# Assessing the Combined Impact of 18 Common Genetic Variants of Modest Effect Sizes on Type 2 Diabetes Risk

**Hana Lango,[1,2] the U.K. Type 2 Diabetes Genetics Consortium, Colin N.A. Palmer,[3] Andrew D. Morris,[4] Eleftheria Zeggini,[5] Andrew T. Hattersley,[1,2] Mark I. McCarthy,[5,6] Timothy M. Frayling,[1,2] and Michael N. Weedon[1,2]**

**OBJECTIVES—**Genome-wide association studies have dramatically increased the number of common genetic variants that are robustly associated with type 2 diabetes. A possible clinical use of this information is to identify individuals at high risk of developing the disease, so that preventative measures may be more effectively targeted. Here, we assess the ability of 18 confirmed type 2 diabetes variants to differentiate between type 2 diabetic case and control subjects.

**RESEARCH DESIGN AND METHODS—**We assessed index single nucleotide polymorphisms (SNPs) for the 18 independent loci in 2,598 control subjects and 2,309 case subjects from the Genetics of Diabetes Audit and Research Tayside Study. The discriminatory ability of the combined SNP information was assessed by grouping individuals based on number of risk alleles carried and determining relative odds of type 2 diabetes and by calculating the area under the receiver-operator characteristic curve (AUC).

**RESULTS—**Individuals carrying more risk alleles had a higher risk of type 2 diabetes. For example, 1.2% of individuals with >24 risk alleles had an odds ratio of 4.2 (95% CI 2.11–8.56) against the 1.8% with 10–12 risk alleles. The AUC (a measure of discriminative accuracy) for these variants was 0.60. The AUC for age, BMI, and sex was 0.78, and adding the genetic risk variants only marginally increased this to 0.80.

**CONCLUSIONS—**Currently, common risk variants for type 2 diabetes do not provide strong predictive value at a population level. However, the joint effect of risk variants identified subgroups of the population at substantially different risk of disease. Further studies are needed to assess whether individuals with extreme numbers of risk alleles may benefit from genetic testing.
***Diabetes* 57:3129–3135, 2008**

From the [1]Genetics of Complex Traits, Institute of Biomedical and Clinical Science, Peninsula Medical School, Exeter, U.K.; [2]Diabetes Genetics, Institute of Biomedical and Clinical Science, Peninsula Medical School, Exeter, U.K.; the [3]Population Pharmacogenetics Group, Biomedical Research Centre, Ninewells Hospital and Medical School, University of Dundee, Dundee, U.K.; the [4]Diabetes Research Group, Division of Medicine and Therapeutics, Ninewells Hospital and Medical School, University of Dundee, Dundee, U.K.; the [5]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, U.K.; and the [6]Oxford Centre for Diabetes, Endocrinology and Medicine, University of Oxford, Churchill Hospital, Oxford, U.K.

Corresponding author: Michael Weedon, michael.weedon@pms.ac.uk.

Received 16 April 2008 and accepted 24 June 2008.

Published ahead of print at http://diabetes.diabetesjournals.org on 30 June 2008. DOI: 10.2337/db08-0504.



Recent genome-wide association (GWA) studies, which assay >300,000 single nucleotide polymorphisms (SNPs) across many thousands of individuals, have led to the discoveries of variants predisposing to many common complex diseases, including type 2 diabetes (1–6), coronary artery disease (7–9), prostate cancer (10,11), Crohn's disease (12–14), and many others (see http://www.genome.gov/26525384 for an up to date list of all GWA studies). The variants identified by these GWA studies are common in the general population (minor allele frequency >1%), but most have, individually, only small effects on disease risk, with odds ratios (ORs) typically <1.3.

Despite the relatively small predisposing effects conferred, these variants provide important, novel insights into disease biology. For example, variants of a number of genes, such as *HHEX*, *CDKN2A/B*, and *CDKAL1*, implicate defects in pancreatic β-cell development and function as important in type 2 diabetes etiology (4,15,16), whereas the discovery that variants in *FTO* are associated with BMI opened up novel areas of investigation for obesity biology (17–19). By gaining further knowledge of the underlying biology, and promoting potential therapeutic and preventative approaches, these insights are likely to be the most important outcome from these GWA studies.

A more immediate clinical utility may be to use the identified risk variants to aid the determination of an individual's risk of developing a particular disease. Several companies, such as deCODE genetics and 23andme, have begun to use SNPs identified from these GWA studies, offering up to 1 million SNP GWA scans (http://www.decodeme.com; https://www.23andme.com) or individual disease-associated SNP tests (http://www.decodediagnostics.com). It is, however, unclear how useful the currently identified variants will be in predicting disease.

One of the disease traits for which the GWA approach has been most successful is type 2 diabetes. Together with candidate gene approaches, 18 common variants, including *FTO* and two independent signals in the *CDKN2A/B* region, have now been convincingly shown to associate with the disease (1–6,20–26). In this study, we aimed to assess the combined discriminatory power of these common, modest effect variants, using >4,900 individuals from the Genetics of Diabetes Audit and Research Tayside Study (GoDARTS).

## RESEARCH DESIGN AND METHODS

**SNP selection and genotyping.** We only included variants that have been convincingly shown to associate with type 2 diabetes. We used variants reviewed by Frayling (27) and those described by Zeggini et al. (5,6), except for the E23K (rs5219; $r^2$ with GWA-SNP rs5215 = 0.89) variant of *KCNJ11* (22)

TABLE 1
Characteristics of study participants

| Variable | Case subjects | Control subjects |
|---|---|---|
| *n* | 2,309 | 2,598 |
| Men (%) | 56 | 51 |
| Age at diagnosis (years) | 55.7 ± 9.0 | NA |
| BMI (kg/m$^2$) | 31.5 ± 6.1 | 26.9 ± 4.5 |
| A1C | 7.8 ± 1.5 | 5.5 ± 0.3 |

Data are *n*, percent, and means ± SD.

and rs7903146 ($r^2$ with GWA-SNP rs7901695 = 0.80) of *TCF7L2* (23,28), where we genotyped a SNP shown to have stronger association with type 2 diabetes, but which were not genotyped on the genome-wide association chips; the *TCF2* locus, where we used rs757210 (26), instead of rs4430796 (24) ($r^2$ = 0.61); and the *ADAM30/NOTCH2* locus, where we used rs2641348 in *ADAM30* as a proxy for rs2934381 ($r^2$ = 0.92).

Genotyping was performed by KBioscience (Hertsfordshire, U.K.), which designed and used assays based on either their proprietary competitive allele-specific PCR (KASPar) method or a modified TaqMan-based assay, details of which are available on the company website (www.kbioscience.co.uk/chemistry/index.htm). Genotyping quality control measures for the SNPs are as described previously (5,6,25).

