Novel biological insights emerging from genetic studies of type 2 diabetes and related metabolic traits
N. Maneka G. De Silva and Timothy M. Frayling

Introduction
Genetic factors play an important role in influencing which individuals in a given population at a given time become diabetic. The increased prevalence of type 2 diabetes (T2D) is probably the most important adverse outcome of the increase in prevalence of obesity. As populations have become fatter, more individuals have become diabetic, but many obese people do not get T2D, whereas many nonobese people do get the condition. In addition, there is strong evidence that genetic factors influence where an individual lies on the population scale of body mass index (BMI). These observations have led to substantial efforts to identify the genes involved in T2D and related metabolic traits, in an effort to increase our biological understanding of a complex social and medical problem.

Despite concerted efforts over the past 15 years, it is only in the past 3 years that substantial progress has been made in identifying some of the genetic variation involved in metabolic traits. This progress has been possible largely due to technological advances. Microarray-chip-based assays can measure a large proportion of the common genetic variation in a single DNA sample. These assays, together with the establishment of large, well characterized patient and control DNA collections, have facilitated genome-wide association studies (GWASs). Largely through GWASs, we now know of approximately 20 gene variants associated with T2D, 10 with body mass index (BMI) and obesity, four with fasting glucose levels in the normoglycaemic population and over 30 with lipid levels. These findings are stimulating many new important areas of research related to metabolic diseases. For T2D and glucose metabolism, we discuss a number of aspects and implications of the genetic findings, including the observations that T2D gene variants are not usually in or near obvious candidate genes, highlighting the poor prior knowledge of the biology of the disease; most T2D gene variants are associated with β-cell function rather than insulin resistance; there is a difference between genes that influence variation in normal glucose levels compared with those influencing onset and progression of diabetes; and there is a genetic link between diabetes and foetal growth.

Summary
Genetic studies in the past 3 years have provided a greatly increased knowledge of the regions of the genome involved in adverse metabolic consequences. There are now over 100 common genetic variants reproducibly associated with metabolic traits, including reduced β-cell function, obesity, increased lipid levels and increased glucose levels. These genetic findings are already altering perceptions of how these traits develop and interact to result in diseases such as T2D.

Keywords
birth weight, fasting glucose, genetics, genome-wide association, type 2 diabetes
with BMI and obesity, four with fasting glucose levels in the normoglycaemic population and over 30 with lipid levels. It is important to note that further research is needed to understand the mechanisms behind how the DNA variants alter gene function. Despite the need for more research, there are a number of important insights already emerging from these recent genetic discoveries.

In this review, we discuss some of these insights, focusing on T2D. We discuss the poor correlation between genes implicated by GWASs and genes thought to be good candidates from previous knowledge of biology; the preponderance of genetic variants that influence β-cell function relative to those influencing insulin resistance; the difference between genes that influence variation in normal glucose levels compared with those influencing onset and progression of diabetes; and the genetic link between diabetes and foetal growth.

**Most gene variants influencing type 2 diabetes risk are not in or near obvious candidate genes**

Of the 20 common genetic variants associated with T2D, only four were identified by a candidate gene approach. The remaining 16 variants were identified by a genome-wide approach \([1^*,2–9,10^*–13^*]\) and few point to clear candidate genes. The results of these genetic studies, therefore, suggest our current knowledge of the cause of T2D is poor, but provide an opportunity to investigate multiple new pathways and mechanisms. The four candidate genes are \(TCF2/HNF1B, KCNJ11, WFS1\) and \(PPARG\) \([14–21]\). Rare mutations in all four of these genes cause a monogenic form of diabetes or severe insulin resistance such as maturity-onset diabetes of the young (MODY; mutations in \(HNF1B\)) \([22]\), permanent neonatal diabetes mellitus (PNDM; mutations in \(KCNJ11\)) \([23]\), Wolfram syndrome (mutations in \(WFS1\)) \([24]\) or severe insulin resistance (mutations in \(PPARG\)) \([25]\). These four genes, therefore, represent examples of genes that have a spectrum of rare and common variation that influences diabetes risk. Of the 16 variants identified by the genome-wide approach, it is possible to retrospectively implicate some genes and details of these are given in Table 1 \([1^*,2–9,10^*–13^*,14,15,17–19,26–34]\), but it is important to note that much further work is needed to be confident that the gene described is the causal gene.

