Dissecting gene–environment contributions to Type 2 diabetes

Braxton D Mitchell*[1]: Braxton Mitchell is professor of medicine and of epidemiology & public health at the University of Maryland School of Medicine in Baltimore, MD, USA. He obtained his PhD degree at the University of Michigan, MI, USA. He is a genetic epidemiologist whose research program focuses on assessing the genetic contributions to common age-related diseases, gene mapping and evaluating the impact of discovered variants at the population level. In addition to diabetes, his current research focuses mainly in the areas of cardiovascular disease, stroke, osteoporosis and osteoarthritis.

Robert L Hanson[2]: Robert Hanson is staff clinician at the Phoenix Epidemiology and Clinical Research Branch of the National Institute of Diabetes and Digestive and Kidney Diseases in Phoenix, AZ, USA. He obtained his MD degree at the University of Kansas (KS, USA) and his MPH degree at Columbia University (NY, USA). He is an internist and epidemiologist and his research program focuses on the genetic epidemiology of Type 2 diabetes, obesity and complications of diabetes.

Q What are gene–environment interactions in the context of diabetes?

Gene–environment interactions represent situations in which there is a genetic effect on disease risk, an environmental effect on disease risk and an additional effect associated with having joint genetic and environmental exposure. A classic example of a gene–environment interaction is phenylketonuria; a disease caused by a mutation in the gene encoding the enzyme phenylalanine hydroxylase, and in which the resulting enzyme deficiency prevents the metabolism of the amino acid phenylalanine. In the presence of a normal protein diet, this mutation causes a build-up of phenylalanine in the body, disruption of metabolism in the brain and development of neurological defects, including mental retardation, brain damage and seizures. While there is no cure for phenylketonuria, these adverse consequences can be prevented by adherence to a low-phenylalanine diet. In other words, the genetic defect can be completely overcome by an environmental intervention.

In the context of Type 2 diabetes, perhaps the best known example of a gene–environment interaction involves the obesity susceptibility gene, *FTO*. The risk-increasing allele at this locus has a frequency of 45–50% in European Caucasians and African–Americans, with each copy of the risk allele associated with a 1–1.5 kg increase in body weight. However, multiple

[1]Departments of Medicine and Epidemiology & Public Health, University of Maryland School of Medicine, 685 W. Baltimore St., MSTF 302. Baltimore, MD 21201, USA
[2]NIDDK/PECRB/DECRS, 1550 E. Indian School Rd, Phoenix, AZ 85104, USA
*Author for correspondence: bmitchel@medicine.umaryland.edu

Future Medicine part of fsg

studies have now demonstrated that the effect of the risk allele on bodyweight is muted in individuals with high physical activity levels; that is, the risk allele has very little effect on individuals who are highly physically active, but rather its effect is restricted mainly to those with a sedentary lifestyle. This is a clear example of how environment and lifestyle can trump genetic susceptibility in Type 2 diabetes.

Q How significantly do these aspects contribute to Type 2 diabetes? To what extent are individuals genetically susceptible to diabetes?

From twin and family studies, it is estimated that genes account for 40–60% of the risk for developing Type 2 diabetes. The big challenge is now to uncover the genetic architecture of Type 2 diabetes, in the hope that identifying the specific genes that predispose to Type 2 diabetes will teach us more about the pathogenesis and molecular causes of the disease, possibly leading to new treatments and prevention strategies.

Our best evidence is that many different genes with very small effects contribute to Type 2 diabetes risk, rather than a small number of genes with large effects. It further appears that these small effect genes all have a low penetrance – that is, disease risk does not automatically increase for carriers of risk alleles, but the effects of these risk alleles (such as in *FTO*) become especially apparent in individuals who are at risk due to their environmental (e.g., lifestyle) risk factors. This brings us back to gene–environment interactions. While we do not have reliable estimates regarding how much of Type 2 diabetes risk is due to gene–environment interactions over and above those due to genes alone and environment alone, there is reason to think that the contribution of gene–environment interactions may be substantial. This speculation is supported by the relatively rapid increase in Type 2 diabetes incidence that is seen worldwide with the transition to more sedentary lifestyles and westernized diets.

