Exposing the exposures responsible for type 2 diabetes and obesity

Paul W. Franks and Mark I. McCarthy

The rising prevalences of type 2 diabetes and obesity constitute major threats to human health globally. Powerful social and economic factors influence the distribution of these diseases among and within populations. These factors act on a substrate of individual predisposition derived from the composite effects of inherited DNA variation and a range of environmental exposures experienced throughout the life course. Although “Western” lifestyle represents a convenient catch-all culprit for such exposures, effective treatment and prevention will be informed by characterization of the most critical, causal environmental factors. In this Review, we examine how burgeoning understanding of the genetic basis of type 2 diabetes and obesity can highlight nongenetic exposures that drive development of these conditions.

About 10% of the global population already has type 2 diabetes (T2D) or is likely to develop it, and ~40% of adults are overweight or obese. Current strategies for prevention are limited in scope and effectiveness, and the persistently high prevalences of both conditions speak to the inadequacies of available therapeutic options.

Individual predisposition to these conditions has a strong genetic basis. Consensus estimates of heritability for obesity and T2D are ~70 and 35%, respectively (1, 2), and scores of genetic variants are now known to influence risk (3, 4). T2D and obesity are, however, also “diseases of lifestyle.” Rates of both have risen sharply over recent decades in tandem with widespread social changes, and these observations are supported by randomized lifestyle intervention trials that demonstrate convincing reductions in body weight and delayed progression to T2D in high-risk adults (5). The environmental exposures driving the development of these conditions must be both impactful, given the rapid shifts in disease prevalence that they have engendered, and pervasive, given that no contemporary industrialized population has been spared.

Epidemiological studies have highlighted many potential environmental “perpetrators” (Fig. 1), the combination of physical inactivity and calorie excess being the most prominent. There are, however, many other plausible environmental factors for which a role has been advanced, including sleep deprivation, endocrine disruptors, and smoking (6). The core limitations of observational studies—confounding, bias, and reverse causality—hinder efforts to determine which among these highly correlated exposures is truly causal (7). Yet clearer definition of these critical exposures is a prerequisite if more effective, targeted interventions are to be implemented at both the personal and the population levels.

The “nature versus nurture” framework for describing the contributions of genetic and environmental influences has been replaced by a more nuanced view that recognizes that the mechanisms through which environmental and genetic variation modify risk may be shared (Fig. 2). Environmental exposures that disturb cellular and physiological processes and influence individual predisposition to diseases such as T2D are likely to do so through active, or reactive, modulation of genome function (through changes in DNA methylation and transcription, for example).

Genetics of T2D and obesity

T2D is the consequence of reduced insulin secretion from pancreatic ß cells, typically observed in the context of insensitivity to the peripheral actions of insulin. Such insulin insensitivity is usually compounded by excess lipid deposition, particularly in nonstandard sites such as the liver and muscle. Physiological and genetic data from humans and rodents support a model whereby multiple concurrent molecular, cellular, and physiological processes contribute to the development of disease (Fig. 3).

Rare variants of large effect are causal for extreme phenotypes such as neonatal diabetes and severe early-onset obesity, but these contribute little to the population burden of T2D and obesity. Genome-wide association studies (GWAS) have identified scores of loci containing common variants robustly associated with T2D and obesity (3, 4), and elucidation of the mechanisms through which these operate provides novel pathophysiological insights. With notable exceptions (8), these common variant signals are of modest effect, collectively explaining only a minority of the overall genetic risk (~20% for T2D and ~5% for body-mass index (BMI)) (3, 4). Much of the remainder can be attributed to a large number of common variant signals with individual effects that are undetectable at stringent levels of statistical significance; for BMI, these underlie ~40% of overall variance (9). Sequence-based analyses are extending discoveries to variants of lower frequency, but the contribution that these make to population variation in the risk of T2D and obesity appears to be limited (10, 11).