**The majority of type 2 diabetes gene variants increase diabetes risk by reducing β-cell function and not increasing insulin resistance**

T2D is characterized by impaired insulin secretion and insulin resistance, and there has been considerable debate about the relative importance of each of these broad mechanisms. Multiple studies of the recently identified T2D risk variants suggest that impaired insulin secretion by the β-cell is the key factor that results in diabetes, rather than insulin resistance. For example, nondiabetic carriers of the \(TCF7L2\) risk allele have impaired insulin secretion but no evidence of a role in insulin resistance \([35–39]\). Variants in other T2D loci at or near \(CDKAL1, CDKN2A/B, HHEX-IDE, IGF2BP2, SLC30A8, KCNJ11, WFS1, JAZF1, TSPAN8, CD123, CAMK1D\) and \(MTNR1B\) are all associated with insulin secretion in nondiabetic populations \([8,40–46]\). In contrast, only the variants in \(IRS1\) are conclusively associated with insulin resistance \([13^*]\), with the exception of the variants in \(FTO\), which alter insulin resistance, along with glucose levels and a range of metabolic traits, but all through a primary effect on BMI \([47]\).

**Gene variants influencing fasting glucose in normoglycaemic range have variable effects on related metabolic traits, including type 2 diabetes**

Common gene variants associated with fasting glucose levels in the normoglycaemic population do not necessarily influence T2D risk \([28,30]\). In contrast, variants strongly associated with T2D risk do not necessarily influence normal variation in fasting glucose levels \([35,36]\). These observations suggest that different genes, and therefore different mechanisms, influence physiological glucose levels compared to β-cell deterioration (either number or functions) and pathophysiologically glucose levels. In reality, there is likely to be a grey zone where people with glucose levels approaching or just in the diabetic range (>7 mmol l\(^{-1}\)) may be either at the top end of a normal range of physiological glucose that shifts slightly upwards with age, or at the bottom end of the pathophysiological range and on their way to diabetes (Fig. 1).

Common genetic variants at four loci are associated with fasting glucose levels at genome-wide levels of significance \([11^*,12^*,28–34,46]\) and yet only one of these loci is clearly associated with T2D risk \([11^*,12^*]\) in the direction predicted – that is the glucose raising allele also increases the risk of T2D. Common variants in or near the glucokinase (\(GCK\)) \([27,31]\), glucose-6-phosphatase catalytic unit 2 (\(G6PC2\)) \([28,30]\), glucokinase regulatory protein (\(GCKR\)) \([29,32,33]\) and melatonin receptor 1B (\(MTNR1B\)) genes \([11^*,12^*,34,46]\) are associated with fasting glucose levels. Details of these associations and genes, including effect sizes are given in Table 1. It is important to note that the association of a variant in or close to a gene does not necessarily mean that it acts on that gene – often several genes cluster in a small region of DNA and variants within one gene could act on other genes nearby. Despite this qualification, the four associations strongly implicated these genes, because all four are clear candidates for a role in glucose homeostasis. The four genes...
### Table 1 Details of gene variants associated with type 2 diabetes and fasting glucose at genome-wide levels of significance