**GoDARTS study and participants.** The GoDARTS study is a substudy of the Diabetes Audit and Research Tayside Study (DARTS) (29), which aims to identify all known diabetes patients in the Tayside region of Scotland using electronic database retrieval. The samples used in this study are a subsample of the type 2 diabetes patients identified and have been described in detail previously (6). Briefly, the GoDARTS study includes individuals of white European descent, living in the Tayside region when recruited. The diagnosis of diabetes in case subjects was based on either current treatment with diabetes-specific medication or laboratory evidence of hyperglycemia if treated with diet alone. Patients with confirmed diagnosis of monogenic diabetes and those treated with regular insulin therapy within 1 year of diagnosis were excluded. Case subjects in this study had an age at diagnosis between 35 and 70 years, inclusive. Control subjects had not been diagnosed with diabetes at the time of recruitment or subsequently and were excluded if there was evidence of hyperglycemia during recruitment (fasting glucose >7.0 mmol/l, A1C >6.4%) or if they were >80 years old. The study was approved by the Tayside Medical Ethics Committee. Informed consent was obtained from all study participants. Table 1 presents the clinical characteristics of subjects used in this study.

**Statistical analysis.** All statistical analyses were performed in StataSE v10.0 for Windows (StataCorp, Brownsville, TX). We used logistic regression for all individual SNP analyses. To test for deviation from a within-loci additive model, we performed likelihood ratio test of an additive model against a general 2 degrees of freedom model. To test for gene-gene interaction across all pairs of loci, we used likelihood ratio tests to compare an additive model to a model with an interaction term. We combined information from multiple SNPs by using an allele count model, where we summed the number of risk alleles carried by each individual. This assumes that each of the alleles has an equal and additive effect on type 2 diabetes risk.

We used logistic regression on the general model (i.e., individual SNP genotypes as indicator variables) to construct the receiver-operating characteristic (ROC) curves and calculate the areas under the curve (AUCs). We also performed these ROC analyses on the allele count model for comparison with the general model.

## RESULTS

Genotyping data on all of the variants were available for 2,309 type 2 diabetic case subjects and 2,598 control subjects. Characteristics of these participants are shown in Table 1. Supplementary Table 1, available in an online appendix at http://dx.doi.org/10.2337/db08-0504, presents a comparison of clinical characteristics for these subjects against the 1,739 who were not successfully genotyped across all SNPs. Individually, the variants have similar effect sizes in this study compared with those reported in other large studies (Table 2) (1–6,20–26), and the range of ORs from 1.00 to 1.36 most likely reflects stochastic variation. Several variants are not associated at $P < 0.05$ in the sample used here but are still included in the analyses because they are confirmed type 2 diabetes risk variants, and the lack of significance is the result of relatively low power in this number of subjects. Based on these and larger datasets, all of the variants appear to have an additive mode of inheritance (1–6,20–26). The *CDKAL1* locus was reported by Steinthorsdottir et al. (4) to fit a recessive model, but other large studies do not support this. There is no evidence of interaction between any of the SNPs based on these data (supplementary Table 2) or on the larger analyses previously published. Therefore, we assumed an additive genetic model. We found no evidence

TABLE 2
Summary of type 2 diabetes variants in 2,598 control subjects and 2,309 case subjects from the Dundee cohort

| SNP | Gene/region | Risk allele frequency | Additive model test *P* value | OR (95% CI) | *P* value |
|---|---|---|---|---|---|
| rs7903146 | *TCF7L2* | 0.30 | 0.70 | 1.36 (1.24–1.48) | $3.97 \times 10^{-12}$ |
| rs5219 | *KCNJ11* | 0.36 | 0.058 | 1.25 (1.15–1.36) | $8.54 \times 10^{-8}$ |
| rs10811661 | *CDKN2A/2B* | 0.85 | 0.24 | 1.21 (1.08–1.35) | $8.82 \times 10^{-4}$ |
| rs1801282 | *PPARG* | 0.87 | 0.46 | 1.21 (1.07–1.36) | $2.18 \times 10^{-3}$ |
| rs2641348* | *ADAM30/NOTCH2* | 0.11 | 0.68 | 1.15 (1.01–1.30) | $3.20 \times 10^{-2}$ |
| rs564398 | *CDKN2A/2B* | 0.59 | 0.95 | 1.13 (1.04–1.22) | $3.61 \times 10^{-3}$ |
| rs4402960 | *IGF2BP2* | 0.33 | 0.76 | 1.12 (1.03–1.22) | $7.62 \times 10^{-3}$ |
| rs8050136 | *FTO* | 0.41 | 0.32 | 1.11 (1.02–1.20) | $1.43 \times 10^{-2}$ |
| rs10946398 | *CDKAL1* | 0.34 | 0.19 | 1.11 (1.02–1.21) | $1.47 \times 10^{-2}$ |
| rs13266634 | *SLC30A8* | 0.70 | 0.60 | 1.10 (1.01–1.20) | $2.57 \times 10^{-2}$ |
| rs7961581 | *TSPAN8/LGR5* | 0.29 | 0.87 | 1.09 (1.00–1.19) | $5.56 \times 10^{-2}$ |
| rs12779790 | *CDC123* | 0.20 | 0.15 | 1.10 (0.99–1.21) | $7.58 \times 10^{-2}$ |
| rs10010131 | *WFS1* | 0.60 | 0.54 | 1.07 (0.99–1.16) | $9.19 \times 10^{-2}$ |
| rs757210 | *TCF2* | 0.37 | 0.18 | 1.07 (0.99–1.16) | $1.09 \times 10^{-1}$ |
| rs4607103 | *ADAMTS9* | 0.77 | 0.60 | 1.05 (0.96–1.16) | $2.89 \times 10^{-1}$ |
| rs1111875 | *HHEX-IDE* | 0.62 | 0.19 | 1.02 (0.94–1.11) | $5.98 \times 10^{-1}$ |
| rs7578597 | *THADA* | 0.91 | 0.33 | 1.04 (0.90–1.19) | $6.07 \times 10^{-1}$ |
| rs864745 | *JAZF1* | 0.50 | 0.50 | 1.00 (0.93–1.09) | $9.70 \times 10^{-1}$ |

Only samples that were successfully genotyped for all 18 variants are included. Additive model test *P* value refers to a test of deviation from additivity of alleles at each SNP. *This SNP falls within the *ADAM30* gene and is a proxy ($r^2$ = 0.92 in HapMap CEU) for rs2934381 in the *NOTCH2* gene, which showed stronger association previously (5).

