New gene variants alter type 2 diabetes risk predominantly through reduced beta-cell function

John R.B. Perry and Timothy M. Frayling

Genetics of Complex Traits, Institute of Biomedical and Clinical Science, Peninsula Medical School, Exeter, UK

Correspondence to Timothy M. Frayling, Professor of Human Genetics, Genetics of Complex Traits, Peninsula Medical School, Magdalen Road, Exeter EX1 2LU, UK
Tel: +44 1392 262935;
e-mail: tim.frayling@pms.ac.uk

Current Opinion in Clinical Nutrition and Metabolic Care 2008, 11:371–377

Purpose of review

Over the past 18 months, the number of gene loci robustly associated with type 2 diabetes has risen from three to 18. In this study, we focus on explaining the genome-wide approach that has led to most of these discoveries and discuss some of the early insights the new gene loci have provided into the aetiology of type 2 diabetes.

Recent findings

Recent genome-wide association studies have provided an important resource for furthering our understanding of type 2 diabetes disease mechanisms. Genes previously unsuspected of playing a role in diabetes are now implicated in the disease process. These include genes in cell cycling control (*CDKN2A/2B*, *CDKAL1*), transcription factors (*TCF7L2*, *HHEX*), and ion channels (*SLC30A8*). These variants are all associated with insulin-secretory defects in the general population and show little if any relationship to insulin resistance. Two common variants (near or in *FTO* and *MC4R*) alter diabetes risk through a primary effect on obesity.

Summary

Recent genome-wide association studies show that there are now 18 gene loci associated with the risk of type 2 diabetes. Most of these T2D gene loci affect insulin secretion.

Keywords

beta-cell dysfunction, genome-wide association, insulin secretion, obesity genetics, type-2 diabetes genetics

Curr Opin Clin Nutr Metab Care 11:371–377

1363-1950

Introduction

The advent of genome-wide association studies (GWASs) in 2007 has resulted in huge advances in complex trait genetics. Many of the genes now implicated were not expected and so provide completely new insights into common causes of a disease. One of the biggest success stories has been in type 2 diabetes (T2D), where very recent progress has taken the number of confirmed gene loci to 18. One of the reasons behind this success, in a condition that has a considerable environmental component, is that researchers of type 2 diabetes were among the first to combine data from multiple GWAS, providing more power to detect disease loci.

Candidate genes

Until the recent GWASs, the identification of type 2 diabetes genes had been frustratingly slow and difficult. Over the past decade, human geneticists had predominately relied on two approaches for gene discovery: Detailed analyses of genomic regions shared between diabetic relatives more often than expected by chance (linkage approach), and candidate gene studies. Of these efforts, only the candidate-gene approach has lead to any reproducible results.

The candidate-gene approach provided several important lessons for gene discovery efforts, including the importance of using dense sets of single-nucleotide polymorphisms (SNPs) to capture a large proportion of the common variation across genes. This allows assessment of variation in introns and considerable distance upstream or downstream of a gene, not just the coding or obviously functional regions. Here, data from the HapMap project [1] helped inform which SNPs were redundant due to correlation between common variants. These studies also demonstrated the importance of using several thousand cases and controls together with even larger replication studies to provide robust evidence of association.

Many common diseases, including type 2 diabetes, have several related rare monogenic forms. If mutations in a gene result in diabetes, it suggests the gene product is crucial for normal insulin secretion or sensitivity. These genes are therefore also excellent candidates for a role in common forms of the disease. For type 2 diabetes, there are several single-gene disorders of relevance, including maturity-onset diabetes of the young (MODY) and permanent neonatal diabetes mellitus (PNDM).

Studying common variation in monogenic diabetes genes proved a successful approach. Common variants in or near four genes involved in monogenic diabetes disorders [2–5] predispose to type 2 diabetes; these genes are *KCNJ11*, *PPARG*, *TCF2/HNF1B* and *WFS1* [6–8,9•,10–14]. A common variant in a fifth monogenic gene, *TCF1*, encoding *HNF1A*, is also a risk factor for type 2 diabetes in the Oji–Cree population, but this variant is only found in this population of native Canadians [15,16].

The noncandidate-gene-based approach to T2D gene discovery

The identification of common risk variants in *TCF7L2* represented an important milestone in T2D genetics. *TCF7L2* was the first T2D gene locus to be identified without any prior knowledge of its function.

TCF7L2 was identified in 2006 by Decode genetics [17]. Following up a region on chromosome 10q identified through linkage analysis, Decode genotyped over 200 polymorphisms across the region in 1185 cases and 931 Icelandic controls. They identified a number of variants at the *TCF7L2* locus in strong linkage disequilibrium with each other, with the most associated SNP present in nearly 28% of chromosomes from individuals without diabetes and nearly 36% of chromosomes from individuals with T2D. This finding has now been replicated in many studies, including those of non-European population groups [18–20].

