Supplementary Online Material: Zeggini *et al*

Samples

WTCCC genome-wide association scan

Cases: Details of the samples used in the WTCCC genome wide scan will be described elsewhere (*S1*). Briefly, a total of 1,999 samples were selected for the WTCCC genome-wide scan, all from the Diabetes UK Warren 2 repository (*S2*). Samples were collected in five main centres (Exeter, London, Oxford, Norwich and Newcastle) but with wide national representation. All were of British/Irish European descent, diagnosed between age 25 and 75. Approximately 30% of cases were explicitly recruited as part of multiplex sibships, and ~25% represented the T2D offspring within parent-offspring "triads" or "duos" (that is, offspring from sibships with only one parent available) (*S3*). The remainder were recruited as isolated cases ascertained for early age at diagnosis compared to the population distribution. Diagnosis of diabetes was based on either current prescribed treatment with diabetes-specific medication or, in the case of those treated with diet alone, historical or contemporary laboratory evidence of hyperglycemia (as defined by the World Health Organization). Other forms of diabetes were excluded by standard clinical criteria based on personal and family history. The final analysis on the scan included a total of 1,924 individuals after all QC-related exclusions.

Controls: a total of 3,004 samples were ascertained as controls from two sources. Equal numbers came from the 1958 Birth Cohort (*S4*), a longitudinal study of individuals born in a single week during March 1958, and the UK Blood Service Collection (blood donors ascertained through the National Blood Transfusion Service; UKBS controls). In both collections, ascertainment was distributed throughout the UK. No relevant phenotypic data are available for the UKBS controls; metabolic and

anthropometric data from the 1958 Birth Cohort are not available under the terms of their use within the WTCCC. The final analysis on the scan included a total of 2,938 individuals after exclusions.

All subjects gave written informed consent and the project protocols were approved by the relevant research ethics committees in the UK.

UKT2D Genetics Consortium replication samples

Replication set (RS) 1: *UK Type 2 Diabetes Genetics Consortium collection (first tranche).* All cases and controls were of European White descent, living in the Tayside region of Dundee when recruited. Cases had T2D diagnosed between the ages of 35-70 years (inclusive). The diagnosis of diabetes was based on either current prescribed treatment with diabetes-specific medication or, in the case of individuals treated with diet alone, laboratory evidence of hyperglycemia as defined by the World Health Organization. Patients were excluded if they had an established (clinical and/or molecular) diagnosis of monogenic diabetes (e.g. maturity-onset diabetes of the young, mitochondrial diabetes) or if they had been treated with regular insulin therapy within 1 year of diagnosis. Controls were aged below 80 years and had not been diagnosed with diabetes at the time of recruitment (or subsequently). Control subjects were excluded from analysis if laboratory investigations at the time of recruitment provided evidence of hyperglycaemia (fasting glucose $\geq$7.0 mmol/l, HbA1c >6.4%). RS1 comprised 2,022 cases and 2,037 controls meeting these criteria. This study was approved by the Tayside Medical Ethics Committee and informed consent was obtained from all subjects.

Replication set (RS) 2: *UK T2D cases and EFSOCH controls*. RS2 included 632 additional T2D cases and 1,750 population controls from the Exeter Family Study of Child Health (EFSOCH). The T2D cases were UK Whites derived from two sources:

202 were taken from a collection of young-onset T2D subjects (diagnosed before age 45y) who had been subjected to extensive analysis to exclude other causes of diabetes (*S5*). The remaining 430 were isolated T2D cases (age of diagnosis below 65y) (*S6*). Both subsets met the same criteria for diagnosing T2D described for the WTCCC cases. These cases were collected in a number of UK centres but ~60% come from SouthWest England. All subjects gave written informed consent and the project protocols were approved by the relevant research ethics committees in the UK. The controls were made up of parents from a consecutive birth cohort (EFSOCH: the Exeter Family Study of Child Health): only those with normal (<6.0mmol/l) fasting glucose and/or normal HbA1c levels (<6.0%; Diabetes Control and Complications trial corrected) were included (*S7*). Ethical approval was given by the North and East Devon Local Research Ethics Committee and informed consent was obtained from the parents of the newborns.

Replication set (RS) 3: *UK Type 2 Diabetes Genetics Consortium collection (second tranche)*. A subset of variants were typed in a further tranche of 1,103 cases and 1,559 controls from the UKT2DGC. These represented consecutive samples collected under exactly the same criteria as described for RS1 and with very similar clinical characteristics.

Basic clinical characteristics of all subjects are summarized in **Table S1**.