Q Which Type 2 diabetes loci have been found recently?

To date, approximately 55 loci have been associated with Type 2 diabetes through large genome-wide association studies (GWAS) and meta-analyses of GWAS that have included very large numbers of cases and controls. Some of these loci have been mapped to genes that influence pancreatic β-cell function, while a few are in genes established to influence insulin action. The mechanism of action of many of the other discovered genes is unknown (for a description of GWAS findings and diabetes, see [1]).

Among European-derived populations, the largest single locus effect discovered is for a variant within the *TCF7L2* gene, which encodes a transcription factor that is active in the Wnt-signaling pathway that was found serendipitously through a large-scale association study of available microsatellite markers. Each copy of the *TCF7L2* risk allele increases the odds of developing Type 2 diabetes by 1.37. The odds ratios for the remaining Type 2 diabetes susceptibility loci range from 1.1 to 1.3. As additional GWAS are carried out and meta-analyses reanalyzed to include these larger sample sizes, more Type 2 diabetes loci will undoubtedly be identified, although one can expect these to have smaller and smaller effect sizes.

Q How useful are GWAS for identifying candidate genes in diabetes research?

The effect sizes of these variants are relatively small and, even collectively, the 55 or so variants discovered so far account for less than 10% of disease susceptibility. Because of their small effect sizes and low penetrance, these variants, even collectively, are not yet useful for clinical prediction, particularly if one accounts for clinical variables such as BMI and family history of diabetes. It must be recognized, however, that most of the diabetes susceptibility variants discovered so far are biased towards high-frequency, low-penetrance variants because these are the ones represented on the large GWAS arrays. However, from these discoveries, what we have learned so far is that most of the discovered GWA variants affect insulin secretion, not insulin action; implying that compromised β-cell function leaves one susceptible to other diabetogenic stresses. We have also learned that some of the identified Type 2 diabetes susceptibility loci harbor other

(rarer) variants that have previously been associated with monogenic forms of disease (e.g., *PPARG*, *KCNJ11*, *HNF1B* and *WFS1*), and that some Type 2 diabetes susceptibility loci have also been implicated in other diseases (e.g., the 9p21 locus and cardiovascular disease/aneurysms, *HNF1B*/*JAZF1* and prostate cancer, and *CDKAL1* and Crohn's disease; for more discussion of GWAS and diabetes, see [2]).

Q What is genomic imprinting in diabetes?

Normally, we inherit two functional copies of our genes, one from each parent. For a subset of autosomal genes, either the maternally or paternally transmitted copy of the gene may be turned off in some tissues, so that only one of the parental copies is expressed. Because only one copy of the gene is expressed, the gene may be particularly susceptible to the effects of sequence variants or epigenomic influences that impact gene function, since the paired copy has been rendered inactive and cannot help compensate. Genomic imprinting is the process by which the maternal or paternal copy of the gene is epigenetically marked or imprinted, rendering it inactive.

Imprinting usually occurs as a result of one copy of the gene being silenced. One way that this can happen is via methylation of the DNA around the gene or via modification of the histones that package and order the DNA. While methylation typically silences gene expression, it can also enhance gene expression. For example, methylation of the paternal copy of a region on chromosome 11p15 prevents an 'insulator' from binding to the region, which allows enhancers to access and promote expression of the imprinted gene *IGF2* on the paternal chromosome.

Many of the genomic imprinting disorders identified so far in humans affect growth and development (e.g., Prader–Willi and Angelman syndromes, Beckwith–Wiedemann syndrome, Russell–Silver syndrome and Albright hereditary osteodystrophy). Abnormalities in an imprinted gene region on chromosome 6q24 are now known to be associated with transient neonatal diabetes mellitus. In addition, strong parent-of-origin effects have been reported for variants in four different loci associated with Type 2 diabetes in a study from Iceland, suggesting that imprinted genes may play an important role in the common form of Type 2 diabetes. High-density arrays that assess methylation patterns across the genome have recently become available and are now starting to be used to assess the role of methylation in Type 2 diabetes (for more discussion of imprinting in diabetes, see[3]).

Q What approaches are used to dissect the genetic and environmental determinants of diabetes and obesity?