To date, variants in *TCF7L2* have the largest risk effect of any of the known T2D susceptibility loci – the 7% of European individuals carrying two copies of the risk allele are almost twice as likely to get diabetes than the 55% of individuals carrying two protective alleles. *TCF7L2* is a component of the WNT signalling pathway, a target of the HHEX gene that is also implicated in T2D. *TCF7L2* is a transcription factor that acts as a nuclear receptor for *CTNNB1* (beta-catenin) [21]. WNT signalling plays a crucial role in cell proliferation, motility and normal embryogenesis, and has been shown to regulate myogenesis and adipogenesis [22,23]. Tight regulation of WNT signalling is also required for normal pancreatic development during embryonic growth [24].

New genotyping technology results in the identification of six new type 2 diabetes gene loci

Six new gene loci were identified in early 2007 through independent GWASs. This took the number of disease susceptibility variants to 11, including the four loci discovered by candidate-gene approaches and TCF7L2. The start of 2008 saw a further increase, with an additional seven genes replicated through combined GWASs (see Table 1 and Fig. 1) [25].

This success was the result of new technology combined with large DNA resources. Genome-wide association scans allow hundreds of thousands of SNPs, spread across the entire genome, to be analysed in a single experiment. New technology in the form of DNA chips allows researchers to capture a large proportion of genomic variation, in thousands of DNA samples. Statistically significant allele frequency differences between cases and controls can then highlight potential polymorphisms, and ultimately genes, which are implicated in disease. This hypothesis-free approach provides the opportunity to identify completely unexpected genes, broadening our understanding of disease mechanisms.

The two main commercial genome-wide genotyping platforms have utilized different methods to design SNP arrays. SNPs on the Illumina panel are selected to capture optimal genomic variation based on HapMap linkage disequilibrium data. Although the genome contains millions of SNP sites, the correlation of SNPs due to linkage disequilibrium allows a reasonable proportion of these to be 'tagged' by a modestly sized subset of these SNPs. Alternatively, SNPs on the Affymetrix chips were selected without any consideration for linkage disequilibrium – and so were randomly distributed across the genome. Although the Affymetrix approach would give a reduction in variation coverage, it benefits from not being population specific and also has an inbuilt redundancy (due to linkage disequilibrium) to help compensate for genotyping failures. The limitation of both chips, however, is that the arrays are not exhaustive, with poor coverage of both rare SNPs [minor allele frequency (MAF) <5%] and structural variation such as copy number variants (CNVs).

On a genotyping platform comprising nearly 500 000 SNPs, an entire spectrum of 'significant' P values will occur through chance. For example, by performing 500 000 statistical tests, we would expect a P value of 2×10^{-6} (1/500 000) to occur once by chance. In addition, the prior probability of any variant being a risk factor for diabetes is very small. These statistical arguments mean that P values lower than approximately 5×10^{-7} are required to provide robust statistical evidence of association.

Table 1 Details of variants and implicated genes associated with type 2 diabetes

Example variant	Closest gene	Mode of identification	Current evidence (*P* value)[a]	Predicted disease mechanism	Odds ratio (per allele)[a]	RAF (UK)
rs5215 (E23K)	*KCNJ11*	Candidate	5×10^{-11}	Beta-cell dysfunction	1.14 (1.10–1.19)	0.35
rs1111875	*HHEX/IDE*	Genome wide	7×10^{-17}	Beta-cell dysfunction	1.15 (1.10–1.19)	0.65
rs13266634	*SLC30A8*	Genome wide	1×10^{-19}	Beta-cell dysfunction	1.15 (1.12–1.19)	0.69
rs10946398	*CDKAL1*	Genome wide	2×10^{-18}	Beta-cell dysfunction	1.14 (1.11–1.17)	0.32
rs7901695	*TCF7L2*	Region wide	1×10^{-48}	Beta-cell dysfunction	1.37 (1.31–1.43)	0.31
rs10811661	*CDKN2A/2B*	Genome wide	8×10^{-15}	Beta-cell dysfunction	1.20 (1.14–1.25)	0.83
rs4402960	*IGF2BP2*	Genome wide	9×10^{-16}	Beta-cell dysfunction	1.14 (1.11–1.18)	0.32
rs8050136	*FTO*	Genome wide	1×10^{-12}	Altered BMI	1.17 (1.12–1.22)	0.4
rs17782313	*MC4R*	Genome wide	0.003	Altered BMI	1.09 (1.03–1.15)	0.24
rs1801282 (P12A)	*PPARG*	Candidate	2×10^{-6}	Unknown	1.14 (1.08–1.20)	0.87
rs4430796	*TCF2*	Candidate	8×10^{-10}	Unknown	1.10 (1.07–1.14)	0.47
rs10010131	*WFS1*	Candidate	4×10^{-11}	Unknown	1.12 (1.09–1.15)	0.60
rs864745	*JAZF1*	Genome wide	5×10^{-14}	Unknown	1.10 (1.07–1.13)	0.50
rs12779790	*CDC123/CAMK1D*	Genome wide	1×10^{-10}	Unknown	1.11 (1.07–1.14)	0.18
rs7961581	*TSPAN8/LGR5*	Genome wide	1×10^{-9}	Unknown	1.09 (1.06–1.12)	0.27
rs7578597	*THADA*	Genome wide	1×10^{-9}	Unknown	1.15 (1.10–1.20)	0.90
rs4607103	*ADAMTS9*	Genome wide	1×10^{-8}	Unknown	1.09 (1.06–1.12)	0.76
rs10923931	*NOTCH2*	Genome wide	4×10^{-8}	Unknown	1.13 (1.08–1.17)	010