Genome-wide scan

Full details of DNA processing, genotyping, allele calling and QC as used in the WTCCC genome wide scan will be described elsewhere (*S1*). Briefly, DNA samples destined for Affymetrix genotyping were sent to the WTCCC laboratories at the Wellcome Trust Sanger Institute and the JDRF/Wellcome Trust Diabetes and Inflammation Laboratory, Cambridge for re-normalization and molecular

fingerprinting (Sequenom multiplex reactions). All T2D case and population control DNA samples sent to Affymetrix for genotyping passed through this pipeline and were evaluated with the same quality thresholds. Samples with low concentrations, evidence of degradation, poor PCR performance, or gender discrepancies with the stated information were rejected. Approved samples were re-arrayed into 96-well plates and shipped to California for genotyping. Whole genome SNP genotyping was performed with the commercial release of the GeneChip® 500K Mapping Array Set at Affymetrix' Service Facility. Arrays not passing the 93% call rate threshold at p 0.33 with the Dynamic Model (DM) algorithm were repeated. Affymetrix delivered successful samples as those having a DM call rate of 93% at p 0.33 for each array, over 90% concordance for SNPs common to the two arrays, cross-chip agreement gender and over 70% identity to the molecular fingerprinting genotypes previously generated. Intensity files generated from the .cel files for each sample were quantile-normalized and genotypes called using a novel calling algorithm (Chiamo++) (*S1*). A total of 75 T2D cases and 66 control individuals were excluded from further analysis due to relatedness, high missing data rates, excess heterozygosity, external discordances, duplicate sample status or non-European ethnic origin. Finally, the WTCCC Data Analysis Group excluded 30,583 autosomal SNPs with >5% missing call rates (>1% missing call rates if the SNP MAF<5% across all individuals), exact Hardy Weinberg equilibrium $p<5.7x10^{-7}$ and genotype frequency differences between the 2 control groups with $p<5.7x10^{-7}$. Analyses based on the remaining 459,448 autosomal SNPs will be reported in the WTCCC paper (*S1*).

Type 2 diabetes-specific quality control

SNP exclusions: For the analyses reported here, we considered only the 393,453 autosomal SNPs for which the minor allele frequency (MAF) in both T2D cases and

controls exceeded 1%, and for which there were no extreme departures from Hardy Weinberg Equilibrium (which we defined as a $p<10^{-4}$ in either cases or controls). Though differential genotype missingness between cases and controls can give rise to false positive results, we found that SNPs displaying substantial rates of differential missingness were generally excluded by other QC steps (such as tests of call rate and HWE). However, we checked that all SNPs prioritized for replication efforts in our study displayed acceptable rates of differential missingness ($p>0.001$). In addition, all SNPs taken forward to replication were demonstrated to map to a unique location in the human genome. We estimate that these SNPs capture at least 67% of common variation genome-wide based on HapMap CEU Phase II genotypes (based on pairwise assessments alone, requiring $r^2>0.8$ and MAF>0.05).

Sample checks: Based on the 393,453 SNPs passing T2D-specific quality control, we repeated checks for population stratification and cryptic relatedness performed in the full WTCCC data set, but now applied specifically to the 1,924 T2D cases and 2,938 controls passing WTCCC quality control. Consistent with the prior WTCCC quality control, all cases and controls were found to be of exclusive European descent. Comparison of the observed p value distribution for single-point Cochran-Armitage tests of association with that expected under the null indicates no substantial confounding from population substructure and genotyping bias (genome control inflation factor, $\lambda=1.08$) (**Figure S6)**. We also looked for evidence of population stratification using Eigenstrat (*S8*) and estimated its effects on association statistics by examining the correlation of χ^2 statistics obtained from a naive trend test against χ^2 statistics corrected for population structure. This analysis showed that there was no evidence that any of the more significant signals on the GWA scan were attributable to the effects of latent structure (**Figure S7**). In addition, adjustment for region of residence in a stratified analysis failed to reveal any evidence that the signals which we report here were the result of population stratification, nor did any

of the signals map close to any markers showing evidence of regional variation in allele frequency. Equivalent analyses are not possible in the replication datasets, as they have not been genome-scanned. However, in each case, there has been close matching of region of origin within the UK which, had substructure been an issue, would have minimized its impact. Finally, the strong replication in independent datasets from different parts of Europe provides further confidence that our T2D association findings are not due to stratification (since it is unlikely that the same biases would operate in each situation).

Because of issues related to X chromosome-specific analysis, a final set of WTCCC data was delayed with respect to the autosomal data and replication studies have not yet been initiated. However, there is no evidence from the WTCCC (*S1*) or other scans (*S9*, *S10*, *S11*) of any compelling susceptibility effects outside the autosomes.