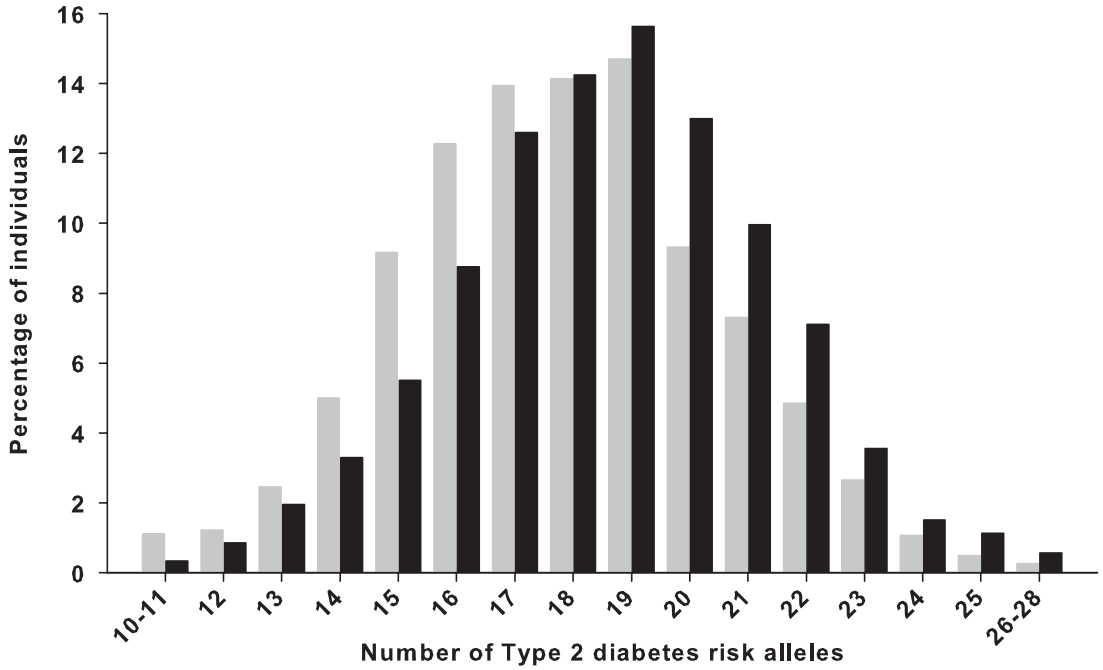


**FIG. 1. Distribution of risk alleles in type 2 diabetic case subjects (black bars) and control subjects (gray bars).**

of any interaction between the individual variants and BMI or age (lowest interaction $P$ values = 0.14 and 0.02, respectively). We performed the analysis with and without the *FTO* variant, the one variant shown to predispose to type 2 diabetes through a primary effect on BMI (18).

**Comparing extremes.** The proportion of case and control subjects grouped according to the number of risk alleles that they carry is shown in Fig. 1. The distribution of risk alleles follows a normal distribution in both case and control subjects, with a shift toward a higher number of risk alleles in the case subjects. There is an increase in ORs for type 2 diabetes with the increasing number of risk alleles against the baseline group of 1.8% of individuals carrying 10–12 risk alleles. Of individuals with ≥25 risk alleles, 1.2% have an OR of 4.2 (95% CI 2.11–8.56) against the baseline reference group. Similarly, 11.5% of this study population carrying ≥22 risk alleles had an OR of 2.3 (1.73–2.93) for type 2 diabetes compared with the 8.2% of individuals with ≤14 risk alleles.

Figure 2 plots the ORs relative to the median number of 18 risk alleles. Those with ≥25 risk alleles were more than twice as likely to have type 2 diabetes (OR 2.18 [95% CI 1.24–3.81]) compared with those with the median number of risk alleles. The *TCF7L2* variant had a stronger effect than the other variants (OR 1.36 compared with 1.00–1.25 for the rest), so these results may be slight underestimates, because the additive model used for the allele counting assumes equal effects across all SNPs.

We performed the same analyses for two subgroups of the cohort, one including only obese individuals (with BMI of ≥30 kg/m$^2$, $n$ = 1,803), the other nonobese individuals (BMI <30 kg/m$^2$, $n$ = 3,083). The results were similar across these subgroups. For example, the 1.4% of obese individuals with >24 risk alleles had an OR of 5.5 (95% CI 2.11–14.36) compared with the 1.9% of obese individuals with <13 risk alleles. The corresponding OR for the nonobese subjects was 3.31 (1.34–8.16), for the 1.8 and 1.1% of individuals with <13 and >24 risk alleles, respectively.

**ROC curve.** We evaluated the discriminatory power of a genetic test based on the 18 type 2 diabetes variants by calculating the area under the ROC curve. Using the general model (as opposed to the additive model, which assumes equal and additive effects), the ROC curve for the 18 type 2 diabetes variants studied here is 0.60 (Fig. 3). We performed the same analysis for the obese and nonobese subgroups of the cohort. The AUCs for the obese and nonobese groups were 0.58 and 0.60, respectively. A similar result was obtained when we removed the *FTO* variant (obese, 0.58; nonobese, 0.59). We also tested whether the risk variants would add to the discriminatory power of BMI, age, and sex alone (AUC 0.78 in our study). A model that includes BMI, age, sex, and the 18 variants has an AUC of 0.80 (Fig. 3); although marginal, the increase in the AUC was statistically significant ($P$ = $2.88 \times 10^{-12}$). The AUC remained virtually the same (AUC = 0.80) when the *FTO* variant was removed from the model.

**The effect of BMI and age.** Supplementary Table 3 presents the individual SNP type 2 diabetes associations adjusted for BMI. As expected, the *FTO* association is weakened on adjusting for BMI (OR 1.00 [95% CI 0.92–1.10]), and the *TCF7L2* association is strengthened (1.46 [1.32–1.61]). Testing the combined effect of the risk variants on clinical features of the type 2 diabetes patients, we found that the number of risk alleles was associated with an earlier age at diagnosis of 0.15 years per risk allele (95% CI −0.29 to −0.01, $P$ = 0.038). We also observed an overall modifying effect on BMI (−0.14 BMI units per risk allele [−0.23 to −0.05], $P = 3.41 \times 10^{-3}$), but this finding is mainly explained by the known association of the *TCF7L2* variant alone with BMI in type 2 diabetic case subjects (30,31). Here, each *TCF7L2* risk allele was associated with a difference in BMI of −0.69 kg/m$^2$ (−1.06 to −0.31, $P = 3.18 \times 10^{-4}$), whereas the combined effect of all other variants without *TCF7L2* could just be detected (−0.10 kg/m$^2$ per risk allele [−0.20 to 0.01], $P$ = 0.036). The difference in BMI and age at diagnosis was more noticeable when we compared individuals with low and high numbers of risk alleles. For example, carriers of ≥23 risk alleles (11.8%) were, on average, diagnosed 4.2 years