[a] Approximate *P* values and odds ratios calculated by meta-analysis of individual study odds ratios from [23,24,27•] except for the signals for HHEX-IDE, which also includes data from [25], CDKAL1, which also includes data from [26••], SLC30A8, which includes data from five GWASs listed above, TCF2, which is based on data from [8,11] and WFS1, which is based on data from [14]. JAZF1, CDC123, TSPAN8, THADA, ADAMTS8 and NOTCH2 data from [28•], MC4R data from [42]. ADAMTS9, ADAM metallopeptidase with thrombospondin type 1 motif 9. BMI, body mass index; CAMK1D, calcium/calmodulin-dependent protein kinase 1D; CDC123, cell division cycle 123 homologue (*Saccharomyces cerevisiae*); CDKAL1, CDK5 regulatory subunit-associated protein 1 like 1; CDKN2, cyclin-dependent kinase inhibitor 2A; FTO, fat mass and obesity associated; HHEX, haematopoietically expressed homeobox; IDE, insulin-degrading enzyme; IGF2BP2, insulin-like growth factor 2 mRNA-binding protein 2; JAZF1, juxtaposed with another zinc finger gene 1; KCNJ11, potassium inwardly rectifying channel, subfamily J, member 11; LGR5, leucine-rich repeat-containing G protein coupled; MC4R, melanocortin 4 receptor; NOTCH2, Notch homologue 2 (*Drosophila*); PPARG, peroxisome proliferator-activated receptor-γ gene; RAF, risk allele frequency; SLC30A8, solute carrier family 30 (zinc transporter), member 8; TCF2, transcription factor 2, hepatic; LF-B3, variant hepatic nuclear factor; TCF7L2, transcription factor 7 like 2 (T-cell specific, HMG box); THADA, thyroid adenoma associated; TSPAN8, tetraspanin 8; WFS1, Wolfram syndrome 1. Adapted from Frayling [25].

Results from the first T2D GWASs were published in early 2007, providing six replicated new susceptibility loci [26••,27•,28•,29,30,31•]. Six studies, describing five separate T2D scans, genotyped a combined total of 18 000 individuals of Northern European ancestry. Large additional replication cohorts were an essential part of each study design, totalling over 55 000 individuals across the five studies.

These efforts were swiftly followed by researchers combining information from three of these GWA scans (WTCCC, DGI and FUSION), combining information in a meta-analysis of over 10 000 samples [32]. This meta-analysis used imputed data to test approximately 2.2 million SNPs for disease association. Imputation algorithms allow the prediction of untyped SNP genotypes based on patterns of haplotype variation from the HapMap dataset [33,34].

