Genetic Factors in Type 2 Diabetes: The End of the Beginning?

Stephen O’Rahilly,1* Inês Barroso,2 Nicholas J. Wareham3

The intensive search for genetic variants that predispose to type 2 diabetes was launched with optimism, but progress has been slower than was hoped. Even so, major advances have been made in the understanding of monogenic forms of the disease which together represent a substantial health burden, and a few common gene variants that influence susceptibility have now been unequivocally identified. Armed with a better understanding of the tools needed to detect such genes, it seems inevitable that the rate of progress will increase and the relevance of genetic information to the diagnosis, treatment, and prevention of diabetes will become increasingly tangible.

Serious efforts to identify genetic variants that predispose to complex human diseases have been under way for 20 years or more. The birth of the global effort to determine the genetic basis of these diseases was accompanied by a fanfare of optimism, but as time has passed and major successes seem more elusive than was anticipated, notes of discord have been increasingly heard. Where are the personalized medicines and life-style interventions based on genotype? Could research funding have been better deployed? Type 2 diabetes is a complex disease that represents a major international public health threat. Although the current rise in its prevalence is almost certainly driven by life-style changes, the inherent susceptibility to the condition is widely considered to be attributable to complex genetic determinants. It is, therefore, an interesting test case in which to think critically about where we are now and how we might progress. To structure our thoughts, we have posed seven questions, which will be applicable, to varying degrees, to most other common diseases.

How Homogenous Is Type 2 Diabetes?

If type 2 diabetes were a homogenous condition with a common underlying molecular pathogenesis, then tracking down the common genetic variants that underlie susceptibility would be relatively easy. Unfortunately, it is more likely that the disease we call type 2 diabetes is heterogeneous and may result from defects in one or more diverse molecular pathways. The control of plasma glucose within tight limits requires precise, life-long coordination of the function of the pancreatic β cell with the responsiveness to insulin of major metabolic tissues such as muscle, liver, and fat. We have long known that there are multiple ways of producing a disease that looks like type 2 diabetes. Stimuli as diverse as glucocorticoid treatment and iron overload, for example, can result in a state of chronically elevated blood glucose that is hard to distinguish from type 2 diabetes. With the elucidation of the molecular bases of a range of monogenic disorders that result in diabetes, it has become clear that many affected members in these families may be diagnosed in middle life and look, superficially at least, much like patients with common type 2 diabetes. The heterogeneous nature of type 2 diabetes suggests that we may need to focus our attention on more phenotypically homogenous subgroups of individuals. It is possible, however, that classical phenotyping approaches will be too crude to act as a useful guide to etiological subgroups. If so, then a revised classification of diabetes will only be engineered the other way around, with improved nosology emerging from a better understanding of molecular pathogenesis.

How Important Is Heredity in Determining the Risk of Type 2 Diabetes?

This apparently simple question is actually very difficult to answer, because the contribution of heredity appears to differ considerably between different populations and in different environments. The best evidence that heredity plays an important role comes from the following observations: (i) Concordance rates for type 2 diabetes and its predecessor, impaired glucose tolerance, are consistently higher in monogygotic than in dizygotic twin pairs; (ii) sibling recurrence rates are consistently higher than population prevalence rates, although the reported excess is modest; (iii) groups of patients labeled as having type 2 diabetes include individuals suffering from unrecognized monogenic and digenic disorders; and (iv) certain common single-nucleotide polymorphisms (SNPs) appear to influence diabetes risk (1).

The influence of environment and lifestyle on risk should not be underestimated, however. Ecological studies, comparisons of migrant populations, cohort studies, and intervention trials clearly demonstrate that factors related to diet and physical activity have a major impact on the development of diabetes (2). More recent attention has focused on the possible effects of prenatal and early postnatal environment on diabetes risk (3). The correlation between low birth weight and later diabetes is consistent across many populations studied, but could, of course, imply that genetic factors influence both birth weight and diabetes risk. Indeed, Hattersley and colleagues have reported an elegant example of this, illustrating how mutations in the glucokinase gene of a fetus lead to reduced intrauterine growth, presumably as a result of relatively low fetal insulin levels, and also to postnatal diabetes (4). To counter this, however, Beck-Nielsen and colleagues have demonstrated that in a group of middle-aged identical twins discordant for type 2 diabetes, the diabetic twin was much more frequently the lighter of the pair at birth (5). This implies that even when genetic factors are held constant, intrauterine factors influencing fetal growth may have long-term implications for metabolic health. The lesson for geneticists interested in this disease is that if they ignore the effects of the pre- and postnatal environment and their capacity to interact with genetic variants, real progress in understanding diabetes may be impeded.