This Review focuses on the application of this improved understanding of genetically driven variation in disease risk to provide mechanistic insights into the causal impact of proposed environmental exposures and thereby to define more effective interventions. These applications may, for example, take genetic variants that mimic environmental exposures and use the principles of Mendelian randomization to determine whether those exposures are likely to be causal for disease (7). Alternatively, they may aim to detect gene-environment interaction effects, whereby the impact of a given genetic variant is modified by the environmental milieu (or the reverse). Data from rodent and cellular models can also provide clues as to mechanism and causation, but their value is crucially dependent on the extent to which these

---

*Corresponding author. Email: mark.mccarthy@dirl.ox.ac.uk

1Department of Clinical Sciences, Genetic and Molecular Epidemiology Unit, Lund University, Malmö, Sweden. 2Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden. 3Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA. 4Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, UK. 5Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK.

---

**Type 2 diabetes and obesity**

**Fig. 1.** Examples of environmental exposures and mechanisms implicated in the development of T2D and obesity. Inclusion in this figure does not indicate that a causal connection has been demonstrated.
models recapitulate exposures and processes that are relevant to humans.

**Genetics and diet**

Obesity is a major risk factor for the development of T2D. Most people with T2D are overweight or obese at the point of diagnosis, and interventions that reduce body weight lower diabetes risk (5). Those who develop T2D despite having a normal body weight tend to have more prominent defects in insulin secretion, which translate into a more rapid requirement for exogenous insulin treatment. In some individuals, this reflects patterns of genetic predisposition that have features of both T1D and T2D and concomitant loss of β-cell capacity caused by autoimmune insult; in others, T2D with a lean body composition simply reflects one end of the spectrum of T2D presentation.

The consensus is that increased energy intake, facilitated by widespread availability of energy-dense foods, has contributed, in concert with lower energy expenditure (e.g., reduced physical activity), to trends of positive energy balance (12). However, human diets are complex, and there are many specific dietary components that have, at various times, been implicated in T2D risk (Fig. 1). The details have been debated, most recently with respect to the relative dangers of diets rich in processed carbohydrates (especially sugars) and fats. Nevertheless, there is no compelling evidence from epidemiological or clinical trial data that any given dietary configuration is more effective at reducing long-term body weight (13), and it is becoming clear that although some types of dietary fat may be metabolically harmful, others may in fact be protective (14).

What insights can genetic data provide? Variants discovered by GWAS to influence overall adiposity are enriched for a role in hypothalamic control of energy balance, with overrepresentation of pathways involved in both food intake and physical activity (3). Common variation at the FTO locus (which accounts for ~1% of population variance in BMI) affects energy balance (15), and BMI-raising FTO alleles correlate with higher dietary protein intake in adults, but not children (15), and with higher total energy intake in children (15) and adults (16). More recently, BMI-associated alleles at this locus have been linked to increased expression of IRX3 and IRX5 during early adipocyte differentiation and a reduction in the potential to dissipate energy through adipocyte browning, raising the possibility that differences in food choice and energy intake associated with some BMI-risk variants are the consequence, rather than the direct cause, of primary alterations in adipose mass (17). Although taste, macronutrient preference, and food patterns are under some degree of genetic control (18), these variants have no evident impact on the risk of T2D or obesity. This highlights the complex ways in which adiposity loci such as FTO may act and illustrates the need for careful partitioning of causal from noncausal relationships.

Most of the ~100 loci known to influence T2D risk (4) do so through primary effects on insulin secretion, pointing to underlying defects in pancreatic islet development and/or function; only a minority act through reducing insulin action. Among these T2D loci, the most obvious mechanistic connection to diet involves coding variation within PPARG. This gene encodes a nuclear receptor implicated in insulin signaling, adipogenesis, and the matching of lipid storage provision to nutritional state. Modest interactions...
between PPARG variants and dietary fat type (mainly polyunsaturated fatty acids (PUFAs)) with respect to T2D risk have been reported, but these remain unconfirmed (29), and there is no evidence of positive clinical outcomes arising from individualized approaches to prevention or management of T2D predicated on PPARG variation.