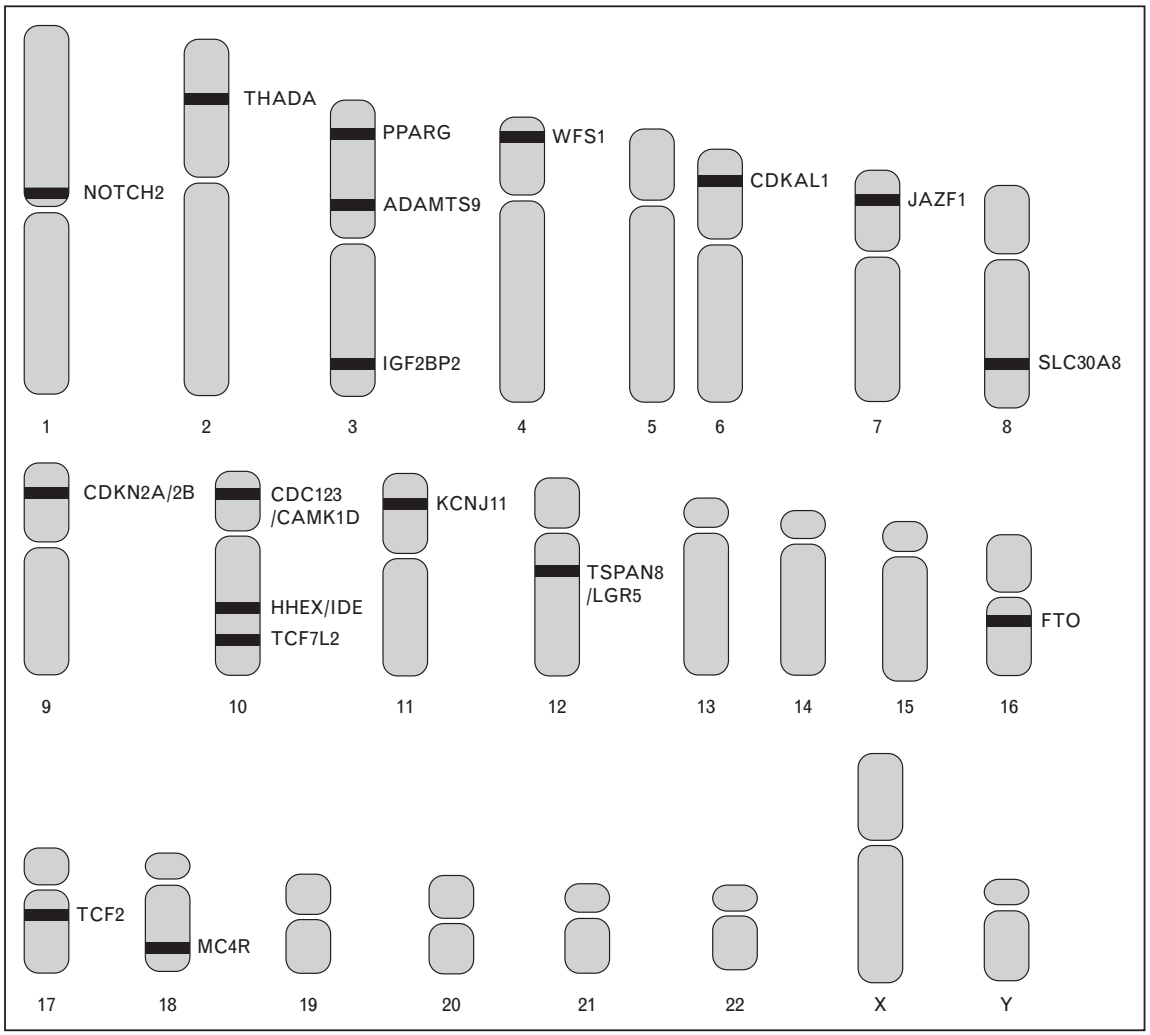
Seven of the 18 common T2D variants alter disease risk through altered beta-cell function

There is increasing evidence to suggest the genetic component to T2D risk acts through beta-cell dysfunction – seven of the gene loci have been associated with impaired insulin secretion in the general population, but there is little if any evidence that any of the loci alter insulin resistance (see Fig. 2).

The common variants in the *TCF7L2* gene predispose to type 2 diabetes by reducing beta-cell function and insulin secretion. Multiple studies of individuals without diabetes show that measures of first-phase insulin release are reduced in individuals carrying T2D risk alleles, but there is no evidence for a role in insulin resistance [35–38]. *TCF7L2* type 2 diabetes risk alleles also influence birth weight [39•]. Each copy of the diabetes risk allele in mothers is associated with a nearly 30 g increase in offspring birth weight. The mechanism behind this is likely to be reduced maternal insulin secretion causing elevated blood glucose, which, in turn, increases foetal insulin levels, an important foetal growth factor.

Risk alleles at the *CDKAL1*, *SLC30A8*, *HHEX-IDE*, *CDKN2A/B*, *KCNJ11* and *IGF2BP2* loci all alter insulin secretion [10,30,40•,41•,42,43]. For each loci, there are several studies comprising a few hundred to 2000 individuals with detailed oral glucose tolerance test (OGTT) or detailed physiological clamp studies. Some studies have shown insulin-secretion defects at one gene and not others, but overall a consistent picture is emerging that these loci alter diabetes risk by reducing the insulin-secretory capacity of the beta-cells. There is no evidence

Figure 1 Genomic location of all 18 T2D disease loci



that any of these loci alter insulin resistance and for the six most recently described loci there are no data yet to address the question of whether insulin secretion or insulin resistance is the primary causal pathway.