Allelic nomenclature

Unless otherwise stated, all alleles are reported relative to the forward strand and all ORs [95%CIs] are predicted per copy of allele 2 at each variant. [Allele 1 and 2 are defined alphabetically with allele 1 being the forward strand allele which appears first in the sequence A,C,G,T]. In **Table 1** and at appropriate points in the text, we report all alleles with respect to the risk allele (as denoted), to allow clear alignment of the effects seen across the multiple studies (and to generate ORs that all exceed 1). Wherever we report results from the DGI and FUSION studies, we have used available LD information (from the HapMap and our own data) to confirm phase and direction of effect. Ancestral allele information was gathered from Entrez SNP and were derived by comparison against chimpanzee sequence.

Analysis

Unless otherwise stated, all results from the main WTCCC scan refer to Cochran-Armitage 1df tests for trend and are based on the 393,453 autosomal SNPs passing T2D-specific quality control. Single-point analyses for SNPs genotyped in the replication sets were carried out in PLINK version 0.99p (*S12*), STATA SE v9 (*S13*) and StatXact (*S14*) (again, using the Cochran-Armitage test for trend). Meta-analyses across UK case-control datasets were carried out using a fixed effects Mantel-Haenszel method for the additive model. We tested the dominant, general (2df) and recessive models for the emerging robust replication signals (**Table S8**). We also tested for deviations from the additive model in a logistic regression framework by calculating the reduction in deviance by fitting the general model. There was no evidence of deviation from the additive model in any dataset for any of the replicated SNPs (**Table S8**).

Combined estimates of the ORs across the UK, DGI and FUSION studies (**Table 1**) were calculated by weighting the logORs of each study by the inverse of their variance. When different SNPs have been genotyped to tag the same signal across studies, we combined data for the strongest proxies (**Table S3**). Utilization of the DGI and FUSION data to prioritize SNPs for replication could conceivably lead to some inflation of the effect estimate when all 3 studies (primary scans and replication sets) are combined (the "winners curse" argument). Since we based our SNP selection on GWA data alone, the association signals from the replication datasets available to DGI, FUSION and ourselves would not be subject to such bias. We therefore also estimated combined ORs based on the UK, DGI and FUSION replication sets only.

Intermediate trait analyses were carried out for body mass index (BMI), waist circumference and age of diagnosis (AoD). BMI and AoD were transformed to normality. All replication SNPs were analysed within T2D case groups by linear

regression, adjusting for gender (implemented in PLINK). T2D association analyses adjusted for BMI and waist circumference (and gender) were carried out using logistic regression in PLINK and STATA SE v9, with three replication sets as strata. These adjustment analyses were only possible in the replication sets, since BMI and waist circumference values for the WTCCC controls are not available. Quantitative trait and adjusted analyses were combined across groups using the inverse variance method for continuous data (**Tables S5-S7**) and the Mantel-Haenszel method for binary data.

Haplotype-based analyses were carried out using two complementary approaches. GENEBPM (*S15*) was used to analyse 5 SNP sliding windows across regions centred on replicating T2D signals. The size of each region was guided by local LD architecture and was delineated to extend beyond the location of flanking recombination hotspots, in order to ensure that all haplotypes relevant to these signals were covered. The GENEBPM algorithm uses Bayesian modelling techniques. Evidence in favour of association of T2D with a single SNP or haplotypes in a sliding window are presented in terms of a Bayes' factor (*S1*), the ratio of marginal likelihoods under models of association and no association. By convention, a log10 Bayes' factor of greater than 2 is taken as decisive evidence of association, although this does not take account of multiple testing.

As a complementary strategy, we also followed a LD block-based approach. For these analyses, haplotype blocks (*S16*) were defined for SNPs in the regions surrounding robustly replicating signals using HaploView v14 (*S17*). We then carried out haplotype-based analyses within blocks using UNPHASED (*S18*).

Pairwise interaction analyses were carried out for replicating signals and established T2D susceptibility genes in PLINK (*S12*), testing deviation of SNP-pair association from the log additive model.

Linkage disequilibrium plots for the Figures (**1, 2, S1, S3-S5**) were constructed using GoldSurfer2 (*S19*), using both HapMap CEU (for the lower panels) and WTCCC case data (for the upper panels). Plots based on HapMap data included pairwise r^2 statistics for all HapMap-genotyped SNPs to provide an overview of the regional genomic architecture, incorporating as much information as possible. Apparent differences in the patterns of LD between the two panels in the relevant figures largely reflect the substantial differences in SNP density between the two datasets. When the HapMap SNP map was pruned to match the SNP content achieved in WTCCC study, patterns of LD between the 2 datasets were almost identical (see (*S1*) for examples of other regions).

Estimates of λ_s were based on the effect sizes from the combined UK data (**Table 1**) and calculated as described in (*S20*).