What Have We Learned from Monogenic Disorders?

Over the past 20 years, several Mendelian disorders with diabetes as a major phenotypic feature have been characterized at the molecular level. The largest subgroup of these monogenic diseases is caused by defects in the pancreatic β cell, resulting in a stable or progressive disorder of insulin secretion (6). Monogenic disorders that primarily impair insulin action either involve molecules in the insulin signal transduction cascade or result in abnormalities of fat tissue development (lipodystrophy) with secondary metabolic derangements leading to insulin resistance (7).

---

1University of Cambridge, Department of Clinical Biochemistry, Addenbrooke’s Hospital, Cambridge CB2 0QQ, UK. 2The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambs CB10 1SA, UK. 3Medical Research Council Epidemiology Unit, Elsie Widdowson Laboratory, Fulbourn Road, Cambridge CB1 9NL, UK.

*To whom correspondence should be addressed. E-mail: so104@medschl.cam.ac.uk
If major mutations (i.e., mutations that cause a substantial functional defect and are normally rare in the population) in these molecules lead to a highly penetrant form of human diabetes, then it seems plausible that more subtle genetic changes affecting the structure or expression of these molecules might play a role in determining susceptibility to type 2 diabetes. Our current state of knowledge regarding genetic variants influencing type 2 diabetes strongly supports this notion (Table 1). Thus, a common missense variant in the Y2 isoform of peroxisome proliferator–activated receptor gamma (PPARγ) [Pro12→Ala12 (Pro12Ala)] has shown an association with diabetes in multiple studies, with a meta-analysis suggesting that the common allele is associated with an increased diabetes risk of 25% (8). Major missense mutations in this gene result in severe, dominantly inherited insulin resistance, diabetes mellitus, and additional features such as partial lipodystrophy and hypertension (9). A common variant (Glu23Lys) in the gene encoding the inwardly rectifying potassium channel Kir6.2 also seems to increase diabetes risk by about 25%. Again, major mutations in this gene lead to an inherited form of severe diabetes or hypoglycemia depending on whether the mutations cause aberrant opening or closing of the channel. On somewhat less secure ground, common variants in the gene encoding the transcription factor hepatocyte nuclear factor 4α (HNF4α), the insulin receptor, and the mitochondrial genome appear to influence type 2 diabetes risk, whereas major mutations in these genes lead to more overt and early-onset metabolic disturbances. Thus, if past performance is any guide to future success, then any monogenic disorder of insulin secretion, termed maturity-onset diabetes of the young (MODY), is notable that in no case has the homologous murine knockout exhibited hyperglycemia (6). Even with something as fundamental as the insulin receptor, there are significant differences between the human and mouse phenotypes. Thus, mice lacking both copies of the insulin receptor are born of normal weight but die rapidly after birth of ketoacidosis (17), whereas humans with analogous null mutations are small at birth and rarely if ever develop ketoacidosis (3). We have demonstrated (4) that humans with the PPARγ Pro467Leu allele develop extreme insulin resistance, diabetes mellitus, and hypertension. Surprisingly, mice with the identical mutation, although they are hypertensive, show completely normal insulin sensitivity and glucose homeostasis (17). Thus, when it comes to the control of intermediary metabolism and plasma glucose levels, there may sometimes be important differences between mice and humans. The extent to which mice accurately model human phenotypes may be highly dependent on the particular phenotype under investigation.