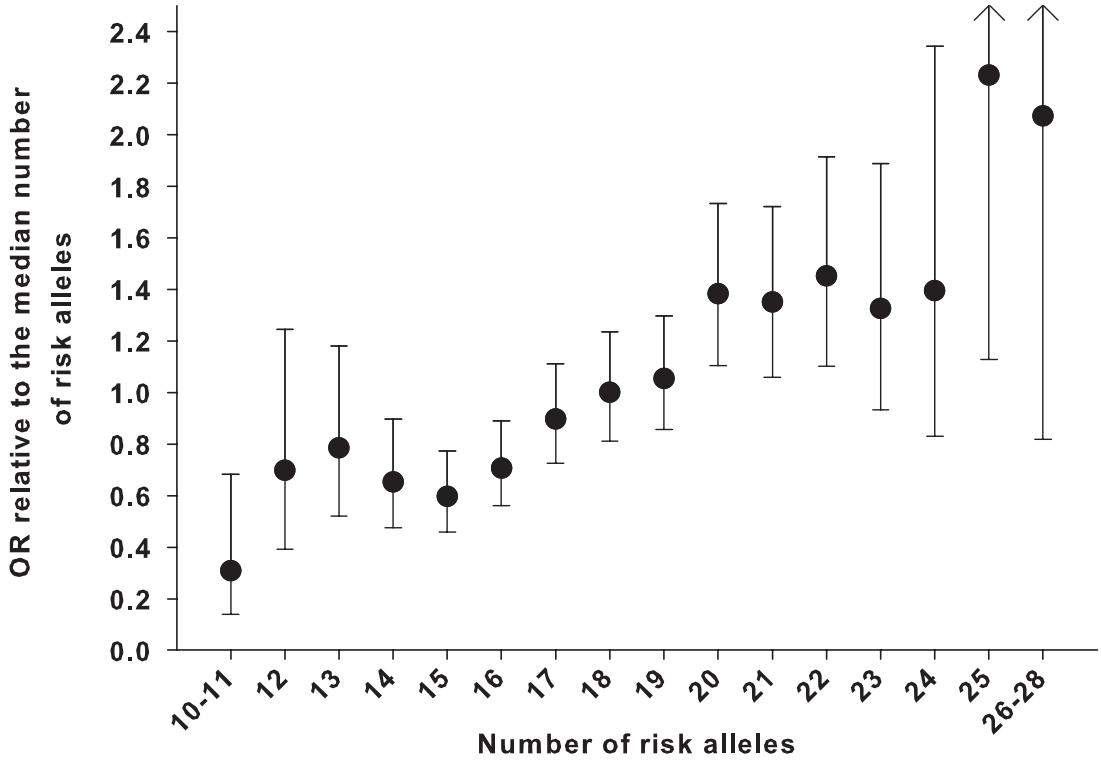


**FIG. 2. A plot showing the increasing ORs with the increasing number of type 2 diabetes risk alleles versus the baseline of 10–11 risk alleles. The ORs are given relative to the median number of 18 risk alleles (●). The vertical bars represent 95% CIs.**

earlier (−6.45 to −1.87, $P = 4.21 \times 10^{-4}$) and had 1.60 kg/m$^2$ lower BMI (−3.35 to 0.08, $P = 0.062$) than those carrying <15 (8.6%) risk alleles.

## DISCUSSION

Recent success in identifying common variants predisposing to type 2 diabetes has led to suggestions that they may be useful in predicting an individual's risk of the disease. In this study, we evaluated the ability of 18 confirmed predisposing variants to discriminate between individuals with and without type 2 diabetes, using the GoDARTS study. The samples used in this study were not enriched for family history or low BMI, factors that may inflate effect sizes. Although the GoDARTS cohort was a part of the Wellcome Trust Case Control Consortium-Type 2 Diabetes GWA Study (5,6), it was only used as a stage 2 replication set for the follow-up of the initial hits. This means that there should be a minimal effect of the "winner's curse" (32), the upward bias of the effect size in the discovery samples compared with subsequent replication studies.

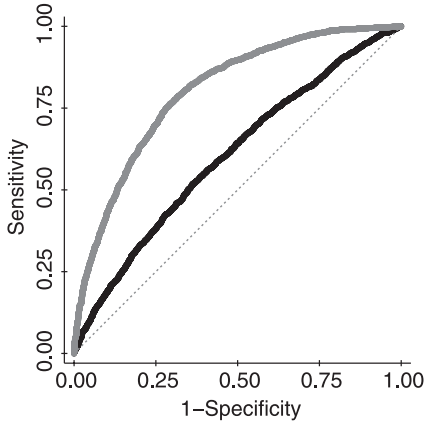


**FIG. 3. ROC plot for a model containing all type 2 diabetes variants, BMI, age, and sex (gray line, AUC = 0.80) and for the 18 variants alone (black line, AUC = 0.60).**

**The combined information identifies individuals at different risks of disease.** By comparing individuals with the fewest type 2 diabetes risk alleles with those carrying the most risk alleles, combining genetic information allowed us to identify subgroups of the population at a distinctly differing risk of disease. For example, we were able to distinguish ~1% of the population carrying >25 risk alleles that had more than four times increased risk of diabetes compared with the 2% with 10–12 risk alleles. The high-risk group also had over twice the odds for type 2 diabetes than those with the median number of risk alleles. These figures were similar in individuals who were obese and not obese, a major risk factor for type 2 diabetes and easily measurable. Obese individuals carrying large numbers of type 2 diabetes risk alleles may therefore be a particular group worth studying to test potential intervention strategies. This may be important given that the escalating rates of obesity and type 2 diabetes suggest that efforts aimed at the whole population are not effective and that intensive, but expensive, lifestyle interventions aimed at increasing exercise and improving diet can result in weight loss and a reduced risk of type 2 diabetes (33–36).

**The current variants are not particularly discriminative but explain only a small amount of the heritability of type 2 diabetes.** Rather than focusing on individuals with "extreme" numbers of risk alleles, at a population level, the utility of genetic tests may be better classified by ROC curves. One of the most important factors in the validity of a genetic test in clinical practice is its ability to discriminate between individuals who will and will not develop the disease. A clinically relevant AUC threshold clearly depends on a whole range of factors (for

example, the cost of the test and the availability of preventative measures), but as an example from current clinical practice, oxidized-LDL cholesterol has an AUC of ~0.80 for coronary artery disease (37), making it a good discriminator between patients and healthy control subjects. The 18 type 2 diabetes variants had an inadequate discriminatory ability with an AUC of 0.60, a slight improvement on the AUC of 0.55 based on *TCF7L2* alone. These data imply that genetic tests for type 2 diabetes (and many other complex diseases) that are offered by several commercial companies currently have limited predictive value. However, there are many more variants to be identified, because these 18 variants only explain a small amount of the heritability of type 2 diabetes: the sibling relative risk for type 2 diabetes is ~3 (38), and the combination of these variants would only account for a sibling relative risk of ~1.07. As more susceptibility variants are found for type 2 diabetes, genetic testing that uses the inexpensive and rapid genotyping technologies may eventually become more clinically useful.

