Exposing the exposures responsible for type 2 diabetes and obesity

Paul W. Franks[1,2,3] & Mark I. McCarthy[4,5]

[1]Department of Clinical Sciences, Genetic and Molecular Epidemiology Unit, Lund University, Malmö, Sweden;

[2]Department of Public Health & Clinical Medicine, Umeå University, Umeå, Sweden;

[3]Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA

[4]Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, UK

[5]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford UK

WORD COUNT: 4410 (including main text, legends, references)

ABSTRACT

The rising prevalences of type 2 diabetes and obesity represent major threats to human health globally. Powerful social and economic factors influence the distribution of these diseases between and within populations. These 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 lifecourse. Whilst "Western" lifestyle represents a convenient "catch-all" culprit for those 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 disease can highlight non-genetic exposures that drive development of these conditions.

Introduction

Approximately 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 prevalence of both conditions speaks 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 ecological 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, since no contemporary industrialized population has been spared.

Epidemiological studies have highlighted many potential environmental "perpetrators" (**Fig 1**), the combination of physical inactivity and caloric 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 behavior(*6*). The core limitations of observational

studies – confounding, bias, reverse causality – hinder efforts to determine which amongst these highly-correlated exposures is truly causal(*7*). Yet, clearer definition of these critical exposures is a pre-requisite if more effective, targeted, interventions are to be implemented at both the personal and the population level.

The "nature vs. 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 which 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 the pancreatic beta-cell, typically seen in the context of insensitivity to the peripheral actions of insulin. The latter is usually compounded by adipose tissue excess, and particularly by deposition of that lipid excess in non-standard sites such as 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 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 BMI/obesity)(*3, 4*). Much of the remainder can be attributed to a large number of common variant signals with individual effects undetectable at stringent levels of statistical significance: for BMI, these underlie ~40% of overall variance(*9*). Sequence-based analyses are extending discovery to variants of lower frequency, but the contribution these make to population variation in risk of T2D and obesity appears limited(*10, 11*).

This review focuses on the application of this improved understanding of genetically-driven variation in risk to provide mechanistic insights into the causal impact of proposed environmental exposures, and explore the potential for those insights to define more effective interventions. These applications may, for example, take genetic variants which 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 (GEI) 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 provide clues to mechanism and causation but their value is crucially dependent on the extent to which the models available recapitulate exposures and processes 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 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 concomitant loss of beta-cell capacity due to autoimmune insult, and patterns of genetic predisposition that have features of both type 1 and type 2 diabetes: in others, lean T2D 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 longterm body weight(*13*), and it is becoming clear that whilst some types of dietary fat may be metabolically harmful, others may in fact be protective(*14*).

What insights can genetic data provide? GWAS-discovered variants influencing overall adiposity are enriched for a role in hypothalamic control of energy balance, with over-representation 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 via adipocyte browning, raising the possibility that differences in food choice and energy intake which associate with some BMI-risk variants are the consequence, rather than the direct cause, of primary alterations in adipose mass(*17*). Certainly, whilst taste, macronutrient preference and food patterns are under some degree of genetic control(*18*), these variants have no evident impact on risk of T2D or obesity. This highlights the complex ways in which adiposity loci like *FTO* may act and illustrates the need for careful dissection of causal from non-causal relationships.

