humans, and mechanistic studies in cells and animals will allow us to define the genetic architecture of the locus and dissect how these genes interact with specific environmental and other factors. The molecular and analytical tools to accomplish these goals are now in hand, but cooperation among investigators will be necessary to amass the requisite numbers of phenotyped and genotyped individuals. Identification of susceptibility genes for obesity and determining how they interact with each other and the environment will lead to new insights into the molecular, cellular, and physiological basis of energy homeostasis, and novel strategies for prevention and treatment.

Introduction: How Complex a Trait is Obesity?

This two-day meeting of experts in the area of genetics, energy homeostasis, and eating behavior highlighted the complexity and multifactorial nature of obesity and its complications. Clearly, both genetics and the environment play important roles, with gene variants having been selected over the course of about 1 billion years of evolution, followed by the expression of these “thrifty” genes in an obesogenic environment over just the last 50–100 years (1). One billion years of selection for thrifty genes would predict that obesity susceptibility genes in humans will be many (perhaps 100–1,000) and most will have modest effects. Furthermore, there will likely be great genetic heterogeneity both within and between populations, and important gene–gene and gene–environment interactions. Clinically, we know that there is great interindividual variation not only in susceptibility to obesity in a similar obesogenic environment, but also in where body fat accretion occurs (subcutaneous vs. visceral), and to what extent the obesity-related comorbid conditions such as insulin resistance, glucose intolerance, hypertension, hyperlipidemia, sleep apnea, and cancer will manifest. Furthermore, there is great interindividual variation in response to lifestyle and pharmacological

Correspondence: Alan R. Shuldiner (E-mail: Ashuldin@medicine.umaryland.edu).

Disclosure

A.R.S. has received consulting fees from Johnson & Johnson and Merck.

interventions to prevent and treat obesity, further demonstrating enormous heterogeneity of the phenotype, which is due in part to genetics. Clearly, finding the obesity susceptibility genes as well as those that predispose to the malignant comorbidities of obesity will be challenging.

Current State of Knowledge

Many candidate genes for obesity based on our current understanding of energy homeostasis have been studied, but results to date have been disappointing (reviewed in ref. 2). This may be due to the fact that our current understanding of the pathophysiology of energy homeostasis is incomplete and thus the correct candidate genes have not been studied. Furthermore, most candidate genes that have been studied have not been investigated thoroughly based on today's understanding of human genome variation. Finally, most of these studies have been severely underpowered to detect modest effects of single gene variants. Genome-wide linkage analysis has similarly been disappointing in identifying replicated linkage signals and true obesity susceptibility gene variants, due in large part to the limited power of linkage analysis (compared to association analysis) to detect common susceptibility alleles.

A Proposed Roadmap Forward

Step 1. Find the obesity susceptibility genes: genome-wide association scans

Recent advances in our knowledge of the human genome and relatively inexpensive high throughput single nucleotide polymorphism (SNP) genotyping and sequencing technologies have made possible powerful genome-wide association studies of complex traits and diseases. With this approach, hundreds of thousands of SNPs are genotyped across the genome in cases and controls, and association between each SNP and disease status is assessed. Genome-wide association studies have proven successful in identifying many complex disease genes including cardiovascular disease (3,4), type 2 diabetes (5,6,7,8), multiple sclerosis (9), and cancer (10,11). To date, two genes that influence body mass, insulin induced gene 2 (*INSIG2*) and fat mass and obesity-associated (*FTO*) genes, have been identified by this approach (12,13). A great strength of genome-wide approaches is that the search for susceptibility genes is agnostic (hypothesis-free) and thus not limited by our narrow understanding of the pathophysiology of the disease. Indeed, genome-wide association studies have identified novel disease susceptibility genes in pathways not previously known to be involved in pathophysiological processes, e.g., *FTO* and obesity; cyclin-dependent kinase inhibitor (*CDKN*) 2A and 2B with type 2 diabetes and myocardial infarction.

The successful genome-wide association scans have also taught us that large numbers of cases and controls are necessary due to the modest effects of gene variants on disease phenotypes that we wish to detect, and also due to type 1 error resulting from multiple comparisons (hundreds of thousands with high-density SNP arrays) (14,15). Large numbers of cases and controls may also overcome “noise” of inaccurate phenotypes, variation in environmental exposures, and other potential confounders. Furthermore, replication cohorts are critical to teasing out true positive signals from the large number of false positive signals of initial genome-wide association scans (16).