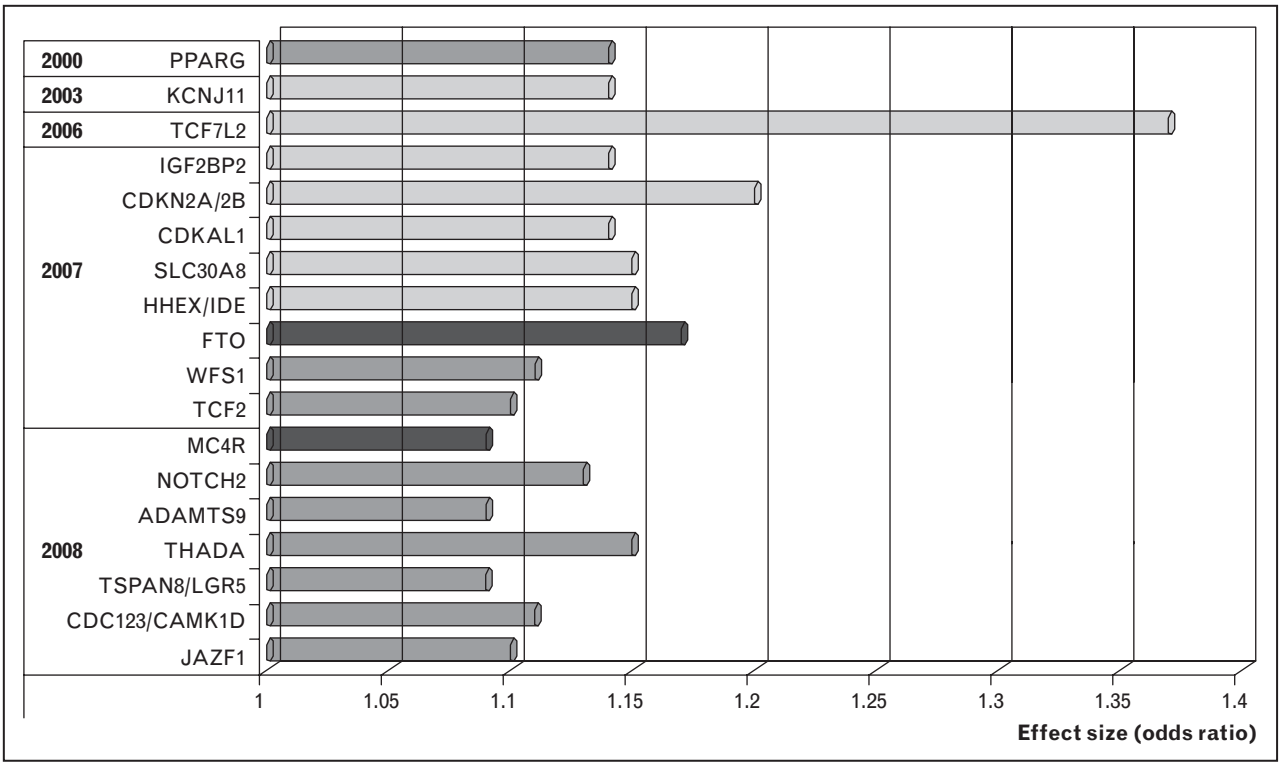
Two type 2 diabetes genes influence fat mass and obesity risk

Two of the 18 loci mediate their effect on T2D risk by altering BMI. Common variants in the *FTO* and *MC4R* genes represent the only polymorphisms reproducibly associated with altered BMI and obesity risk.

FTO was the first of these gene loci to be discovered [44•]. A common variant in intron 1 of the *FTO* gene was the second strongest locus associated with type 2 diabetes after *TCF7L2* in the UK T2D GWAS. Subsequent analyses later confirmed the locus-altered BMI, which, in turn, conferred diabetes risk. Using 30 000 individuals, Frayling *et al.* [44•] reported that 16% of the European population carrying two copies of the T2D risk allele were approximately 2–3 kg heavier than the 35% with no risk alleles. *FTO* genotypes alter metabolic traits inline with the effect on BMI, indicating that *FTO* variants do not independently alter other metabolic traits [45]. Early evidence suggests that the gene may influence appetite [46]; however, more functional research is required to confirm this.

Common variants near the *MC4R* gene were also recently found to be associated with fat mass and obesity risk [47]. *MC4R* variants were identified during a BMI GWAS on approximately 17 000 individuals, which ranked variants in *FTO* at the top. The common variant associated with BMI is 188 kb from the *MC4R* gene but is a compelling biological candidate because rare mutations are the commonest cause of monogenic obesity in humans [48].