Most of the ~100 loci known to influence T2D risk(*4*) do so via primary effects on insulin secretion, pointing to underlying defects in pancreatic islet development and/or function: only a minority act via reducing insulin action. Amongst 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 [PUFA]) with respect to T2D-risk have been reported but these remain unconfirmed(*19*), 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 which influence vitamin D metabolism can be used to define subgroups that will have experienced lifelong differences in 25(OH)D exposure. Since 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 whilst 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 which 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 other 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 for regulatory inflammatory and immune function, argues strongly that the chronic inflammation characteristic of these conditions is a reaction to, rather than a cause of, these conditions. More recently, identification of variants which influence sleep behavior is enabling 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 accommodate to a cold climate and a marine diet rich in omega-3 PUFA. 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*). Homozygote carriers of the nonsense p.Arg684ter allele in the *TBC1D4* gene, common in the Inuit but rare elsewhere, are at several-fold increased risk of T2D. The underlying mechanism for this risk appears to involve muscle-selective loss of the long isoform of TBC1D4 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, GEI). Here, we restrict use of "gene environment interaction" to situations of evident non-additivity (that is, where the joint effects of a pair of specified genetic and non-genetic exposures is significantly greater or less than the sum of their individual effects). Identification of robust (i.e. independently replicated) GEI effects could provide the basis for personalization of disease prevention and management. However, the detection of GEIs 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 food consumption(*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 ~4M person-years of follow-up in 340K 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 trails 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 how participants behave during the rest of the day – the food they eat, the physical activities they pursue, the quality of their sleep – is likely to contribute to heterogeneous responses(*36*). These limitations are hard to overcome, as 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, in 2,843 adults from the Diabetes Prevention Program, found no interaction between genetic measures of T2D-risk and intervention with either metformin or lifestyle(*37*). Overall, on current evidence, 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, non-exercise activity thermogenesis and food-related thermogenesis) represent obvious candidates with respect to obesity risk. There is, however, little to indicate that the T2D- and obesity-risk variants identified by GWAS directly influence these processes, and many of the genes implicated by earlier candidate gene studies (e.g., those encoding the uncoupling proteins), have not been substantiated in the much larger studies. Whilst 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 impacts weight gain or metabolism.

Interactions between BMI-associated variation and measures of physical activity that influence adult adiposity appear 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

inactive carriers(*39*) and there have been similar interactions involving sets of obesity-associated variants(*40*). However, these 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 following 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 non-existent, of small magnitude, or because our research methods and available datasets 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 history of antibiotic exposure and dietary intake, has been tied to individual risk as well as the sharp rise in prevalence 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 which integrates personal clinical (biochemistry, anthropometry, physical activity), behavioral (dietary preferences) and microbiome 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 amongst those who are obese or diabetic, though data are inconsistent(*48, 49*). A variety of mechanisms for the metabolic effects of microbiome diversity has 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 which 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 significant impact on microbiome integrity(*44*). Characterization of the impact of host genome variation on microbiome diversity and content(*51*) will provide genetic instruments to support efforts to define, much more precisely than has been hitherto possible, the extent to which genetic variants which influence individual risk of T2D and obesity do so through direct, or indirect, impact on the gut microbiome.

Genetics and early life environment

Genetic and environmental exposures offer sharply contrasting explanations for the widely-replicated associations between low birthweight (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 nutrition and placental function) on the risk of metabolic disease decades later. This hypothesis, consistent with observational studies in humans exposed to severe nutritional restriction during early life, has been supported by experimental studies in rodents. These studies have focused attention on the detection of methylation signatures which might convey the "memory" of early-life events across the lifecourse(*52*). However, most of the T2D- and obesity-associated methylation signals detected in blood-based epigenome studies have either failed to replicate or appear reactive or confounded, but 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 where maternal obesity and gestational diabetes are frequent, the relationship between early growth and adult T2D is best described as "U"-shaped(*56*). The elevated T2D-risk in those with high birthweight 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 due to the additional metabolic burden imposed upon the developing endocrine pancreas.

Whilst 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 which therefore increase risk of T2D in later life, will also, given insulin's trophic effects, also tend to exhibit reduced fetal growth (reproducing the low birthweight 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 birthweight arm(*58*). Such genetically-mediated links between early growth and subsequent metabolic dysfunction are clearly documented in families segregating rare monogenic forms of diabetes, such as glucokinase-MODY(*58*).

Common alleles implicated in T2D disproportionately influence variation in birthweight, though the directional relationships are complex(*59*). At some loci, such as *MTNR1B* and *GCK*, children carrying the T2D-risk allele have increased birthweight, reflecting a predominant effect of those variants on maternal hyperglycemia. At others, such as *ADCY5* and *CDKAL1*, the T2D-risk allele lowers birthweight, 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, and a maternal history with elevated, birthweight(*60*).

The model that emerges is one where the relationship between early growth and later disease is influenced by an intimate weave of genetic and environmental mechanisms connecting both extremes of early growth and birthweight 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 non-genetic influences that affect fetal nutrition. This relationship may turn out to be even more mechanistically complex, subject to the contribution of transgenerational 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 clearly represents a highly desirable goal that genetics might help 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 genetic variants that reliably predict individual metabolic or adiposity response to common exposures, and which could justifiably be used to personalize lifestyle interventions.