GWAS have been used to identify most genetic loci, while environmental risk factors are generally identified through classic risk-factor epidemiology. Traditionally, family-based studies have been used to dissect the relative contribution of genetic and environmental determinants of diseases, such as diabetes and obesity. The extent to which phenotypic similarity tracks with genetic similarity within families is used to infer the potential influence of genetic factors on the trait. In classic twin studies, which compare concordance rates between monozygotic and dizygotic twins, one can separate out the contributions to disease susceptibility from genetic factors, environmental factors shared within families and environmental factors unique to the individual. With the advent of GWAS, methods have recently been developed to assess the influence of genetic factors based on the genetic similarity among individuals across all molecular markers, even in 'unrelated' individuals. Analyses of traits such as height and BMI using these methods have revealed that the effects at markers that did not meet the stringent statistical criteria for genome significance explain a substantial portion of the variation in disease susceptibility. Future inclusion of sequence information in these analyses, which will provide more information about rare variants, will probably further increase this proportion. In the context of gene–environment interaction, however, we may find that a large portion of the variability in susceptibility cannot be assigned strictly to 'genetic' or 'environmental' factors, but instead may be attributable to both factors acting simultaneously.

Q What has been the most striking/significant research in this field to date?

The most striking findings are just how many Type 2 diabetes loci have been identified in a relatively short time and the fact that the effect sizes for most of the variants are modest. This suggests that the underlying genetics of Type 2 diabetes (and of obesity) may be reflective of 'polygenic' effects. At present, the extent to which the genetic architecture of Type 2 diabetes reflects the action of many common variants with relatively small effects or the action of many relatively rare variants with relatively large effects (though restricted to few individuals) is uncertain. As studies using DNA sequencing technologies are undertaken in larger groups of individuals, this will become clearer.

Q How have studies in this field benefited clinical practice for diabetics?

At present, there is little direct clinical utility from genetic studies of Type 2 diabetes, at least with respect to 'bedside' genetic testing. Exceptions include identification of individuals with monogenic forms of diabetes, such as maturity-onset diabetes in the young or neonatal diabetes, as some genetic forms of these disorders will dictate appropriate therapy. Results from the Diabetes Prevention Program suggest that for the currently known diabetes susceptibility variants, those individuals with and without the variants respond equally well to lifestyle intervention, and it is helpful for clinicians to understand that high-risk patients may benefit from preventive measures, regardless of their risk profile at these variants. Ultimately, benefits from genetic research into diabetes and obesity may come more through insight into the molecular and physiological processes leading to these diseases than through direct clinical genetic testing.

Q What challenges remain in this field? What progress do you foresee happening in the next 10 years?

Epigenetics is an exciting new area of research in the field. Epigenetic changes, which can be passed on from parents to their offspring, refer to modifications of the genome that affect gene expression and/or phenotype but do not involve a change to the nucleotide sequence. These may be mediated by processes such as DNA methylation or histone modification. New technologies for epigenomic profiling have now made it possible to measure epigenetic changes to the genome in a much more comprehensive fashion than before, thus opening up opportunities for new studies to evaluate the causes and consequences of epigenetic changes. Such studies are likely to add new insights into the mechanisms by which psychosocial factors, pregnancy, sleep quality and a host of other factors contribute to diabetes and obesity susceptibility, possibly via their effects on the epigenome. These epigenetic factors may provide a critical link between environmental and genetic influences on disease, possibly identifying some of the molecular mechanisms underlying gene–environment interactions.

Financial & competing interests disclosure

BD Mitchell acknowledges the support of the NIH (grant numbers P60DK079637 and P30DK072488). RL Hanson is supported by the NIDDK intramural research program. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

References

1 Prokopenko I, McCarthy MI, Lindgren CM. Type 2 diabetes: new genes, new understanding. *Trends Genet.* 24, 613–621 (2008).

2 Billings LK, Florez JC. The genetics of Type 2 diabetes: what have we learned from GWAS? *Ann. NY Acad. Sci.* 1212, 59–77 (2010).

3 Mitchell BD, Pollin TI. Genomic imprinting in diabetes. *Genome Med.* 2(8), 55 (2010).