**The use of genetic information in addition to age, sex, and BMI.** For many complex diseases, there are already well-established risk factors that can be used to predict someone's chances of developing the disease. Incorporating genetic information may be justified on the basis that current preventative measures are expensive and that prevention at a population level is not effective, so the more selective we can be the better. In type 2 diabetes, family history, age, BMI, ethnicity, and lifestyle all contribute to an individual's risk of the disease. In our study, the AUC for BMI, age, and sex (we did not have family history data) combined was 0.78, a moderate diagnostic value. The genetic risk variants had a poor discriminatory ability alone (AUC = 0.60) and only marginally increased the discriminatory power of the test when combined with BMI, age, and sex (AUC = 0.80), suggesting that they add little to the already known predictive factors.

**Risk variants modify clinical characteristics of individuals with type 2 diabetes.** Type 2 diabetes often occurs in individuals who are not overweight or obese, and can be diagnosed at a relatively young age. This may be because these individuals have a stronger genetic risk component than more "typical" type 2 diabetes patients. Therefore, we tested the extent to which patients with the stronger genetic predisposition tended to be leaner, and how much younger they were at diagnosis. There were notable differences between the 11.8 and 8.6% of the population carrying either high or low numbers of disease-predisposing alleles, respectively. Patients with high genetic risk had an average BMI of 30.3 kg/m$^2$ compared with 31.9 kg/m$^2$ in those with low genetic risk and were diagnosed at an average age of 55.2 years, compared with 59.3 years for patients with relatively low genetic risk. These results support an important role for genetic predisposition to type 2 diabetes in nonobese, young-onset case subjects.

**Weighting variants and the optimal ROC curve.** The simple allele count model we used for some of our analyses of "extremes" assumes that each risk allele has the same effect size and that the effects are additive both within and between loci. Although we found no strong evidence for deviation from additivity, clearly some SNPs have stronger effects than others. This is most evident for *TCF7L2*, where the allelic OR is 1.37, significantly larger than any of the other variants. One way to overcome this is to weigh SNPs differently; however, we decided not to do this in this study for a number of reasons. First, all of our AUC analyses are based on a general model, in which the assumption of equal effects is not made. Second, as Janssens et al. (39) previously showed, when the ORs of the individual variants are relatively low (as here), there is little difference in the discriminative accuracy of the test based on the simple allele count model and a model that allows each variant to have a different effect size (the AUCs here are 0.583 and 0.603, respectively, although this was statistically significant [$P = 0.001$]). Third, it is unclear what the most appropriate weights to use would be. Fourth, an allele count model provides important advantages for simplicity and visualization of the results.

Recently, Lu and Elston (40) proposed using an optimal ROC analysis approach rather than the standard approach that we have used. Although the authors proved theoretically that their method is more powerful, the results presented by Lu and Elston (40) showed that the two methods produce the same results when there are few loci and no interactive effects. Because we still have only a relatively few loci, there is no evidence of any nonadditive effects within or between loci, and the ROC curve is concave (40), the two methods should produce the same results. We tested this using the 10 SNPs that were significant (at $P < 0.05$) in our study. Using these variants, the results were the same for both methods (AUC for the Lu and Elston method, 0.596; AUC for the standard method, 0.596).

**Strengths and limitations of our study.** Our study was relatively large in terms of the number of samples, and the number of common variants used. We had >2,000 case subjects and >2,000 control subjects after excluding individuals who were not successfully genotyped for all of the variants included in the study. The 18 variants we used had all been convincingly shown in previous studies to associate with type 2 diabetes.

One of the main limitations of our study is that it was not prospective, and therefore, we are unable to truly determine the predictive power of these variants. Although the results of this study only apply to the Tayside population, it is likely, based on previous data (41–43), that our prediction estimates are reasonably accurate and that the effect sizes observed are likely to be representative of those in similar populations. A second limitation is that although the results are applicable to the Tayside and similar populations, they may not apply to populations of substantially different ethnic origin or those exposed to different social and environmental circumstances. A third limitation concerns the caveat that the majority of the type 2 diabetes–associated SNPs identified to date and used in this study are not the causal variants. This means that the predictive power of these susceptibility loci is likely to be an underestimate. Fine mapping and sequencing approaches are needed to identify the variants causal to these associations, which often have stronger effects than the currently identified variants. These follow-up studies may also reveal additional causal variants at these loci that cannot be detected by GWA methods because of, for example, low frequency, but that may have higher penetrance and therefore would be much more powerful predictors.

In conclusion, the combined information from the currently known susceptibility variants allows us to identify subgroups of the population at substantially increased odds of getting type 2 diabetes. These individuals could be targeted with more effective preventative measures. On a

population level, these variants appear to be of limited use in discriminating between individuals who will and will not develop type 2 diabetes. As more variants are identified, tests with better predictive performance should become available and could eventually become a valuable addition to clinical practice.

## ACKNOWLEDGMENTS

M.N.W. is a Vandervell Foundation Research Fellow. E.Z. is a Wellcome Trust Research Career Development Fellow. This work has received funding from Wellcome Trust.

We thank all study participants.

