research, which has been spawned by well-designed and adequately powered genetic association studies in humans.

*Genome-wide association studies* GWAS have both validated known genes and revealed several novel type 2 diabetes loci. The first GWAS reproduced the robust association of *TCF7L2* (OR 1.65, $p<10^{-7}$) while also identifying *SLC30A8* (OR 1.26, $p<10^{-6}$) and *HHEX* (OR 1.21, $p<10^{-5}$) as two new type 2 diabetes-associated genomic regions [74]. Three additional high-density GWAS, which shared results and were published jointly (representing an aggregate of 32,000 samples), confirmed the known *TCF7L2*, *KCNJ11* and *PPARG* associations as well as the recently published *HHEX* and *SCL30A8* associations. They also discovered the novel diabetes loci *CDKAL1* (OR 1.12, $p<10^{-10}$), *IGF2BP2* (OR 1.14, $p<10^{-15}$) and *CDKN2A/B* (OR 1.20, $p<10^{-14}$) [75–77]. Simultaneously, the deCODE investigators and their collaborators corroborated the strong signal of *TCF7L2* and replicated the *HHEX* and *SCL30A8* findings, while independently identifying *CDKAL1* as an additional locus [78]. An initial association of a variant in the *FTO* gene with diabetes was subsequently discovered to be mediated via its impact on obesity [79]. Suggestive variants reported by GWAS conducted at lower density and/or in smaller samples [80–84] await replication in additional cohorts before they reach similar levels of genome-wide statistical significance ($p\sim5\times10^{-8}$).

Interestingly, many of the newly discovered variants appear to influence insulin secretion rather than insulin resistance (Table 1). For example, Steinthorsdottir et al. demonstrated insulin secretion defects in risk allele carriers at *CDKAL1* and *SLC30A8* [78]. Pascoe et al. performed 75 g OGTTs and hyperinsulinaemic–euglycaemic clamps in 1,276 European participants, and showed that variants in *CDKAL1* and *HHEX* are associated with decreased pancreatic beta cell function as measured by the 30 min insulin response, even after correction for insulin resistance; no variant was associated with insulin sensitivity [85]. Grarup et al. simultaneously reported that variants of *HHEX*, *CDKN2A/B* and *IGF2BP2* were associated with type 2 diabetes, and SNPs within the *HHEX* and *CDKN2A/B* loci impaired glucose-induced insulin release in a Danish population sample and in healthy Danes [86]. Staiger et al. found that the major alleles of the *SLC30A8* SNP rs13266634 and the *HHEX* SNP rs7923837 were associated with reduced insulin secretion stimulated by orally administered glucose, but not with insulin resistance [87]. A quantitative trait analysis of GWAS-identified diabetes susceptibility loci was recently completed by Palmer and colleagues in the Insulin Resistance Atherosclerosis (IRAS) Family Study [88]. In this cohort of 1,268 Hispanic and 581 African-American participants, the diabetes risk allele at *CDKAL1* was associated with a diminished acute insulin response during the IVGTT, and risk variants at *SLC30A8* and *IGF2BP2* appeared to be associated with a reduced disposition index. In sum, with few exceptions, the overwhelming majority of newly discovered diabetes genes appear to influence pancreatic beta cell function (see Fig. 1).

## Are there insulin resistance genes?

Although rare mutations can cause extreme insulin resistance [89], there seems to be a dearth of common genetic variants in established pathways of insulin action that contribute to type 2 diabetes. Indeed, as outlined above, other than *PPARG* (discovered by candidate gene approaches [37]) and *FTO* (discovered by GWAS [79]), most of the newly identified loci are associated with insulin secretion defects. There are several possible explanation for the relative scarcity of insulin resistance genes discovered via GWAS (see text box: Potential reasons for the relative scarcity of insulin resistance genes found via GWAS approaches).