Inherent challenges associated with the accurate assessment of nutritional intake complicate efforts to define the contribution of diet to the development of T2D and obesity. Genetic data can help to address some of these challenges, particularly with respect to the effects of micronutrients. Vitamin D (25(OH)D) deficiency, for example, has long been touted as a cause of T2D on the basis of abundant observational evidence (20) and experiments showing positive effects of vitamin D supplementation on insulin secretion (21). Genetic variants that influence vitamin D metabolism can be used to define population subgroups that will experience lifelong differences in 25(OH)D exposure. Because allocation to the high- and low-exposure groups reflects the chance segregation of alleles at fertilization (hence the term Mendelian randomization), such groups should be, subject to some critical assumptions, matched for environmental and other factors that might otherwise confound interpretation (7). Comparisons between such genotype-defined groups indicate that although BMI has a causal impact on 25(OH)D levels (22), there is little or no causal relationship between variation in 25(OH)D levels and T2D (23). This is consistent with recent randomized controlled trial data that indicate no clinically relevant effects of supplemental vitamin D on glycemic indices in people with or without T2D (24).

Increasing numbers of similar genetic “instruments” are being identified that serve as proxies for environmental exposures relevant to obesity and T2D. For example, failure to detect overlap between the sets of genetic variants influencing T2D and obesity and those influencing regulatory inflammatory and immune function argues strongly that the chronic inflammation characteristic of T2D and obesity is a reaction to, rather than a cause of, these conditions. More recently, identification of variants that influence sleep behaviors has enabled dissection of causal relationships between sleep disturbance and metabolic disease (25).

A further opportunity for genetic insight is afforded by populations with distinct patterns of environmental exposure. Greenlandic Inuits, for example, have needed to survive in a cold climate on a marine diet rich in omega-3 PUFAs. This has driven genetic adaptation, with selection for variation at loci that influence fatty acid metabolism and brown fat differentiation (26). Some of these historically advantageous adaptations now seem to promote obesity and T2D (8). Homozygous carriers of the nonsense p.Arg684Ter allele in the TBCID4 gene, which is common among the Inuit but rare elsewhere, are at a severalfold increased risk of T2D. The underlying mechanism for this risk appears to involve muscle-selective loss of the long isoform of TBCID4, leading to reduced insulin-stimulated GLUT4-mediated glucose uptake into muscle and marked postprandial (but not fasting) hyperglycemia.

In overfeeding studies in twins (27), phenotypic responses to dietary interventions demonstrate strong familial clustering in weight change; this may reflect the modifying effects of genetic variants on the response to dietary manipulation (that is, gene-environment interaction). Here we restrict the use of the term gene-environment interaction to situations of evident nonadditivity (that is, where the joint effects of a pair of specified genetic and nongenetic exposures are significantly greater or less than the sum of their individual effects).

Identification of robust (independently replicated) gene-environment interaction effects could provide the basis for personalization of disease prevention and management. However, the detection of gene-environment interactions in humans is prone to multiple sources of bias and confounding (28), and power is constrained by imprecision in the measurement of exposures and outcomes (29).

Nevertheless, there is some evidence that “healthy” diets modify the impact of individual BMI-associated variants in observational studies (30) and clinical trials (31). Interactions have been reported between BMI-associated genetic risk scores and diverse exposures including sugar-sweetened beverages (32), fried foods (33), and television viewing (34), though replication data are sparse. The most comprehensive epidemiological study of gene-diet interactions in T2D, a prospective study involving ~4 million person-years of follow-up in 340,000 participants (35), provided no evidence that a Mediterranean diet influenced the individual or collective effects of known T2D variants.

Clinical trials are often thought to overcome the limitations of epidemiology that might lead to confounded results; but few trials account adequately for the effects that adherence and/or compensatory behaviors might have on metabolic traits. Lifestyle interventions typically occupy <5% of waking hours, and while participants behave during the rest of the day—the food they eat, the physical activities they pursue, and the quality of their sleep—is likely to contribute to heterogeneous responses (36). These limitations are hard to overcome, because lifestyle interventions, unlike drug interventions, cannot be easily masked, and the ubiquitous monitoring of behavior remains challenging. Nevertheless, the most comprehensive trial-based assessment of gene-lifestyle interactions in T2D incidence, which involved 2843 adults from the Diabetes Prevention Program, found no interaction between genetic measures of T2D risk and intervention with either metformin or lifestyle changes (37). Overall, on the evidence as it stands now, there is no compelling basis for using gene-diet interaction data to support clinically useful individualization of management for these conditions.