<table>
<thead>
<tr>
<th>SNP</th>
<th>Nearest gene</th>
<th>Discovery method</th>
<th>Odds ratio (95% CI)/per allele effect size mmol/l (95% CI)</th>
<th>Possible disease mechanism</th>
<th>Possible function of nearest gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs10922931</td>
<td>NOTCH2</td>
<td>Genome-wide</td>
<td>1.13 (1.08–1.17)</td>
<td>Unknown</td>
<td>Encodes a receptor for membrane bound ligands that regulate cell fate determination that have a possible role in vascular, renal and hepatic development</td>
</tr>
<tr>
<td>rs7578597</td>
<td>THADA</td>
<td>Genome-wide</td>
<td>1.15 (1.10–1.20)</td>
<td>Unknown</td>
<td>Implicated in thyroid adenomas</td>
</tr>
<tr>
<td>rs2943641</td>
<td>IRS1</td>
<td>Genome-wide</td>
<td>1.19 (1.13–1.25)</td>
<td>Insulin resistance</td>
<td>Transmits signals to intracellular pathways upon activation by insulin and insulin-like growth factor receptors</td>
</tr>
<tr>
<td>rs1801282</td>
<td>PPARG</td>
<td>Candidate</td>
<td>1.19 (1.12–1.25)</td>
<td>Insulin resistance</td>
<td>Transcription factor involved in adipocyte differentiation and glucose homeostasis</td>
</tr>
<tr>
<td>rs4402960</td>
<td>IGF2BP2</td>
<td>Genome-wide</td>
<td>1.14 (1.11–1.18)</td>
<td>β-cell dysfunction</td>
<td>Binds insulin-like growth factor involved in pancreatic development</td>
</tr>
<tr>
<td>rs4607103</td>
<td>ADAMTS9</td>
<td>Genome-wide</td>
<td>1.09 (1.06–1.12)</td>
<td>Unknown</td>
<td>Metallopeptase expressed in skeletal muscle and pancreas. Implicated in the control of organ shape development and inhibition of angiogenesis</td>
</tr>
<tr>
<td>rs10010131</td>
<td>WFS1</td>
<td>Candidate</td>
<td>1.12 (0.90–1.15)</td>
<td>β-cell dysfunction</td>
<td>Transmembrane protein primarily located in the endoplasmic reticulum. Regulator of cellular Ca(2+) homeostasis</td>
</tr>
<tr>
<td>rs10946398</td>
<td>CDKL1</td>
<td>Genome-wide</td>
<td>1.14 (1.11–1.17)</td>
<td>Β-cell dysfunction</td>
<td>Unknown but a possible role in the inhibition of the CDK5/p35 complex in pancreatic β-cells</td>
</tr>
<tr>
<td>rs864745</td>
<td>JAZF1</td>
<td>Genome-wide</td>
<td>1.10 (1.07–1.13)</td>
<td>β-cell dysfunction</td>
<td>Transcriptional repressor</td>
</tr>
<tr>
<td>rs13266534</td>
<td>SLC30A8</td>
<td>Genome-wide</td>
<td>1.15 (1.12–1.19)</td>
<td>β-cell dysfunction</td>
<td>Zinc transporter expressed exclusively in β-cells</td>
</tr>
<tr>
<td>rs10811861</td>
<td>CDK9A2/2B</td>
<td>Genome-wide</td>
<td>1.20 (1.14–1.25)</td>
<td>β-cell dysfunction</td>
<td>A tumour suppressor/controls cell cycle. Both inhibit cyclin-dependent kinase 4, a regulator of β-cell replication</td>
</tr>
<tr>
<td>rs12779790</td>
<td>CDC123/CAK1D</td>
<td>Genome-wide</td>
<td>1.11 (1.07–1.14)</td>
<td>β-cell dysfunction</td>
<td>Cell cycle regulator/Possible regulator of granulocyte function</td>
</tr>
<tr>
<td>rs1111875</td>
<td>HHEX/IDE</td>
<td>Genome-wide</td>
<td>1.15 (1.11–1.19)</td>
<td>β-cell dysfunction</td>
<td>Transcription factor involved in developmental processes/involved in peptide signalling and breaking down insulin and other peptides.</td>
</tr>
<tr>
<td>rs7901695</td>
<td>TCFL7L2</td>
<td>Region-wide</td>
<td>1.37 (1.31–1.43)</td>
<td>β-cell dysfunction</td>
<td>Transcription factor involved in Wnt signalling pathway</td>
</tr>
<tr>
<td>rs2237892</td>
<td>KCNJ11</td>
<td>Genome-wide</td>
<td>1.25 (1.11–1.5)</td>
<td>β-cell dysfunction</td>
<td>Encodes the pore-forming subunit of the voltage-gated potassium channel</td>
</tr>
<tr>
<td>rs5215(E229K)</td>
<td>KCNJ11</td>
<td>Candidate</td>
<td>1.14 (1.10–1.19)</td>
<td>β-cell dysfunction</td>
<td>Encodes β-cell ATP-sensitive potassium (KATP) channel</td>
</tr>
<tr>
<td>rs7961581</td>
<td>TSFAN8/LGR5</td>
<td>Genome-wide</td>
<td>1.09 (1.06–1.12)</td>
<td>β-cell dysfunction</td>
<td>Cell-surface glycoprotein. Regulates cell development, activation and motility/signal receptor that may be involved in growth and differentiation</td>
</tr>
<tr>
<td>rs8050136</td>
<td>FTO</td>
<td>Candidate</td>
<td>1.17 (1.12–1.22)</td>
<td>Altered BMI</td>
<td>Unknown: associated with BMI</td>
</tr>
<tr>
<td>rs4430796</td>
<td>TCF2</td>
<td>Candidate</td>
<td>1.10 (1.07–1.14)</td>
<td>β-cell dysfunction</td>
<td>Transcription factor</td>
</tr>
<tr>
<td>rs560887</td>
<td>G6PC2</td>
<td>Genome-wide</td>
<td>−0.06 (−0.08–0.05)</td>
<td>β-cell function</td>
<td>Possibly dephosphorylate glucose-6-phosphate produced by glucokinase (by similarity to liver glucose-6-phosphatase)</td>
</tr>
<tr>
<td>rs780094</td>
<td>GCKR</td>
<td>Genome-wide</td>
<td>−0.028</td>
<td>Insulin resistance</td>
<td>Inhibits glucokinase by binding to it</td>
</tr>
<tr>
<td>rs1799884</td>
<td>GCK</td>
<td>Genome-wide</td>
<td>0.06</td>
<td>β-cell dysfunction</td>
<td>Catalyses the first step in glucose utilization</td>
</tr>
<tr>
<td>rs10830963</td>
<td>MTNR1B</td>
<td>Genome-wide</td>
<td>1.09 (1.05–1.12)</td>
<td>β-cell dysfunction</td>
<td>Encodes the high affinity receptor for melatonin.</td>
</tr>
</tbody>
</table>