### Why Do We Want to Discover the Precise Genetic Basis of Type 2 Diabetes Risk?

Given the massive investment of time, money, and energy in the search for genetic determinants for type 2 diabetes, it is reasonable to ask “Why did we bother starting in the first place?” Intellectual curiosity is a respectable motivation, but it would be difficult to justify the required resources on that basis alone. The principal motivation for this venture lies in the potential for improvement of human health. We could take the attitude that we already know about many of the environmental and lifestyle risk factors for type 2 diabetes and that intensive programs targeted at altering those factors can substantially reduce diabetes incidence in at-risk individuals (2). The rapid and widespread implementation of life-

<table>
<thead>
<tr>
<th>Gene</th>
<th>Monogenic disease</th>
<th>OMIM</th>
<th>Polygenic type 2 diabetes</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPARγ</td>
<td>Familial partial lipodystrophy (FPLD3)</td>
<td>604367</td>
<td>Pro12Ala, OR = 1.25</td>
<td>(8)</td>
</tr>
<tr>
<td>KCNJ11</td>
<td>Permanent neonatal diabetes mellitus (PNDM)</td>
<td>606176</td>
<td>Glu23Lys OR = 1.18</td>
<td>(18–20)</td>
</tr>
<tr>
<td>HNF4A</td>
<td>MODY1</td>
<td>125850</td>
<td>Thr103Ile late-onset diabetes in Japanese (OR = 4.3), 5’ SNPs increased risk in Finnish (OR = 1.33) and Ashkenazim (OR = 1.4), protective haplotype in UK Caucasians (OR = 0.83)</td>
<td>(20–23)</td>
</tr>
<tr>
<td>Mitochondrial genome</td>
<td>Diabetes and deafness maternally inherited (MIDD)</td>
<td>520000</td>
<td>Mitochondrial DNA 16189, OR = 1.6</td>
<td>(24)</td>
</tr>
<tr>
<td>HNF1A</td>
<td>MODY3</td>
<td>600496</td>
<td>Excess paternal transmission of class III VNTR (69% versus expected 50%), 3p+9 in UK Caucasian (OR = 2.02 recessive model only)</td>
<td>(25)</td>
</tr>
<tr>
<td>INS</td>
<td>Diabetes-type hyperglycemia with hyperinsulinemia</td>
<td>176730</td>
<td>Val985Met in the Netherlands (OR = 1.87), IVS6–2 in UK Caucasians (OR = 1.32) and haplotype in UK Caucasians (OR = 1.34)</td>
<td>(20, 27)</td>
</tr>
</tbody>
</table>

How Helpful Have Animal Models Been?

The power of modern murine genetics means that naturally occurring or induced mutations resulting in a relevant phenotype can be positionally identified and, in addition, targeted mutagenesis can reveal the potential contribution of any gene to glucose homeostasis. The importance of murine genetics has been emphatically shown in the field of energy balance and obesity, a topic of great importance to type 2 diabetes given that most type 2 diabetic humans are obese. The discovery of the leptin and melanocortin pathways controlling energy balance have largely been driven by murine genetics (10). The relevance of these pathways to the control of energy balance in humans has been repeatedly demonstrated, although there are some notable exceptions where humans and mice diverge (11).
style interventions to reduce diabetes is not, however, incompatible with the desire to improve the effectiveness and precision of treatment and prevention through greater understanding of the genetic and molecular etiology of the disease.

How could genetic information help us? The great hope that inspired most efforts in disease gene discovery is that such genes will then become therapeutic targets and that drugs targeted to molecules that are fundamentally involved in disease causation will be more effective than the cruder therapies we now use. An alternative strategy is to use murine genetics to identify molecules critical for the normal control of metabolism and glucose homeostasis and to infer from that how pharmacological manipulation of such molecules would offer therapeutic benefit. In this scenario, genetics is being used as a tool to aid the discovery of drugs. Perhaps more immediately applicable is the notion that the presence of particular genetic variants in an individual will influence his/her response to particular therapies (“pharmacogenomics”), as in the case of patients with MODY resulting from HNF1α mutations who preferentially respond to sulphonylurea therapy (6). In this case, we are looking at genetics as a guide to better drug use in the clinic. Greater understanding of the basis of disease also allows for more accurate provision of prognostic information to patients as the risk of future complications and likely disease trajectory is more accurately determined. Individuals with glucokinase mutations, for example, have stable hyperglycemia with limited risk of the microvascular complications, whereas those with mutations in the HNFs have a more progressive course (6).

Perhaps the most exciting future for genetics in type 2 diabetes is not necessarily in pharmacogenomics or provision of prognostic information but in understanding how specific genes interact with diet, exercise, and other lifestyle factors in the control of intermediary metabolism. We look forward to the day when the genetic technology will aid the identification of individuals at high risk and will also help to determine which particular combination and type of diet and exercise program and perhaps pharmacotherapy can be optimally used to prevent the onset of hyperglycemia. In the short term, because we know that family history and obesity interact to increase risk, it might be sensible to target preventive measures at sedentary people with a family history of diabetes (14).