## REFERENCES

1. Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PIW, Chen H, Roix JJ, Kathiresan S, Hirschhorn JN, Daly MJ, Hughes TE, Groop L, Altshuler D, Almgren P, Florez JC, Meyer J, Ardlie K, Bostrom KB, Isomaa B, Lettre G, Lindblad U, Lyon HN, Melander O, Newton-Cheh C, Nilsson P, Orho-Melander M, Rastam L, Speliotes EK, Taskinen MR, Tuomi T, Guiducci C, Berglund A, Carlson J, Gianniny L, Hackett R, Hall L, Holmkvist J, Laurila E, Sjogren M, Sterner M, Surti A, Svensson M, Svensson M, Tewhey R, Blumenstiel B, Parkin M, DeFelice M, Barry R, Brodeur W, Camarata J, Chia N, Fava M, Gibbons J, Handsaker B, Healy C, Nguyen K, Gates C, Sougnez C, Gage D, Nizzari M, Gabriel SB, Chirn GW, Ma QC, Parikh H, Richardson D, Ricke D, Purcell S: Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 316:1331–1336, 2007
2. Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, Erdos MR, Stringham HM, Chines PS, Jackson AU, Prokunina-Olsson L, Ding CJ, Swift AJ, Narisu N, Hu T, Pruim R, Xiao R, Li XY, Conneely KN, Riebow NL, Sprau AG, Tong M, White PP, Hetrick KN, Barnhart MW, Bark CW, Goldstein JL, Watkins L, Xiang F, Saramies J, Buchanan TA, Watanabe RM, Valle TT, Kinnunen L, Abecasis GR, Pugh EW, Doheny KF, Bergman RN, Tuomilehto J, Collins FS, Boehnke M: A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* 316:1341–1345, 2007
3. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, Boutin P, Vincent D, Belisle A, Hadjadj S, Balkau B, Heude B, Charpentier G, Hudson TJ, Montpetit A, Pshezhetsky AV, Prentki M, Posner BI, Balding DJ, Meyre D, Polychronakos C, Froguel P: A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 445:881–885, 2007
4. Steinthorsdottir V, Thorleifsson G, Reynisdottir I, Benediktsson R, Jonsdottir T, Walters GB, Styrkarsdottir U, Gretarsdottir S, Emilsson V, Ghosh S, Baker A, Snorradottir S, Bjarnason H, Ng MCY, Hansen T, Bagger Y, Wilensky RL, Reilly MP, Adeyemo A, Chen YX, Zhou J, Gudnason V, Chen GJ, Huang HX, Lashley K, Doumatey A, So WY, Ma RCY, Andersen G, Borch-Johnsen K, Jorgensen T, van Vliet-Ostaptchouk JV, Hofker MH, Wijmenga C, Christiansen C, Rader DJ, Rotimi C, Gurney M, Chan JCN, Pedersen O, Sigurdsson G, Gulcher JR, Thorsteinsdottir U, Kong A, Stefansson K: A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nat Genet* 39:770–775, 2007
5. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, de Bakker PIW, Abecasis GR, Almgren P, Andersen G, Ardlie K, Bostrom KB, Bergman RN, Bonnycastle LL, Borch-Johnsen K, Burtt NP, Chen H, Chines PS, Daly MJ, Deodhar P, Ding C-J, Doney ASF, Duren WL, Elliott KS, Erdos MR, Frayling TM, Freathy RM, Gianniny L, Grallert H, Grarup N, Groves CJ, Guiducci C, Hansen T, Herder C, Hitman GA, Hughes TE, Isomaa B, Jackson AU, Jorgensen T, Kong A, Kubalanza K, Kuruvilla FG, Kuusisto J, Langenberg C, Lango H, Lauritzen T, Li Y, Lindgren CM, Lyssenko V, Marvelle AF, Meisinger C, Midthjell K, Mohlke KL, Morken MA, Morris AD, Narisu N, Nilsson P, Owen KR, Palmer CNA, Payne F, Perry JRB, Pettersen E, Platou C, Prokopenko I, Qi L, Qin L, Rayner NW, Rees M, Roix JJ, Sandbaek A, Shields B, Sjogren M, Steinthorsdottir V, Stringham HM, Swift AJ, Thorleifsson G, Thorsteinsdottir U, Timpson NJ, Tuomi T, Tuomilehto J, Walker M, Watanabe RM, Weedon MN, Willer CJ, Illig T, Hveem K, Hu FB, Laakso M, Stefansson K, Pedersen O, Wareham NJ, Barroso I, Hattersley AT, Collins FS, Groop L, McCarthy MI, Boehnke M, Altshuler D: Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet* 40:638–645, 2008
6. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, Timpson NJ, Perry JRB, Rayner NW, Freathy RM, Barrett JC, Shields B, Morris AP, Ellard S, Groves CJ, Harries LW, Marchini JL, Owen KR, Knight B, Cardon LR, Walker M, Hitman GA, Morris AD, Doney ASF, McCarthy MI, Hattersley AT: Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* 316:1336–1341, 2007
7. Helgadottir A, Thorleifsson G, Manolescu A, Gretarsdottir S, Blondal T, Jonasdottir A, Jonasdottir A, Sigurdsson A, Baker A, Palsson A, Masson G, Gudbjartsson DF, Magnusson KP, Andersen K, Levey AI, Backman VM, Matthiasdottir S, Jonsdottir T, Palsson S, Einarsdottir H, Gunnarsdottir S, Gylfason A, Vaccarino V, Hooper WC, Reilly MP, Granger CB, Austin H, Rader DJ, Shah SH, Quyyumi AA, Gulcher JR, Thorgeirsson G, Thorsteinsdottir U, Kong A, Stefansson K: A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science* 316:1491–1493, 2007
8. McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, Cox DR, Hinds DA, Pennacchio LA, Tybjaerg-Hansen A, Folsom AR, Boerwinkle E, Hobbs HH, Cohen JC: A common allele on chromosome 9 associated with coronary heart disease. *Science* 316:1488–1491, 2007
9. Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, Dixon RJ, Meitinger T, Braund P, Wichmann HE, Barrett JH, Konig IR, Stevens SE, Szymczak S, Tregouet D, Iles MM, Pahlke F, Pollard H, Lieb W, Cambien F, Fischer M, Ouwehand W, Blankenberg S, Balmforth AJ, Baessler A, Ball SG, Strom TM, Braenne I, Gieger C, Deloukas P, Tobin MD, Ziegler A, Thompson JR, Schunkert H: Genomewide association analysis of coronary artery disease. *New Engl J Med* 357:443–453, 2007
10. Gudmundsson J, Sulem P, Manolescu A, Amundadottir LT, Gudbjartsson D, Helgason A, Rafnar T, Bergthorsson JT, Agnarsson BA, Baker A, Sigurdsson A, Benediktsdottir KR, Jakobsdottir M, Xu JF, Blondal T, Kostic J, Sun JL, Ghosh S, Stacey SN, Mouy M, Saemundsdottir J, Backman VM, Kristjansson K, Tres A, Partin AW, Albers-Akkers MT, Marcos JGI, Walsh PC, Swinkels DW, Navarrete S, Isaacs SD, Aben KK, Graif T, Cashy J, Ruiz-Echarri M, Wiley KE, Suarez BK, Witjes JA, Frigge M, Ober C, Jonsson E, Einarsson GV, Mayordomo JI, Kiemeney LA, Isaacs WB, Catalona WJ, Barkardottir RB, Gulcher JR, Thorsteinsdottir U, Kong A, Stefansson K: Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. *Nat Genet* 39:631–637, 2007
11. Yeager M, Orr N, Hayes RB, Jacobs KB, Kraft P, Wacholder S, Minichiello MJ, Fearnhead P, Yu K, Chatterjee N, Wang ZM, Welch R, Staats BJ, Calle EE, Feigelson HS, Thun MJ, Rodriguez C, Albanes D, Virtamo J, Weinstein S, Schumacher FR, Giovannucci E, Willett WC, Cancel-Tassin G, Cussenot O, Valeri A, Andriole GL, Gelmann EP, Tucker M, Gerhard DS, Fraumeni JF, Hoover R, Hunter DJ, Chanock SJ, Thomas G: Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. *Nat Genet* 39:645–649, 2007
12. Hampe J, Franke A, Rosenstiel P, Till A, Teuber M, Huse K, Albrecht M, Mayr G, De La Vega FM, Briggs J, Gunther S, Prescott NJ, Onnie CM, Hasler R, Sipos B, Folsch UR, Lengauer T, Platzer M, Mathew CG, Krawczak M, Schreiber S: A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. *Nat Genet* 39:207–211, 2007
13. Parkes M, Barrett JC, Prescott NJ, Tremelling M, Anderson CA, Fisher SA, Roberts RG, Nimmo ER, Cummings FR, Soars D, Drummond H, Lees CW, Khawaja SA, Bagnall R, Burke DA, Todhunter CE, Ahmad T, Onnie CM, McArdle W, Strachan D, Bethel G, Bryan C, Lewis CM, Deloukas P, Forbes A, Sanderson J, Jewell DP, Satsangi J, Mansfield JC, Cardon L, Mathew CG: Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nat Genet* 39:830–832, 2007
14. Rioux JD, Xavier RJ, Taylor KD, Silverberg MS, Goyette P, Huett A, Green T, Kuballa P, Barmada MM, Datta LW, Shugart YY, Griffiths AM, Targan SR, Ippoliti AF, Bernard EJ, Mei L, Nicolae DL, Regueiro M, Schumm LP, Steinhart AH, Rotter JI, Duerr RH, Cho JH, Daly MJ, Brant SR: Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nat Genet* 39:596–604, 2007
15. Grarup N, Rose CS, Andersson EA, Andersen G, Nielsen AL, Albrechtsen A, Clausen JO, Rasmussen SS, Jorgensen T, Sandbaek A, Lauritzen T, Schmitz O, Hansen T, Pedersen O: 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* 56:3105–3111, 2007
16. Pascoe L, Tura A, Patel SK, Ibrahim IM, Ferrannini E, Zeggini E, Weedon MN, Mari A, Hattersley AT, McCarthy MI, Frayling TM, Walker M: Common variants of the novel type 2 diabetes genes CDKAL1 and HHEX/IDE are associated with decreased pancreatic β-cell function. *Diabetes* 56:3101–3104, 2007
17. Dina C, Meyre D, Gallina S, Durand E, Korner A, Jacobson P, Carlsson LMS, Kiess W, Vatin V, Lecoeur C, Delplanque J, Vaillant E, Pattou F, Ruiz J, Weill J, Levy-Marchal C, Horber F, Potoczna N, Hercberg S, Le Stunff C, Bougneres P, Kovacs P, Marre M, Balkau B, Cauchi S, Chevre JC, Froguel