Odds ratios (95% confidence intervals) from [26] except for PPARG, which is from [14]; JAZF1, CDC123/CAK1D, TSFAN8/LGR5, THADA, ADAMTS9 and NOTCH2 from [9]; KCNJ11 [11*]; MTNR1B from [11*]; IRS1 from [13*]; GCK from [27], G6PC2 from [28] and GCKR from [29]. For fasting glucose associated variants effect sizes reported are in mmol/l and per minor allele. 95% CIs are given in parentheses wherever reported. ADAMTS9, ADAM metalloepitclease with thrombospondin type 1 motif 9; CAMK1D, calcium/calmodulin-dependent protein kinase 1D; CDC123, cell division cycle 123 homologue (Saccharomyces cerevisiae); CDKL1, CDK5 regulatory subunit-associated protein-like-1; CDK9A2/2B, cyclin-dependent kinase inhibitor 2A; FTO, fat mass and obesity associated; GCK, glucokinase; GCKR, glucokinase regulator; G6PC2, glucose-6-phosphatase catalytic subunit-related protein 2; HHEX, haemato poietically expressed homeobox; IDE, insulin-degrading enzyme; IGF2BP2, insulin-like growth factor 2 mRNA-binding protein 2; IRS1, insulin receptor substrate 1; JAZF1, juxtaposed with another zinc finger gene 1; KCNJ11, potassium inwardly rectifying channel, subfamily J, member 11; KCNQ1, potassium voltage-gated channel subfamily KQT member 1; LGR5, leucine-rich repeat containing G protein coupled; MTNR1B, melatonin receptor type 1B; NOTCH2, Notch homologue 2 (Drosophila); PPARG, peroxisome proliferator-activated receptor-γ gene; SLC30A8, solute carrier family 30 (zinc transporter), member 8; TCF2, transcription factor 2, hepatic; TCF7L2, transcription factor 7 like 2; THADA, thyroid adenoma associated; TSFAN8, tetraspanin 8; WFS1, Wolfram syndrome 1.