Figure 2 Effect sizes, risk mechanism and year of discovery for all 18 T2D disease loci



The *Y*-axis gives year published for evidence that gene loci reached genome-wide level significance. ADAMTS9, ADAM metallopeptidase with thrombospondin type 1 motif 9; CAMK1D, calcium/calmodulin-dependent protein kinase 1D; CDC123, cell division cycle 123 homologue (*Saccharomyces cerevisiae*); CDKAL1, CDK5 regulatory subunit-associated protein 1-like 1; CDKN2, cyclin-dependent kinase inhibitor 2A; FTO, fat mass and obesity associated; HHEX, haematopoietically expressed homeobox; IDE, insulin-degrading enzyme; IGF2BP2, insulin-like growth factor 2 mRNA-binding protein 2; JAZF1, juxtaposed with another zinc finger gene 1; KCNJ11, potassium inwardly rectifying channel, subfamily J, member 11; LGR5, leucine-rich repeat-containing G protein coupled; MC4R, melanocortin 4 receptor; 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; LF-B3, variant hepatic nuclear factor; TCF7L2, transcription factor 7 like 2 (T-cell specific, HMG box); THADA, thyroid adenoma associated; TSPAN8, tetraspanin 8; WFS1, Wolfram syndrome 1. □ Beta-cell; ▒ unknown; ■ BMI.

A genetic overlap between type 2 diabetes and prostate cancer

Several epidemiological studies have highlighted an inverse relationship between T2D and prostate cancer risk [49]. One of the most interesting findings of recent GWASs is that common variation in the T2D susceptibility genes *TCF2* and *JAZF1* also influence prostate cancer risk [9•,50].

For *TCF2*, the risk allele of the T2D susceptibility variant has a protective effect on prostate cancer susceptibility, whereas for *JAZF1* there are two different variants predisposing to the two diseases, which are not correlated with each other. The mechanisms by which these genes influence disease are unknown; however, it is possible that they influence cell-cycle control. It is possible that on the one hand a particular allele predisposes to increased cell proliferation, while the alternative allele could influence cell senescence or death, potentially affecting beta-cell mass and function. This theory might also explain the number of cell-cycle genes that have been implicated in T2D susceptibility.

Next steps

Having identified a handful of loci responsible for T2D susceptibility, the next stage is to identify the mechanisms by which these genes act on disease risk. Further physiological studies, alongside research by clinical, cellular, animal and molecular biologists will hopefully lead to further advances in our understanding of the disease. Alongside these efforts, there still remain many more susceptibility loci to find, as only a small fraction of heritability can be explained by the known variants.

Further meta-analysis of studies and analysing additional samples will raise the number of disease loci and our

knowledge of disease. For the majority of T2D loci, it is unknown where the causal variant lies and, in some cases, it is unclear which gene might be implicated (e.g., *HHEX/IDE* locus). Deep sequencing and additional genotyping will therefore be required to pinpoint functional variant(s) within these genes.

The analysis of additional types of genomic variation, in particular CNVs will also play an important role in future T2D genetics studies. Known CNVs cover more than 15% of the human genome and are often not well correlated with SNPs. CNVs have been shown to associate with various human traits and diseases, including SLE and HIV-1/AIDS susceptibility [51,52]; however, any potential role in T2D is still unknown.

Conclusion

Type 2 diabetes GWASs over the last 2 years have accelerated our understanding of disease aetiology. There are now 18 gene loci reproducibly associated with T2D, identified through large collaborative efforts. These studies have highlighted genes with no previous biological knowledge, implicating new pathways in T2D genetics. There is now a growing amount of evidence to suggest the genetic component to T2D risk acts through beta-cell dysfunction rather than insulin action.

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 World Literature section in this issue (pp. 524–525).

1 The International HapMap Project. Nature 2003; 426:789–796.

2 Barroso I, Gurnell M, Crowley VE, *et al.* Dominant negative mutations in human PPARgamma associated with severe insulin resistance, diabetes mellitus and hypertension. Nature 1999; 402:880–883.

3 Gloyn AL, Pearson ER, Antcliff JF, *et al.* Activating mutations in the gene encoding the ATP-sensitive potassium-channel subunit Kir6.2 and permanent neonatal diabetes. N Engl J Med 2004; 350:1838–1849.

4 Inoue H, Tanizawa Y, Wasson J, *et al.* A gene encoding a transmembrane protein is mutated in patients with diabetes mellitus and optic atrophy (Wolfram syndrome). Nat Genet 1998; 20:143–148.

5 Nishigori H, Yamada S, Kohama T, *et al.* Frameshift mutation, A263fsinsGG, in the hepatocyte nuclear factor-1beta gene associated with diabetes and renal dysfunction. Diabetes 1998; 47:1354–1355.

6 Altshuler D, Hirschhorn JN, Klannemark M, *et al.* The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. Nat Genet 2000; 26:76–80.

7 Florez JC, Burtt N, de Bakker PI, *et al.* Haplotype structure and genotype–phenotype correlations of the sulfonylurea receptor and the islet ATP-sensitive potassium channel gene region. Diabetes 2004; 53:1360–1368.