**Potential reasons for the relative scarcity of insulin resistance genes found via GWAS approaches**

1. Study design. Deliberate matching of cases and controls for BMI or selecting for lean cases may bias the scan against variants that increase diabetes risk via their effects on fat accumulation
2. Lower heritability of insulin resistance traits indicates a stronger influence of environmental covariates
3. There may be fewer variants that affect insulin resistance
4. Variants that affect insulin resistance may be less frequent in the population (i.e. more difficult to detect with current arrays)
5. Variants that affect insulin resistance may have more modest effect sizes (i.e. more difficult to distinguish from surrounding statistical noise)
6. Ascertainment criteria may impose constraints around insulin resistance measures that restrict the variance in the trait
7. The measures of insulin sensitivity studied in large human studies may be poorly correlated with insulin resistance at the tissue or molecular level

**Key limitations of current research on novel biomarkers for diabetes, and suggested improvements**

| Area | Limitations | Possible solutions |
|---|---|---|
| Residual confounding | Residual confounding is common: many parameters examined in isolation have some overlap of origin with other proposed novel risk markers so that levels correlate | More comprehensive studies examining a range of parameters from common sources |
| Measurement variation | Consideration of analytical/biological variation limited | Repeat measures in a subgroup may be helpful to correct for regression dilution, and over longer periods could be used to assess dynamic changes |
| Statistics | Statistical methods and reporting of results vary tremendously, making comparisons difficult | Standardisation of reporting. All studies should consider reporting associations with a 1 SD change in a novel parameter, if appropriate |
| Prediction vs association | Lack of studies examining extent of prediction afforded by novel risk factors beyond known predictors or simple algorithms to detect individuals at high diabetes risk | More use of ROC analyses, C statistics or similar statistical methods to place results into meaningful clinical context. There is no guarantee that novel risk factors will improve risk prediction |
| Differing populations | Associations of risk factors with incident diabetes will vary according to age, sex and ethnicity—more research needed | More studies in different population subgroups needed |
| Diabetes vs vascular risk markers | Risk factors for diabetes may not necessarily predict vascular events, or do so with equal strength | Further research comparing and contrasting risk factors for diabetes and vascular events is needed |
| Causal inferences | Association studies in isolation cannot prove causality, regardless of their robustness or comprehensiveness | Appraisal against criteria for causality, and combination of biomarkers with genetic polymorphisms (Mendelian randomisation) and, where possible, interventions with specific agents, are needed to advance claims of causality |

0.73 to 0.77 for fasting or post-load glucose measures [10]. Combinations of obesity and easily available biochemical/clinical measures have reasonable predictive ability for diabetes, perhaps better than an OGTT. In a Swedish study, a combination of $HbA_{1c}$, fasting glucose and BMI achieved a specificity of between 93% and 97% and a sensitivity of between 52% and 66%, with comparable results on addition of a positive family history, OGTT or triacylglycerol [11]. Similarly, a multivariable model with readily available clinical variables (age, sex, BMI, ethnicity, family history, systolic blood pressure, fasting glucose and HDL-cholesterol) achieved a greater AUC (0.84) than the 2 h glucose value alone (AUC 0.78), while addition of 2 h glucose measurement to the clinical model increased the AUC modestly from 0.84 to 0.86, thus not justifying its greater cost and inconvenience [12]. More recently, a simple clinical model including family history of diabetes, obesity, blood pressure, lipids and impaired fasting glucose produced an AUC of 0.85 for prediction of incident diabetes in the Framingham Offspring Study [13]. Notably, more complex clinical models that included OGTT, fasting insulin, and C-reactive protein levels or homeostasis model assessment (HOMA) indexes of insulin sensitivity and beta cell sensitivity did not improve this AUC [13].

*Use of risk scores or questionnaires for diabetes risk prediction* As a further development, the concept of ‘risk scores’, calculated using routinely available or easily collectable data, has emerged as an appealing tool for predicting both undiagnosed prevalent diabetes and the risk of future incident diabetes. The Cambridge Diabetes Risk Score uses general practice record data (age, sex, BMI, history of antihypertensive or steroid medication, family history and smoking history) to give a reasonable prediction of prevalent undiagnosed diabetes (AUC 0.80) [14]. Using categorical variables for age, BMI, waist circumference, history of antihypertensive drugs or high blood glucose, physical activity and daily consumption of fruit/berries/vegetables, the Finnish Diabetes Risk Score achieved an AUC of between 0.85 and 0.87 for 10 year incident diabetes [15]. The recent German Diabetes Risk Score, comprising age, height, waist circumference, history of hypertension, physical activity, smoking and dietary factors,