P: Variation in FTO contributes to childhood obesity and severe adult obesity. *Nat Genet* 39:724–726, 2007
18. Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, Perry JRB, Elliott KS, Lango H, Rayner NW, Shields B, Harries LW, Barrett JC, Ellard S, Groves CJ, Knight B, Patch AM, Ness AR, Ebrahim S, Lawlor DA, Ring SM, Ben-Shlomo Y, Jarvelin MR, Sovio U, Bennett AJ, Melzer D, Ferrucci L, Loos RJF, Barroso I, Wareham NJ, Karpe F, Owen KR, Cardon LR, Walker M, Hitman GA, Palmer CNA, Doney ASF, Morris AD, Smith GD, Hattersley AT, McCarthy MI: A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 316:889–894, 2007
19. Scuteri A, Sanna S, Chen WM, Uda M, Albai G, Strait J, Najjar S, Nagaraja R, Orru M, Usala G, Dei M, Lai S, Maschio A, Busonero F, Mulas A, Ehret GB, Fink AA, Weder AB, Cooper RS, Galan P, Chakravarti A, Schlessinger D, Cao A, Lakatta E, Abecasis GR: Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. *PLoS Genet* 3:1200–1210, 2007
20. Altshuler D, Hirschhorn JN, Klannemark M, Lindgren CM, Vohl M-C, Nemesh J, Lane CR, Schaffner F, Bolk S, Brewer C, Tuomi T, Gaudet D, Hudson TJ, Daly M, Groop L, Lander ES: The common PPARγ Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat Genet* 26:76–80, 2000
21. Florez JC, Burtt N, de Bakker PI, Almgren P, Tuomi T, Holmkvist J, Gaudet D, Hudson TJ, Schaffner SF, Daly MJ, Hirschhorn JN, Groop L, Altshuler D: Haplotype structure and genotype-phenotype correlations of the sulfonylurea receptor and the islet ATP-sensitive potassium channel gene region. *Diabetes* 53:1360–1368, 2004
22. Gloyn AL, Weedon MN, Owen KR, Turner MJ, Knight BA, Hitman G, Walker M, Levy JC, Sampson M, Halford S, McCarthy MI, Hattersley AT, Frayling TM: Large-scale association studies of variants in genes encoding the pancreatic β-cell KATP channel subunits Kir6.2 (KCNJ11) and SUR1 (ABCC8) confirm that the KCNJ11 E23K variant is associated with type 2 diabetes. *Diabetes* 52:568–572, 2003
23. Grant SFA, Thorleifsson G, Reynisdottir I, Benediktsson R, Manolescu A, Sainz J, Helgason A, Stefansson H, Emilsson V, Helgadottir A, Styrkarsdottir U, Magnusson KP, Walters GB, Palsdottir E, Jonsdottir T, Gudmundsdottir T, Gylfason A, Saemundsdottir J, Wilensky RL, Reilly MP, Rader DJ, Bagger Y, Christiansen C, Gudnason V, Sigurdsson G, Thorsteinsdottir U, Gulcher JR, Kong A, Stefansson K: Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet* 38:320–323, 2006
24. Gudmundsson J, Sulem P, Steinthorsdottir V, Bergthorsson JT, Thorleifsson G, Manolescu A, Rafnar T, Gudbjartsson D, Agnarsson BA, Baker A, Sigurdsson A, Benediktsdottir KR, Jakobsdottir M, Blondal T, Stacey SN, Helgason A, Gunnarsdottir S, Olafsdottir A, Kristinsson KT, Birgisdottir B, Ghosh S, Thorlacius S, Magnusdottir D, Stefansdottir G, Kristjansson K, Bagger Y, Wilensky RL, Reilly MP, Morris AD, Kimber CH, Adeyemo A, Chen Y, Zhou J, So WY, Tong PCY, Ng MCY, Hansen T, Andersen G, Borch-Johnsen K, Jorgensen T, Tres A, Fuertes F, Ruiz-Echarri M, Asin L, Saez B, van Boven E, Klaver S, Swinkels DW, Aben KK, Graif T, Cashy J, Suarez BK, van Vierssen O, Frigge ML, Ober C, Hofker MH, Wijmenga C, Christiansen C, Rader DJ, Palmer CNA, Rotimi C, Chan JCN, Pedersen O, Sigurdsson G, Benediktsson R, Jonsson E, Einarsson GV, Mayordomo JI, Catalona WJ, Kiemeney LA, Barkardottir RB, Gulcher JR, Thorsteinsdottir U, Kong A, Stefansson K: Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes. *Nat Genet* 39:977–983, 2007
25. Sandhu MS, Weedon MN, Fawcett KA, Wasson J, Debenham SL, Daly A, Lango H, Frayling TM, Neumann RJ, Sherva R, Blech I, Pharoah PD, Palmer CNA, Kimber C, Tavendale R, Morris AD, McCarthy MI, Walker M, Hitman G, Glaser B, Permutt MA, Hattersley AT, Wareham NJ, Barroso I: Common variants in WFS1 confer risk of type 2 diabetes. *Nat Genet* 39:951–953, 2007
26. Winckler W, Weedon MN, Graham RR, McCarroll SA, Purcell S, Almgren P, Tuomi T, Gaudet D, Bostrom KB, Walker M, Hitman G, Hattersley AT, McCarthy MI, Ardlie KG, Hirschhorn JN, Daly MJ, Frayling TM, Groop L, Altshuler D: Evaluation of common variants in the six known maturity-onset diabetes of the young (MODY) genes for association with type 2 diabetes. *Diabetes* 56:685–693, 2007
27. Frayling TM: Genome-wide association studies provide new insights into type 2 diabetes aetiology. *Nat Rev Genet* 8:657–662, 2007
28. Helgason A, Palsson S, Thorleifsson G, Grant SF, Emilsson V, Gunnarsdottir S, Adeyemo A, Chen Y, Chen G, Reynisdottir I, Benediktsson R, Hinney A, Hansen T, Andersen G, Borch-Johnsen K, Jorgensen T, Schafer H, Faruque M, Doumatey A, Zhou J, Wilensky RL, Reilly MP, Rader DJ, Bagger Y, Christiansen C, Sigurdsson G, Hebebrand J, Pedersen O, Thorsteinsdottir U, Gulcher JR, Kong A, Rotimi C, Stefansson K: Refining the impact of TCF7L2 gene variants on type 2 diabetes and adaptive evolution. *Nat Genet* 39:218–225, 2007
29. Morris AD, Boyle DI, MacAlpine R, Emslie-Smith A, Jung RT, Newton RW, MacDonald TM: The diabetes audit and research in Tayside Scotland (DARTS) study: electronic record linkage to create a diabetes register. DARTS/MEMO Collaboration. *BMJ* 315:524–528, 1997
30. Florez JC, Jablonski KA, Bayley N, Pollin TI, de Bakker PIW, Shuldiner AR, Knowler WC, Nathan DM, Altshuler D: TCF7L2 polymorphisms and progression to diabetes in the Diabetes Prevention Program. *New Engl J Med* 355:241–250, 2006
31. Weedon MN: The importance of TCF7L2. *Diabet Med* 24:1062–1066, 2007
32. Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN: Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet* 33:177–182, 2003
33. Lindstrom J, Ilanne-Parikka P, Peltonen M, Aunola S, Eriksson JG, Hemio K, Hamalainen H, Harkonen P, Keinanen-Kiukaanniemi S, Laakso M, Louheranta A, Mannelin M, Paturi M, Sundvall J, Valle TT, Uusitupa M, Tuomilehto J: Sustained reduction in the incidence of type 2 diabetes by lifestyle intervention: follow-up of the Finnish Diabetes Prevention Study. *Lancet* 368:1673–1679, 2006
34. Knowler WC, Barrett-Connor E, Fowler SE, Hamman RF, Lachin JM, Walker EA, Nathan DM: Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *New Engl J Med* 346:393–403, 2002
35. Tuomilehto J, Lindstrom J, Eriksson JG, Valle TT, Hamalainen H, Ilanne-Parikka P, Keinanen-Kiukaanniemi S, Laakso M, Louheranta A, Rastas M, Salminen V, Uusitupa M, Aunola S, Cepaitis Z, Moltchanov V, Hakumaki M, Mannelin M, Martikkala V, Sundvall J: Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. *New Engl J Med* 344:1343–1350, 2001
36. Pan XR, Li GW, Hu YH, Wang JX, Yang WY, An ZX, Hu ZX, Lin J, Xiao JZ, Cao HB, Liu PA, Jiang XG, Jiang YY, Wang JP, Zheng H, Zhang H, Bennett PH, Howard BV: Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance: the Da Qing IGT and Diabetes Study. *Diabetes Care* 20:537–544, 1997
37. Johnston N, Jernberg T, Lagerqvist B, Siegbahn A, Wallentin L: Improved identification of patients with coronary artery disease by the use of new lipid and lipoprotein biomarkers. *Am J Cardiol* 97:640–645, 2006
38. Kobberling J, Tillil H: Empirical risk figures for first degree relatives of non-insulin-dependent diabetics. In *The Genetics of Diabetes Mellitus*. Kobberling J, Tattersall R, Eds. London, Academic Press, 1982, p. 201–210
39. Janssens AC, Moonesinghe R, Yang Q, Steyerberg EW, van Duijn CM, Khoury MJ: The impact of genotype frequencies on the clinical validity of genomic profiling for predicting common chronic diseases. *Genet Med* 9:528–535, 2007
40. Lu Q, Elston RC: Using the optimal receiver operating characteristic curve to design a predictive genetic test, exemplified with type 2 diabetes. *Am J Hum Genet* 82:641–651, 2008
41. Lyssenko V, Almgren P, Anevski D, Orho-Melander M, Sjogren M, Saloranta C, Tuomi T, Groop L: Genetic prediction of future type 2 diabetes. *PloS Med* 2:1299–1308, 2005
42. Lyssenko V, Lupi R, Marchetti P, Del Guerra S, Orho-Melander M, Almgren P, Sjogren M, Ling C, Eriksson KF, Lethagen AL, Mancarella R, Berglund G, Tuomi T, Nilsson P, Del Prato S, Groop L: Mechanisms by which common variants in the TCF7L2 gene increase risk of type 2 diabetes. *J Clin Invest* 117:2155–2163, 2007
43. Weedon MN, McCarthy MI, Hitman G, Walker M, Groves CJ, Zeggini E, Rayner NW, Shields B, Owen KR, Hattersley AT, Frayling TM: Combining information from common type 2 diabetes risk polymorphisms improves disease prediction. *PLoS Med* 3:1877–1882, 2006