8 Gloyn AL, Weedon MN, Owen KR, *et al.* Large-scale association studies of variants in genes encoding the pancreatic beta-cell KATP channel subunits Kir6.2 (KCNJ11) and SUR1 (ABCC8) confirm that the KCNJ11 E23K variant is associated with type 2 diabetes. Diabetes 2003; 52:568–572.

9 • Gudmundsson J, Sulem P, Steinthorsdottir V, *et al.* Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes. Nat Genet 2007; 39:977–983.
The describes a common variant in the TCF2/HNF1B gene, one allele of which predisposes to type 2 diabetes and the other to prostate cancer.

10 Nielsen EM, Hansen L, Carstensen B, *et al.* The E23K variant of Kir6.2 associates with impaired post-OGTT serum insulin response and increased risk of type 2 diabetes. Diabetes 2003; 52:573–577.

11 Sandhu MS, Weedon MN, Fawcett KA, *et al.* Common variants in WFS1 confer risk of type 2 diabetes. Nat Genet 2007; 39:951–953.

12 Winckler W, Weedon MN, Graham RR, *et al.* Evaluation of common variants in the six known maturity-onset diabetes of the young (MODY) genes for association with type 2 diabetes. Diabetes 2007; 56:685–693.

13 Florez JC, Jablonski KA, McAteer J, *et al.* Testing of diabetes-associated WFS1 polymorphisms in the Diabetes Prevention Program. Diabetologia 2007; 51:451–457.

14 Franks PW, Rolandsson O, Debenham SL, *et al.* Replication of the association between variants in WFS1 and risk of type 2 diabetes in European populations. Diabetologia 2007; 51:458–463.

15 Hegele RA, Cao H, Harris SB, *et al.* Hepatocyte nuclear factor-1 alpha G319S. A private mutation in Oji-Cree associated with type 2 diabetes. Diabetes Care 1999; 22:524.

16 Hegele RA, Cao H, Harris SB, *et al.* The hepatic nuclear factor-1alpha G319S variant is associated with early-onset type 2 diabetes in Canadian Oji-Cree. J Clin Endocrinol Metab 1999; 84:1077–1082.

17 Grant SF, Thorleifsson G, Reynisdottir I, *et al.* Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. Nat Genet 2006; 38:320–323.

18 Hayashi T, Iwamoto Y, Kaku K, *et al.* Replication study for the association of TCF7L2 with susceptibility to type 2 diabetes in a Japanese population. Diabetologia 2007; 50:980–984.

19 Helgason A, Palsson S, Thorleifsson G, *et al.* Refining the impact of TCF7L2 gene variants on type 2 diabetes and adaptive evolution. Nat Genet 2007; 39:218–225.

20 Horikoshi M, Hara K, Ito C, *et al.* A genetic variation of the transcription factor 7-like 2 gene is associated with risk of type 2 diabetes in the Japanese population. Diabetologia 2007; 50:747–751.

21 Smith U. TCF7L2 and type 2 diabetes – we WNT to know. Diabetologia 2007; 50:5–7.

22 Etheridge SL, Spencer GJ, Heath DJ, Genever PG. Expression profiling and functional analysis of wnt signaling mechanisms in mesenchymal stem cells. Stem Cells 2004; 22:849–860.

23 Ross SE, Hemati N, Longo KA, *et al.* Inhibition of adipogenesis by Wnt signaling. Science 2000; 289:950–953.

24 Papadopoulou S, Edlund H. Attenuated Wnt signaling perturbs pancreatic growth but not pancreatic function. Diabetes 2005; 54:2844–2851.

25 Frayling TM. Genome-wide association studies provide new insights into type 2 diabetes aetiology. Nat Rev Genet 2007; 8:657–662.

26 •• 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.
The largest genome-wide association study – seven diseases and a joint control group from the UK.

27 • Saxena R, Voight BF, Lyssenko V, *et al.* Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science 2007; 316:1331–1336.
One of three T2D GWASs published together that identified new type 2 diabetes genes.

28 • Scott LJ, Mohlke KL, Bonnycastle LL, *et al.* A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 2007; 316:1341–1345.
One of three T2D GWASs published together that identified new type 2 diabetes genes.

29 Sladek R, Rocheleau G, Rung J, *et al.* A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 2007; 445:881–885.

30 Steinthorsdottir V, Thorleifsson G, Reynisdottir I, *et al.* A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. Nat Genet 2007; 39:770–775.

31 • Zeggini E, Weedon MN, Lindgren CM, *et al.* Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science 2007; 316:1336–1341.
One of three T2D GWASs published together that identified new type 2 diabetes genes.