Genetics and energy expenditure

The processes that contribute to overall energy expenditure (including those related to basal metabolism, exercise, nonexercise activity thermogenesis, and food-related thermogenesis) are obvious candidates with respect to obesity risk. There is, however, little evidence to indicate that the T2D- and obesity-risk variants identified by GWAS directly influence these processes, and many of the candidate genes implicated by earlier studies (e.g., those encoding the uncoupling proteins) have not been substantiated in the much larger studies. Although intervention studies have demonstrated that phenotypic responses to exercise are familial (38), there has been little success...
in identifying specific variants that, at the population level, influence exercise tolerance or modulate how exercise influences weight gain or metabolism.

Interactions between BMI-associated variation and measures of physical activity that influence adult adiposity appear to be more robust than those involving dietary exposures or diabetes outcomes. The BMI effect associated with FTO variation has consistently been shown to be weaker in physically active than in inactive carriers (39), and there have been similar interactions found for sets of obesity-associated variants (40). However, such studies are challenging to perform and interpret (41) and need further replication. Despite promising epidemiological data, the largest clinical trial analysis found no evidence that FTO variation influences weight loss after lifestyle intervention (31).

Our assessment is that there is only meager evidence to date that common genetic variation modifies the effects of lifestyle exposures with respect to the development or management of obesity or T2D. This may be because the interaction effects are nonexistent or of small magnitude, or because our research methods and available data sets are insufficient to characterize the complexity of the interactions (Fig. 2).

**Genetics and the microbiome**

There has been an explosion of interest in the role of the gut microbiome in the development of T2D and obesity. Variation in the diversity and composition of gut microflora, in part reflecting personal histories of antibiotic exposure and dietary intake, has been tied to individual risk, as well as to the sharp rise in the prevalences of these conditions (42, 43). In addition, the metabolic benefits of metformin and bariatric surgery have been ascribed to their impact on the microbiome (44, 45). In rodents, manipulation of the microbiome (e.g., through fecal transplantation) can lead to weight loss and diabetes remission (42), though evidence that similar interventions are effective in humans remains limited (46). However, an algorithm that integrates personal clinical (biochemistry and anthropometry), behavioral (dietary preferences and physical activity), and microbiota data has been shown to predict an individual’s metabolic response to food intake and to provide dietary recommendations that limit glycemic excursions after meals (47).

Several studies have detected marked shifts in microbiome content among those who are obese or diabetic, though the data are inconsistent (48, 49). A variety of mechanisms for the metabolic effects of microbiome diversity have been proposed, including impacts on short-chain fatty acid production, bile acid metabolism, and inflammation. However, these studies tend not to distinguish between microbiome variation that is causal for T2D and/or obesity and that which is a function of the disease or its treatment or merely a consequence of correlated exposures. The range of environmental factors influencing gut microbiota is considerable (50), and, in the case of T2D, early reports of disease-associated variation in microbiome content proved to be confounded by metformin treatment, which has a marked impact on microbiome integrity (44). Characterization of the impact of host genome variation on microbiome diversity and content (51) will provide genetic instruments that will support efforts to define, much more precisely than has hitherto been possible, the extent to which genetic variants that influence individual risk of T2D and obesity do so through direct, or indirect, impacts on the gut microbiome.

**Genetics and early life environment**

Genetic and environmental exposures offer sharply contrasting explanations for the widely replicated associations between low birth weight (and early growth) and increased propensity to develop obesity, T2D, and cardiovascular disease in later life (52). The dominant explanation has been provided by the developmental origins (or “fetal programming”) hypothesis, which attributes this relationship to the long-term effects of restricted intrauterine nutrient availability (reflecting maternal nutritional and placental function) on the risk of metabolic disease decades later. This hypothesis, which is consistent with observational studies in humans exposed to severe nutritional restriction during early life, is also supported by experimental studies in rodents. These studies have focused attention on the detection of methylation signatures that might convey the “memory” of early life events across the life course (52). However, most of the T2D- and obesity-associated methylation signals detected in blood-based epigenomic studies have either failed to be replicable or appear to be reactive or confounded, not causal (53). One exception may involve TXNIP, which encodes a thioredoxin-reducing protein implicated in diverse metabolic processes including nutrient sensing, islet function, and energy expenditure (54): Methylation in this region has been associated with both prevalent and incident T2D (55), though, as yet, not with early growth restriction.