### Notes
- Variants identified primarily by type 2 diabetes studies [1*1–2, *9, 10*–13, 15, 17–19]
- Variants identified primarily by fasting glucose studies [1*1, 12*, 27–34]

---

**Nutrition and metabolism**
deserve particular discussion because each variant appears to have different effects on a range of other metabolic traits, including T2D.

(1) Glucokinase (GCK) is an enzyme expressed in the pancreas and the liver and is involved in the first rate-limiting step of glycolysis [48]. Common variation in the promoter of GCK is associated with fasting glucose across the age ranges, including in children [31]. The association appears physiologically to be very similar to the effect of rare mutations in GCK, that cause an approximately 2 mmol l⁻¹ increase in fasting glucose throughout life [49–52], the difference being that the common variant has a small effect on fasting glucose and the rare mutations have a large effect.

(2) Two independent GWASs identified common variants near G6PC2 and ATP-binding cassette subfamily B member 11 (ABC811) genes associated with fasting glucose concentrations [28,30]. The allele that raises fasting glucose is not associated with increased risk of T2D. It is possible that the causal variant is acting through the ABC811 locus, which is primarily expressed in the liver [30]. However, in-vitro functional studies have suggested that a common promoter variant in G6PC2 is likely to explain the association results seen in GWASs [53] and G6PC2 is the better candidate [30]. G6PC2 is expressed almost exclusively in the pancreatic β-cells [54] and G6PC2-knockout mice have reduced fasting plasma glucose [55]. The structural similarity between G6PC2 and the hepatic enzyme glucose-6-phosphatase [54,56], which catalyses the terminal step in gluconeogenesis, has led to the suggestion that G6PC2 alters the set point for glucose-stimulated insulin secretion but that additional factors may be required for the impairment of β-cell function that leads to diabetes [28].

(3) The GCKR protein regulates the activity of GCK by reversibly binding to it in the presence of fructose 6-phosphate and fructose 1-phosphate, which in turn increase or reduce GCKR protein activity, respectively [57]. Overexpression of GCKR in mice leads to increased GCK activity and lower fasting glucose [58] and GCKR-deficient mice show normal GCK activity but impaired glucose clearance [59,60]. This variant is also associated with several other important metabolic phenotypes, and not necessarily in the expected direction, given epidemiological correlations. For example, the allele associated with reduced fasting glucose is also associated with increased triglyceride levels [4,29,32,33], C-reactive protein levels [29,61] and apolipoprotein B (ApoB) levels [29]. The allele associated with reduced fasting glucose levels is also associated with a modest reduced risk of T2D [29,32,33], although the level of significance is not comparable to the other loci discovered through T2D GWASs. Recent functional studies have shown that the association observed with the P446L variant is mediated by increased GCK activity in the liver resulting in increased glycolytic flux and hence reduced levels of fasting glucose [62]. It is also suggested that increased glycolysis results in elevated levels of malonyl-CoA, a substrate for de-novo lipo-
genetic and an inhibitor of the fatty oxidating enzyme, carnitine-palmitoyl transferase-1, which in turn results in increased triglyceride levels [62].

(4) Melatonin receptor 1B (MTNR1B), highly expressed in the brain and retina [63] and also in human and rodent islets [64] may regulate the circadian rhythm of insulin secretion through the release of melatonin [65,66]. A recent study showed that the risk allele was associated with an impairment of early insulin response and both non-diabetic carriers of the risk allele and diabetic patients have increased levels of MTNR1B expression in islets [34]. Furthermore, the glucose-stimulated insulin release was inhibited in the presence of melatonin suggesting that melatonin has a direct inhibitory effect on insulin secretion [34]. Another study has also shown that the risk allele is associated with impairment in glucose-stimulated insulin secretion [46]. These genetic associations provide evidence that altered sleep patterns and melatonin levels may play an aetiological role in the development of T2D, as opposed to a possibly confounded epidemiological association.