32 Zeggini E, Scott LJ, Saxena R, *et al.* Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat Genet 2008 [Epub ahead of print].

33 Frazer KA, Ballinger DG, Cox DR, *et al.* A second generation human haplotype map of over 3.1 million SNPs. Nature 2007; 449:851–861.

34 Marchini J, Howie B, Myers S, *et al.* A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 2007; 39:906–913.

35 Cauchi S, Meyre D, Choquet H, *et al.* TCF7L2 variation predicts hyperglycemia incidence in a French general population: the data from an epidemiological study on the Insulin Resistance Syndrome (DESIR) study. Diabetes 2006; 55:3189–3192.

36 Loos RJ, Franks PW, Francis RW, *et al.* TCF7L2 polymorphisms modulate proinsulin levels and beta-cell function in a British Europid population. Diabetes 2007; 56:1943–1947.

37 Marzi C, Huth C, Kolz M, *et al.* Variants of the transcription factor 7-like 2 gene (TCF7L2) are strongly associated with type 2 diabetes but not with the metabolic syndrome in the MONICA/KORA surveys. Horm Metab Res 2007; 39:46–52.

38 Munoz J, Lok KH, Gower BA, *et al.* Polymorphism in the transcription factor 7-like 2 (TCF7L2) gene is associated with reduced insulin secretion in non-diabetic women. Diabetes 2006; 55:3630–3634.

39 • Freathy RM, Weedon MN, Bennett A, *et al.* Type 2 diabetes TCF7L2 risk genotypes alter birth weight: a study of 24 053 individuals. Am J Hum Genet 2007; 80:1150–1161.
The first confirmed example of a diabetes gene variant that influences birth weight.

40 • Grarup N, Rose CS, Andersson EA, *et al.* Studies of association of variants near the HHEX, CDKN2A/B, and IGF2BP2 genes with type 2 diabetes and impaired insulin release in 10 705 Danish subjects: validation and extension of genome-wide association studies. Diabetes 2007; 56:3105–3111.
One of the first studies to highlight the new type 2 diabetes genes as primarily operating through reduced beta-cell function.

41 • Pascoe L, Tura A, Patel SK, *et al.* Common variants of the novel type 2 diabetes genes CDKAL1 and HHEX/IDE are associated with decreased pancreatic beta-cell function. Diabetes 2007; 56:3101–3104.
One of the first studies to highlight the new type 2 diabetes genes as primarily operating through reduced beta-cell function.

42 Staiger H, Machicao F, Stefan N, *et al.* Polymorphisms within novel risk loci for type 2 diabetes determine beta-cell function. PLoS ONE 2007; 2:e832.

43 Staiger H, Stancakova A, Zilinskaite J, *et al.* A candidate type 2 diabetes polymorphism near the HHEX locus affects acute glucose-stimulated insulin release in European populations: results from the EUGENE2 study. Diabetes 2008; 57:514–517.

44 • Frayling TM, Timpson NJ, Weedon MN, *et al.* A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science 2007; 316:889–894.
The first confirmed association between a common variant and obesity.

45 Freathy RM, Timpson NJ, Lawlor DA, *et al.* Common variation in the FTO gene alters metabolic traits to the extent expected, given its effect on BMI. Diabetes 2008 [Epub ahead of print].

46 Gerken T, Girard CA, Tung YC, *et al.* The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. Science 2007; 318:1469–1472.

47 Loos RJ, Lindgren CM, Li S, *et al.* Association studies involving over 90 000 samples demonstrate that common variants near to MC4R influence fat mass, weight and risk of obesity. Nat Genet (in press).

48 Farooqi IS, Keogh JM, Yeo GS, *et al.* Clinical spectrum of obesity and mutations in the melanocortin 4 receptor gene. N Engl J Med 2003; 348:1085–1095.

49 Kasper JS, Giovannucci E. A meta-analysis of diabetes mellitus and the risk of prostate cancer. Cancer Epidemiol Biomarkers Prev 2006; 15: 2056–2062.

50 Thomas G, Jacobs KB, Yeager M, *et al.* Multiple loci identified in a genome-wide association study of prostate cancer. Nat Genet 2008; 40:310–315.

51 Yang Y, Chung EK, Wu YL, *et al.* Gene copy-number variation and associated polymorphisms of complement component C4 in human systemic lupus erythematosus (SLE): low copy number is a risk factor for and high copy number is a protective factor against SLE susceptibility in European Americans. Am J Hum Genet 2007; 80:1037–1054.

52 Mamtani M, Rovin B, Brey R, *et al.* CCL3L1 gene-containing segmental duplications and polymorphisms in CCR5 affect risk of systemic lupus erythematosus. Ann Rheum Dis 2007 [Epub ahead of print]. doi: 10.1136/ard.2007.078048.