In populations in which maternal obesity and gestational diabetes are frequent, the relationship...
between early growth and adult T2D is best described as U-shaped (56), in that both high and low birth weights are linked to T2D in later life. The elevated T2D risk in those with high birth weight likely reflects the impact of maternal hyperglycemia. Excessive placental transfer of glucose from hyperglycemic mothers not only promotes fetal growth (insulin is a major trophic factor in early life) but also drives a direct, non-genetic increase in offspring propensity to T2D (57), possibly because of the additional metabolic burden imposed on the developing endocrine pancreas.

Although the fetal programming hypothesis is alluring, the effects of shared genetic variants offer a complementary explanation for these observed relationships. Carriers of variant alleles that compromise insulin secretion or action and that therefore increase risk of T2D in later life, will also, given insulin’s trophic effects, tend to exhibit reduced fetal growth (reproducing the low-birth-weight arm of the U). Those same risk alleles may, when present in the mother, contribute to maternal hyperglycemia, providing a potential mechanistic explanation for the high-birth-weight arm of the U (58). Such genetically mediated links between early growth and subsequent metabolic dysfunction have been well documented in families segregating rare monogenic forms of diabetes, such as glucokinase MODY (maturity onset diabetes of the young) (58).

Common alleles implicated in T2D disproportionately influence variation in birth weight, though the directional relationships are complex (59). Children carrying the T2D-risk allele at some loci, such as MTNR1B and GCK, have higher birth weights, reflecting a predominant effect of those variants on maternal hyperglycemia. At other loci, such as ADCA5 and CDKAL1, the T2D-risk allele lowers birth weight, a pattern consistent with direct fetal growth restriction. These explanations fit with the epidemiological data: In 236,000 UK Biobank participants, a paternal history of T2D was associated with reduced birth weight and a maternal history with elevated birth weight (60).

The model that emerges is one in which the relationship between early growth and later disease is influenced by an innate wean of genetic and environmental mechanisms connecting both extremes of early growth and birth weight to subsequent T2D (Fig. 4). The direct effects of fetal genotype on both early growth and later T2D are modulated by the countervailing effects of the same genotypes in the mother (acting at least in part through the fetal environment) and by other nongenetic influences that affect fetal nutrition. This relationship may turn out to be even more mechanistically complex, subject to the contribution of translational epigenetic influences (61) and/or environmentally triggered polyphenism (62).

Where next?

One-size-fits-all diet and exercise recommendations for weight loss and diabetes prevention elicit uneven and unpredictable responses, and the development of effective personalized approaches is a highly desirable goal that genetics might help us to achieve. However, at present, we lack definitive insight into the specific components of modern lifestyles that are most responsible for T2D and obesity risk, as well as the genetic variants that reliably predict individual metabolic or adiposity responses to common exposures, and that could justifiably be used to personalize lifestyle interventions.

Improved specification of the genetic basis of disease predisposition (through whole-genome sequencing and more targeted temporal assessments of exposures and outcomes [combining near-continuous objective measurements of movement, sleep, and diet collected with biomarker technologies]), conducted in ever larger biobanks and health care settings, will allow for the detection of mechanistic overlap and interaction and greater clarity regarding key causal exposures. An important opportunity exists in tying these data to robust, accessible, molecular signatures of individual disease trajectory, which are able to capture the summed, actual (rather than predicted) impact of genetic and environmental influences on a given individual at a given point in time. As well as quantifying risk, these signatures may provide a more nuanced classification of disease etiology and support more precise, personalized diagnostic assignment.

The discovery of such signatures will require investment in the analysis of longitudinal repeated-measures data from large population samples and trials. By enabling more accurate specification of individual disease risk and molecular pathology, such signatures have the potential to support more effective targeting of preventative and therapeutic strategies. By combining these approaches—definition of the key causal exposures and individualization of prognostic and diagnostic information—we can hope to move closer to the desired goals of effective prevention and treatment of T2D and obesity.

REFERENCES AND NOTES

Exposing the exposures responsible for type 2 diabetes and obesity
Paul W. Franks and Mark I. McCarthy

Science 354 (6308), 69-73.
DOI: 10.1126/science.aaf5094