How Might We Progress Our Understanding of the Genetic Basis of Diabetes? It is likely that future progress in the identification of true diabetes susceptibility genes will come from a variety of directions and that it will require synthesis of different types of research evidence. The clearest example of success from family-based linkage studies and positional cloning in complex disease has been the identification of CARD15 variants in inflammatory bowel disease, in which a relatively uncommon variant of large effect was detected by fine mapping of a region underlying a linkage peak (15). However, this may not be the case with other variants and other diseases if the effects of the variants are less pronounced and the linkage peaks less distinct. Thus, to detect the effect of the PPARG2 Pro12Ala polymorphism, which is only associated with a modest risk of diabetes, the use of linkage studies would have required an impractically large number of families. The only type 2 diabetes susceptibility gene thus far identified with a positional approach is that encoding the protease Calpain 10 (16). Although data supporting the proposition that this is a true susceptibility gene are accumulating, it may be too early to conclude that its contribution has been established beyond reasonable doubt. Recognizing the importance of study size, an important characteristic of current positional cloning-based efforts is the establishment of global collaborative consortia.

The continuing study of murine models with both gene targeting and random mutagenesis strategies is very likely to continue to aid the identification of molecules important for the control of glucose homeostasis and thereby provide previously unrecognized therapeutic targets. Extreme human phenotypes will remain a powerful route to finding genes that may be relevant for common human phenotypes. Thus far, we have implied that if major missense or nonsense mutations in a particular gene cause simple Mendelian forms of diabetes, then common SNPs in or around that gene—usually thought to affect gene expression—might alter the risk of the more common form of the disease. The effect on risk would be modest in any one individual but widespread in the population. Recent work by Hobbs and colleagues (17) examining genetic determinants of plasma high-density lipoprotein (HDL) concentrations raises the possibility of an alternative model. These authors sequenced genes in which homozygous mutations were known to cause rare dyslipidemic syndromes in subjects from the normal population who were either below the 5th or above the 95th percentile for plasma HDL cholesterol. Remarkably, there was a high frequency of rare missense variants in these genes among subjects in the lowest fifth percentile for plasma HDL, suggesting that major phenotypic effects of multiple different rare alleles contribute substantially to low HDL cholesterol in the general population. If the same is true for the physiological traits that underlie type 2 diabetes (e.g., insulin sensitivity), then strategies relying on common SNPs alone will not be sufficient to find all the important genetic contributors to type 2 diabetes risk.

Case-control studies, combined with confirmation of transmission disequilibrium in family-based studies, are likely to provide the mainstay of future association studies of possible “diabetogenes.” However, much more attention will need to be paid to study size and the careful characterization of cases and controls. Examining studies for heterogeneity and in particular biased estimation of the magnitude of effect by study size will be important because of potential publication bias.

Is It Possible to Determine the Genetic Basis of Diabetes Without Considering Key Life-style Factors? Traditionally genetic investigators have tackled gene-environment interaction as a second-order question, to be examined once susceptibility alleles have been identified. However, in a phenotype such as type 2 diabetes, for which life-style factors are likely to have a major impact, the neglect of environmental exposures may mean that important effects of particular alleles are not only attenuated but even obscured if the effect of genotype differs markedly according to life-style. Our epidemiological methods for measuring dietary factors and physical activity are imprecise, and measurement precision is critical to the ability to detect interaction. The quantitation of environmental exposures is therefore a major challenge but one that should not be ignored. Simple cross-sectional case-control studies of gene-life-style interaction will not work because of biased estimation of life-style in people who have been given the diagnostic label of diabetes. The optimal approach will involve case-control studies nested within major epidemiological cohort studies. These need to be large, have careful measurement of life-style factors at baseline and critically amass sufficient person-years of follow-up before they are of use. Cohorts of this nature are planned but will need many years before sufficient incident cases occur. In the meantime, we can make use of existing opportunities in multicohort studies that were established in the past.