Some gene variants that influence type 2 diabetes risk and fasting glucose also influence birth weight

Recent studies have provided the first strong evidence that genetic variation could partially explain the association between reduced foetal growth and T2D. Since the first description in 1990, many epidemiological studies have described an association between reduced birth weight and the development of T2D and other metabolic traits later in life [67–71]. Despite these observations, the cause of the association between reduced foetal growth and adverse metabolic traits in later life is not known. The foetal-programming hypothesis suggests the association is caused by undernutrition in the uterus, resulting in reduced foetal growth and permanent reprogramming of foetal metabolism [72,73]. In contrast, the ‘foetal insulin hypothesis’ suggests that genetic variants that influence insulin resistance or secretion may also affect birth weight through effects on insulin-mediated foetal growth [74]. Therefore, genetic factors that affect insulin secretion or sensitivity in the fetus may have an indirect effect on birth weight and direct effect on diabetes risk – the association could be two phenotypes of the same genotype. Maternal genetic factors may also have an indirect effect on offspring birth weight if they influence mother’s insulin secretion and sensitivity [75].

Recent studies show that three common genetic variants associated with altered glycaemia are associated with birth weight. Two of these, GCK [27,31] and TCF7L2 [76], influence birth weight through maternal genotype. A third, CDKAL1, influences birth weight through foetal genotype [77*]. When present in the mother, the glucose-raising allele in GCK results in an average 32 g [95% confidence interval (CI) 11–53 g, P = 0.002] increase in offspring birth weight [31]. No independent effect of foetal genotype was observed [31]. When present in the mother, each additional TCF7L2 T2D risk allele results in a 30 g (95% CI 15–45 g, P = 2.8 × 10^{-5}) increase in offspring birth weight [76]. Analysis of diabetes-related traits by the same study has suggested that the mechanism behind this is likely to be reduced maternal insulin secretion associated with TCF7L2 variants leading to maternal hyperglycaemia, which in turn stimulates foetal insulin secretion and hence insulin-mediated foetal growth. Again, no independent effect of foetal genotype was observed.

More recently, the T2D risk variant in the CDKAL1 locus has been associated with birth weight [77*,78]. Unlike the TCF7L2 and GCK polymorphisms, the presence of diabetes risk variants in CDKAL1 in the foetus was associated with reduced birth weight, but there is no evidence of a maternal genotype effect [77*]. The polymorphism at the CDKAL1 locus is associated with reduced insulin secretion [41] and therefore it is likely that reduced insulin secretion by the fetus results in reduced foetal growth and low birth weight [77*]. There is some evidence that the T2D risk allele at the HHEX/IDE locus lowers birth weight when present in the fetus [77*,79] and suggestive evidence that some risk alleles may have a stronger effect on T2D risk in individuals of low birth weight compared with those of higher birth weights [79].

The genetic link between variation in birth weight and metabolic diseases requires additional study. Two questions of importance arise. First, why do some variants robustly associated with glycaemic traits associate with birth weight through the maternal genotype, some through the foetal genotype and some apparently not at all? The answer to this question may lead to important insights into the timing and mechanisms of how these variants and the genes they act on function. Second, are there genetic links between birth weight and other metabolic traits, such as lipid levels, hypertension and coronary heart disease?

Conclusion

Over the past 3 years, genetic studies have provided a greatly improved knowledge of the genetic variants that influence the risk of T2D and related metabolic traits. Further research is needed to understand which genes are responsible for the many robust statistical associations. Despite this caveat, the results from GWASs are providing many interesting insights into the disease and related metabolic traits.
References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:
• of special interest
• of outstanding interest

Additional references related to this topic can also be found in the Current Literature section in this issue (p. 87).


2 One of the two studies to report the first T2D-associated variant discovered using a non-European cohort and subsequently replicated in other Asian and European populations.


7 The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3000 shared controls. Nature 2007; 447:661–678.


11 One of the two studies to report the first T2D-associated variant discovered using a non-European cohort and subsequently replicated in other Asian and European populations.


13 One of the first studies to report a fasting glucose-associated variant that is also associated with increased risk of T2D.


15 One of the first two studies to report a fasting glucose-associated variant that is also associated with increased T2D.


17 This study reports the most recent addition to the GWAS T2D variants and shows that its primary effect is through insulin resistance.


50 Nutrition and metabolism