The first step toward identifying obesity susceptibility genes should be to perform a genome-wide association scan for the relatively crude obesity surrogate, BMI, in a large number of subjects in the setting of the real-world heterogeneous environment to uncover those obesity susceptibility genes that are germane to the common forms of obesity and that interact with the relevant exposures of the real-world environment. There are many large population-based studies that have been conducted throughout the world in which BMI has been measured. Some of these studies have collected DNA and/or have already performed genome-wide SNP genotyping. Thus, a genome-wide association scan for BMI in tens (or even hundreds) of

thousands of individuals can be performed relatively quickly and with relatively little additional expense. This approach has recently been used to identify two novel gene variants associated with height (17,18), a notoriously heterogeneous polygenic trait for which environmental context plays an important role.

Step 2. Define the genetic architecture of the obesity susceptibility genes

Although genome-wide association scans have been quite successful in identifying SNPs associated with disease, in the vast majority of instances, these SNPs are not the functional variants, but rather in linkage disequilibrium with the causative variant. Thus detecting a robust statistical association is the first step toward defining the gene and its genetic architecture, and exhaustive analysis of the obesity-associated chromosomal region is a next step. Deep resequencing will likely divulge not only the common pathogenic variant that was in linkage disequilibrium with the initial association signal, but also additional disease-associated variants including common and rare SNPs not tagged by the initial genome-wide scan, as well as insertion/deletion and copy number variants. Complete characterization of the genetic architecture of a given obesity susceptibility gene will require its detailed analysis across multiple ethnic populations.

Step 3. Define the functional consequences of the causative variants

As described earlier, the hypothesis-free approach of genome-wide association scans will undoubtedly identify novel genes not previously thought to be involved in energy homeostasis. Thus for many of the obesity susceptibility genes identified, it will be necessary to define how these genes and their variants alter known and unknown pathways. These studies will require well-designed hypothesis-driven experimentation using molecular and cellular *in vitro* study designs (e.g., gene expression and/or functional analysis of protein sequence-altering variants), model organisms (e.g., knockouts and transgenics), and detailed physiological studies in humans recruited by genotype.

Step 4. Identify the relevant gene–environment interactions

Despite remarkable progress in our understanding of the human genome, there has been relatively little progress with regard to technologies to accurately measure chronic environmental exposures relevant to obesity, e.g., diet, eating behavior, physical activity, built environment, social networks. Given the inaccuracies of current technologies to measure environmental exposures, our ability to directly query the entire genome for gene–environment interactions will likely be fruitless. One approach to identifying gene–environment and gene–behavior interactions relevant to obesity is to first identify the obesity susceptibility genes. Once *bona fide* genes for disease susceptibility are identified, hypothesis-driven studies of gene–environment interactions can be performed. Prospective and interventional studies of humans of differing genotypes and epidemiological studies in which exposure variables have been measured will uncover the important gene–environment interactions, including insights into variability in drug response (pharmacogenomics). For example, prediabetic subjects with the transcription factor 7-like 2 (*TCF7L2*) risk allele can ameliorate their increased risk through diet and exercise (19), and those with type 2 diabetes may be less likely to respond to sulfonylureas (20). Similarly, we (21) and others (22) have shown that increased physical activity can attenuate the increased risk for obesity in subjects with the *FTO* risk allele.

Step 5. Translation to clinical practice

The quest to identify and understand most or all of the important obesity susceptibility genes is driven by their application to clinical practice to improve the quality and quantity of life for the millions of adults and children who are afflicted or at risk. Obesity susceptibility gene panels may be useful in a clinical setting to determine which individuals are most likely to

develop obesity and obesity-associated comorbidities. This may aid in targeting interventions to prevent or delay its onset. Identification of novel genes and pathways may pave the way for the design of new drugs based on these targets (reverse pharmacogenomics). This would be a welcome step forward since current pharmacotherapy for obesity is inadequate. However, many questions remain. For example, what will the predictive value of such an obesity susceptibility panel be? If an individual knows they have a certain obesity susceptibility genotype, would they be more or less likely to engage in a preventive intervention such as diet or exercise? Will physicians have the knowledge to interpret the results of predictive gene panels and provide appropriate recommendations to their patients? Answers to these questions will require well-designed randomized clinical trials, thoughtful strategies to teach complex disease genetic medicine to medical students, efforts to retool practicing physicians, and education of the lay public.

Conclusions

Obesity is likely the result of complex interactions between many “thrifty” genotypes in context with our obesogenic environment. Efforts to date to identify functional obesity-susceptibility variants have been disappointing. Clearly the field has underestimated the complexity of the genetic underpinnings of obesity. Recent genome-wide association studies have identified gene variants for several complex diseases including two (to date) for obesity. Further application of this powerful approach, especially in large number of subjects acquired through collaborative efforts of multiple investigators, will likely lead to the identification of many more novel susceptibility genes. Subsequent study of these genes to elucidate their genetic architecture, their roles in energy homeostasis, and gene–gene and gene–environment interactions will provide new insights and opportunities for translation into more effective modalities for treatment and prevention.