Improved specification of the genetic basis of disease predisposition (through whole genome sequencing), and more detailed temporal assessments of exposures and outcomes (combining near-continuous objective measures of movement, sleep and diet 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, able to capture the summated, actual (rather than predicted) impact of genetic and environmental influences in 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.

FIGURE LEGENDS

Figure 1: Examples of environmental exposures and mechanisms implicated in development of T2D and obesity. Inclusion on this figure does not indicate that a causal connection has been demonstrated.

Figure 2: Environmental exposures act in a range of ways to perturb genome function. The joint effects of genetic and environmental factors can, for example, drive variations in ligand binding efficiency, membrane channeling, DNA replication and repair, or methylation. Variations in the "intrinsic environment", reflected in metabolite, proteome or microbiome profiles, for example, may also perturb genome function, thus generating complex feedback loops for gene-environment interactions.

Figure 3: Multiple processes contribute to the development of type 2 diabetes. Examples of T2D-associated loci for which the evidence points to specific mechanisms are indicated. At some of these loci, the specific effector gene has not yet been defined so the gene labels are purely indicative.

Figure 4: Mechanisms underlying the observed relationships between extremes of fetal growth and subsequent risk of T2D. Distinct genetic and non-genetic processes are implicated in both arms of the U-shaped curve that describes the relationship between birthweight and future T2D risk.

References:

1. G. Hemani *et al.*, *Am J Hum Genet* **93**, 865-875 (2013).
2. G. Willemsen *et al.*, *Twin Res Hum Genet* **18**, 762-771 (2015).
3. A. E. Locke *et al.*, *Nature* **518**, 197-206 (2015).
4. A. Mahajan *et al.*, *Nat Genet* **46**, 234-244 (2014).
5. W. C. Knowler *et al.*, *Lancet* **374**, 1677-1686 (2009).
6. S. W. Keith *et al.*, *Int J Obes (Lond)* **30**, 1585-1594 (2006).
7. G. D. Smith, N. Timpson, S. Ebrahim, *Ann Med* **40**, 524-541 (2008).
8. I. Moltke *et al.*, *Nature* **512**, 190-193 (2014).
9. J. Yang *et al.*, *Nat Genet* **47**, 1114-1120 (2015).
10. V. Agarwala, J. Flannick, S. Sunyaev, D. Altshuler, *Nat Genet* **45**, 1418-1427 (2013).
11. C. Fuchsberger *et al.*, *Nature* **536**, 41-47 (2016).
12. L. J. Orozco *et al.*, *Cochrane Database Syst Rev*, CD003054 (2008).
13. F. M. Sacks *et al.*, *N Engl J Med* **360**, 859-873 (2009).
14. F. Qian, A. A. Korat, V. Malik, F. B. Hu, *Diabetes Care* **39**, 1448-1457 (2016).
15. T. Tanaka *et al.*, *Am J Clin Nutr* **97**, 1395-1402 (2013).
16. Q. Qi *et al.*, *Hum Mol Genet* **23**, 6961-6972 (2014).
17. M. Claussnitzer *et al.*, *N Engl J Med* **373**, 895-907 (2015).
18. T. Pallister *et al.*, *Twin Res Hum Genet* **18**, 793-805 (2015).
19. L. Palla, J. P. Higgins, N. J. Wareham, S. J. Sharp, *Am J Epidemiol* **171**, 1225-1232 (2010).
20. L. Ceglia *et al.*, *Eur J Nutr*, (E pub Nov 2, 2015).
21. S. Inomata, S. Kadowaki, T. Yamatani, M. Fukase, T. Fujita, *Bone Miner* **1**, 187-192 (1986).
22. K. S. Vimaleswaran *et al.*, *PLoS Med* **10**, e1001383 (2013).
23. Z. Ye *et al.*, *Lancet Diabetes Endocrinol* **3**, 35-42 (2015).
24. Y. H. Krul-Poel *et al.*, *Diabetes Care* **38**, 1420-1426 (2015).
25. S. E. Jones *et al.*, *PLoS Genet* **12**, e1006125 (2016).
26. M. Fumagalli *et al.*, *Science* **349**, 1343-1347 (2015).
27. C. Bouchard *et al.*, *N Engl J Med* **322**, 1477-1482 (1990).
28. P. W. Franks, G. Pare, *Curr Diab Rep* **16**, 57 (2016).
29. M. Y. Wong, N. E. Day, J. A. Luan, N. J. Wareham, *Stat Med* **23**, 987-998 (2004).
30. J. A. Nettleton *et al.*, *Hum Mol Genet* **24**, 4728-4738 (2015).
31. G. D. Papandonatos *et al.*, *Diabetes* **64**, 4312-4321 (2015).
32. Q. Qi *et al.*, *N Engl J Med* **367**, 1387-1396 (2012).
33. Q. Qi *et al.*, *BMJ* **348**, g1610 (2014).
34. Q. Qi *et al.*, *Circulation* **126**, 1821-1827 (2012).
35. C. Langenberg *et al.*, *PLoS Med* **11**, e1001647 (2014).
36. M. I. Goran, E. T. Poehlman, *Am J Physiol* **263**, E950-957 (1992).
37. M. F. Hivert *et al.*, *Diabetes* **60**, 1340-1348 (2011).
38. C. Bouchard, T. Rankinen, *Med Sci Sports Exerc* **33**, S446-451; discussion S452-443 (2001).
39. T. O. Kilpelainen *et al.*, *PLoS Med* **8**, e1001116 (2011).
40. S. Ahmad, T. V. Varga, P. W. Franks, *Hum Hered* **75**, 106-115 (2013).
41. S. Ahmad *et al.*, *PLoS Genet* **9**, e1003607 (2013).
42. S. Ussar *et al.*, *Cell Metab* **22**, 516-530 (2015).
43. L. M. Cox *et al.*, *Cell* **158**, 705-721 (2014).
44. K. Forslund *et al.*, *Nature* **528**, 262-266 (2015).
45. A. Palleja *et al.*, *Genome Med* **8**, 67 (2016).
46. A. Vrieze *et al.*, *Gastroenterology* **143**, 913-916.e917 (2012).