Conclusions In the hunt for genetic factors that influence susceptibility to type 2 diabetes, no single approach is likely to be successful, and the optimal strategy will involve a variety of different study designs and analytical approaches, including the use of model organisms. Once we are properly positioned to examine the relationship between genetic polymorphisms and responses to different therapies and preventive strategies, the era of personalized, genome-based medicine will have begun in
How Obesity Causes Diabetes: Not a Tall Tale

Mitchell A. Lazar

The epidemic of obesity-associated diabetes is a major crisis in modern societies, in which food is plentiful and exercise is optional. The biological basis of this problem has been explored from evolutionary and mechanistic perspectives. Evolutionary theories, focusing on the potential survival advantages of “thrifty” genes that are now maladaptive, are of great interest but are inherently speculative and difficult to prove. Mechanistic studies have revealed numerous fat-derived molecules and a link to inflammation that, together, are hypothesized to underlie the obesity-diabetes connection and thereby represent prospective targets for therapeutic intervention.

Type 2 diabetes stems from the failure of the body to respond normally to insulin, called “insulin resistance,” coupled with the inability to produce enough insulin to overcome this resistant state. This common form of diabetes is often associated with obesity, and the current epidemics of these two conditions are seemingly related (1). This is glaringly evident in children, who are increasingly plagued by obesity and in whom the prevalence of type 2 diabetes (formerly termed “adult onset”) is approaching that of type 1 diabetes (formerly termed “juvenile onset”) (2). The epidemic of diabetes has a huge associated cost in terms of healthcare dollars as well as human morbidity and mortality (3). Recent studies predict that one in three Americans born in the year 2000 will develop diabetes in their lifetime (4), and a similarly ominous future confronts nearly all developed nations. Here, I discuss the relationship between obesity and diabetes, first in terms of the evolutionary forces that might explain their increased incidence in the modern world and then in terms of the pathogenic pathways that link the two conditions and inform rational strategies for prevention and therapy.

Why We Have Epidemics of Obesity and Diabetes: An Evolutionary Perspective

The evolutionary perspective has successfully guided much of modern biology, yet it is not always definitive. Take, for example, the giraffe’s long neck, which would seem to provide a competitive advantage for obtaining food, thus favoring survival and reproduction of the species. However, in his essay “The Tallest Tale,” Gould argued that the weight of scientific evidence favors alternative selective pressures as having led to the giraffe’s long neck, including combat advantages, sight of predators, and efficient heat loss (5).

There are no known survival advantages of morbid obesity, and increased body fat is associated with increased mortality (6). Hence, natural selection is unlikely to have favored obesity per se. On the other hand, during periods of prolonged famine that plagued early human hunter-gatherers, a survival advantage would have been conferred by genes that favor the economical use and storage of energy: so-called “thrifty” genes (7). The existence of thrifty genes was initially proposed by Neel, who focused on the efficient use of glucose as a biological fuel; he suggested that evolutionary pressure to preserve glucose for use by the brain during starvation led to a genetic propensity toward insulin resistance in peripheral tissues (8). Biological systems store energy most efficiently as fat and, hence, another function of thrifty genes is to promote an increase in adipose tissue. In the modern setting of sedentary lifestyles and unrestricted access to high-caloric foods, thrifty genes have been suggested to underlie the twin epidemics of obesity and diabetes (7).

Human obesity has a clear genetic component but is rarely monogenic (9). Thus, there are likely to be multiple thrifty genes, and the inheritance of several polymorphisms leading to small differences in expression can make populations more or less susceptible to obesity and diabetes (10). Several candidate thrifty genes have been proposed and are reviewed elsewhere (11). In principle, there could be separate sets of thrifty genes that promote body fat deposition or insulin resistance. Indeed, this concept is supported by a paradox: Insulin actually increases the production and storage of fatty acids in adipose tissue, thereby exacerbating obesity, whereas tissues such as muscle are insensitive to insulin (12). Nevertheless, Occam’s Razor (the principle that plurality of causes should not be postulated unless absolutely necessary) argues for thrifty genes that both increase energy storage and cause insulin resistance.

Perhaps the best thrifty gene candidate is the gene that encodes leptin, a hormone produced by adipose tissue and the absence of which leads to obesity and insulin resistance in rodents and humans (13). Leptin functions physiologically as a signal of energy stores, inhibiting food intake and accelerating energy...