Acknowledgments

This publication was sponsored by the National Cancer Institute (NCI) to present the talks from the “Gene–Nutrition and Gene–Physical Activity Interactions in the Etiology of Obesity” workshop held on 24–25 September 2007. The opinions or assertions contained herein are the views of the authors and are not to be considered as official or reflecting the views of the National Institutes of Health.

References

1. Neel JV. Diabetes mellitus: a thrifty genotype rendered detrimental by ‘progress’. *Am J Hum Genet* 1962;4:352–353.
2. Rankinen T, Zuberi A, Chagnon YC, et al. The human obesity gene map: the 2005 update. *Obesity* 2006;14:529–544. [PubMed: 16741264]
3. Samani NJ, Erdmann J, Hall AS, et al. Genomewide association analysis of coronary artery disease. *N Engl J Med* 2007;357:443–453. [PubMed: 17634449]
4. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;447:661–678. [PubMed: 17554300]
5. Zeggini E, Weedon MN, Lindgren CM, et al. Replication of genome-wide association signals in U.K. samples reveals risk loci for type 2 diabetes. *Science* 2007;316:1336–1341. [PubMed: 17463249]
6. Sladek R, Rocheleau G, Rung J, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 2007;445:881–885. [PubMed: 17293876]
7. Scott LJ, Mohlke KL, Bonnycastle LL, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* 2007;316:1341–1345. [PubMed: 17463248]
8. Saxena R, Voight BF, Lyssenko V, et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 2007;316:1331–1336. [PubMed: 17463246]
9. Hafler DA, Compston A, Sawcer S, et al. Risk alleles for multiple sclerosis identified by a genomewide study. *N Engl J Med* 2007;357:851–862. [PubMed: 17660530]

10. Hunter DJ, Kraft P, Jacobs KB, et al. A genome-wide association study identifies alleles in *FGFR2* associated with risk of sporadic postmenopausal breast cancer. *Nat Genet* 2007;39:870–874. [PubMed: 17529973]
11. Broderick P, Carvajal-Carmona L, Pittman AM, et al. A genome-wide association study shows that common alleles of *SMAD7* influence colorectal cancer risk. *Nat Genet* 2007;39:1315–1317. [PubMed: 17934461]
12. Herbert A, Gerry NP, McQueen MB, et al. A common genetic variant is associated with adult and childhood obesity. *Science* 2006;312:279–283. [PubMed: 16614226]
13. Frayling TM, Timpson NJ, Weedon MN, et al. A common variant in the *FTO* gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 2007;316:889–894. [PubMed: 17434869]
14. Hunter DJ, Kraft P. Drinking from the fire hose—statistical issues in genomewide association studies. *N Engl J Med* 2007;357:436–439. [PubMed: 17634446]
15. Manolio TA, Rodriguez LL, Brooks L, et al. New models of collaboration in genome-wide association studies: the Genetic Association Information Network. *Nat Genet* 2007;39:1045–1051. [PubMed: 17728769]
16. Altshuler D, Daly M. Guilt beyond a reasonable doubt. *Nat Genet* 2007;39:813–815. [PubMed: 17597768]
17. Sanna S, Jackson AU, Nagaraja R, et al. Association of human height with genetic variants in the *GDF5-BFZB* locus: genome-wide scans and replication. *Nat Genet* 2008;40:198–203. [PubMed: 18193045]
18. Weedon MN, Lettre G, Freathy RM, et al. A common variant of *HMGA2* is associated with adult and childhood height in the general population. *Nat Genet* 2007;39:1245–1250. [PubMed: 17767157]
19. Florez JC, Jablonski KA, Bayley N, et al. *TCF7L2* polymorphisms and progression to diabetes in the Diabetes Prevention Program. *N Engl J Med* 2006;355:241–250. [PubMed: 16855264]
20. Pearson ER, Donnelly LA, Kimber C, et al. Variation in *TCF7L2* influences therapeutic response to sulfonylureas: a GoDARTs study. *Diabetes* 2007;56:2178–2182. [PubMed: 17519421]
21. Rumpersaud E, Mitchell BD, Pollin TI, et al. Physical activity modifies the effect of common *FTO* gene variants on body mass index and obesity. *Arch Intern Med* 2008;168:1791–1797. [PubMed: 18779467]
22. Andreasen CH, Stender-Petersen KL, Mogensen MS, et al. Low physical activity accentuates the effect of the *FTO* rs9939609 polymorphism on body fat accumulation. *Diabetes* 2008;57:95–101. [PubMed: 17942823]