47. D. Zeevi *et al.*, *Cell* **163**, 1079-1094 (2015).
48. F. H. Karlsson *et al.*, *Nature* **498**, 99-103 (2013).
49. J. Qin *et al.*, *Nature* **490**, 55-60 (2012).
50. G. Falony *et al.*, *Science* **352**, 560-564 (2016).
51. J. K. Goodrich *et al.*, *Cell* **159**, 789-799 (2014).
52. K. M. Godfrey, P. D. Gluckman, M. A. Hanson, *Trends Endocrinol Metab* **21**, 199-205 (2010).
53. K. J. Dick *et al.*, *Lancet* **383**, 1990-1998 (2014).
54. H. Parikh *et al.*, *PLoS Med* **4**, e158 (2007).
55. J. C. Chambers *et al.*, *Lancet Diabetes Endocrinol* **3**, 526-534 (2015).
56. O. Naess *et al.*, *Eur Heart J* **34**, 3427-3436 (2013).
57. D. J. Pettitt *et al.*, *Diabetes Care* **31**, 2126-2130 (2008).
58. A. T. Hattersley, J. E. Tooke, *Lancet* **353**, 1789-1792 (1999).
59. M. Horikoshi *et al.*, *Nat Genet* **45**, 76-82 (2013).
60. J. S. Tyrrell, H. Yaghootkar, R. M. Freathy, A. T. Hattersley, T. M. Frayling, *Int J Epidemiol* **42**, 1714-1723 (2013).
61. G. E. Blake, E. D. Watson, *Curr Opin Chem Biol* **33**, 101-107 (2016).
62. K. Dalgaard *et al.*, *Cell* **164**, 353-364 (2016).

ACKNOWLEDGEMENTS

MMcC is a Wellcome Trust Senior Investigator. He acknowledges funding from the Wellcome Trust (098381), European Commission (HEALTH-F4-2007-201413), Medical Research Council (MR/L020149/1; MR/M004422/1), and National Institutes of Health (DK098032; DK105535) for research related to the contents of this review. PWF acknowledges funding from the European Commission (CoG-2015_681742_NASCENT), Swedish Research Council (Distinguished Young Researchers Award in Medicine), Swedish Heart-Lung Foundation, and the Novo Nordisk Foundation for research related to the contents of this review.