SNP selection for replication

Based on strength of association in the primary single-point analysis of 393,453 SNPs, we selected (in the "first wave") at least one SNP from each region achieving $p<10^{-5}$ (Cochran-Armitage test for trend) for replication. However, even the WTCCC scan (featuring comparison of ~2,000 cases and ~3,000 controls) only has reasonable power for the detection of relatively large signals (e.g. 80% power at $\alpha=10^{-5}$ for a log additive OR of 1.3 and MAF of 20%). Given the distribution of effects expected on both theoretical and empirical grounds, most genuine susceptibility variants are likely to have effect sizes smaller than this (*S21, S22*). Furthermore, even variants with large effects may generate modest evidence for association when there is incomplete LD with directly genotyped SNPs. We therefore undertook a "second wave" of replication, for which we developed a heuristic approach to the

prioritisation of signals generating p values between 10^{-2} and 10^{-5} in the primary GWA scan, taking account of the following additional criteria:

(1) Corroborating evidence for association with T2D in the companion DGI (*S9*) and FUSION (*S10*) scans. Specifically, we up-weighted signals displaying at least a trend for association in the same direction in both scans as well as signals displaying evidence for association in the same direction in either scan. To define overlapping signals, we considered FUSION (Illumina array) SNPs that had $r^2>0.5$ with our (Affymetrix chip) variants, where the same SNP was not represented in both platforms.

(2) Biological candidacy of the gene. We assessed candidacy by informal searches of the literature and bioinformatics databases, in addition to applying a systematic search of publicly available data, using the GeneSniffer program (www.genesniffer.org).

(3) Identification of multiple independent associations within the same locus (defined as $r^2<0.4$).

These criteria would have led to selection of SNPs within both *KCNJ11* and *PPARG* for "second wave" replication despite the modest evidence for association based on the original WTCCC scan. Indeed, despite concerns that differences in ethnic origin, ascertainment schemes, genotyping platforms and analysis plans across the three studies would result in effect size heterogeneity, the enhanced signals observed at known susceptibility variants in *KCNJ11* (p=0.0013 in WTCCC, $p\sim5.0\times10^{-11}$ in combined analysis of all 3 studies) and *PPARG* (p=0.0013 in WTCCC, $p\sim1.7\times10^{-6}$ in combined analysis) (**Tables 1, S3**), provided encouragement that this approach would highlight additional loci with high prior odds of association.

Information outlining why each of the "second-wave" SNPs has been selected is provided in **Table S9**.

Genotype cluster plots for all SNPs prioritized for replication were inspected in each of the T2D cases, UKBS and 1958 birth cohort control groups separately. Only SNPs assays demonstrating acceptable clustering were taken forward to validation and replication.

In addition, we also genotyped rs13266634, in the *SLC30A8* gene, in all UK samples, on the basis of collective evidence for association with type 2 diabetes in the recently published French genome-wide scan (*S11*) and the FUSION study (*S10*).

Validation genotyping

All of the SNPs from the primary WTCCC scan that were prioritized for replication and demonstrated acceptable clustering in the WTCCC data (as well as selected proxies thereof) were re-genotyped for 295 of the WTCCC samples by a second genotyping method (at Kbiosciences, see next section). Overall, there was a 99.3% concordance rate (203 discrepancies out of 29,073 informative comparisons) between the methods. One SNP was monomorphic when genotyped by Kbioscience, accounting for 135 discrepancies; excluding this SNP, the concordance rate is 99.8% (28,721/28,789). 96% of SNPs had genotype concordance rates >99.5% between the two methods. In general, for each signal of interest, we designed replication assays for at least two SNPs, and chose the SNP with the better performance on validation (call rate, concordance) for replication genotyping.

Replication genotyping

Genotyping of RS1, 2 and 3 was performed by Kbiosciences (Herts., UK). Kbiosciences designed and used assays based on either their proprietary competitive allele specific PCR system (KASPar) method or modified TaqMan assays, details of which are available on their website

(http://www.kbioscience.co.uk/chemistry/index.htm). All assays were validated prior to use, using a standard 96-well validation plate used by Kbiosciences and 295 samples from the WTCCC study (see validation genotyping, above).

Expression studies

RNA samples were obtained from Clontech (Oxford, UK) or AMS Bioscience, Abingdon, UK. Isolated islet cells were purchased from the National Disease Resource Interchange [NDRI] (Philadelphia, USA). All RNA samples were DNAse treated prior to reverse transcription using the TURBO DNAse kit (Ambion, Huntingdon, UK). RNA samples were reverse transcribed using the Thermoscript RT-PCR System (Invitrogen, Paisley, UK) in a total volume of 20μl. cDNA samples were treated with 1u RNAse H at 37°C for 30 minutes (Invitrogen, Paisley, UK) to remove RNA prior to amplification. Assays to beta 2 microglobulin (*B2M*) and *CDKN2A* were purchased from Applied Biosystems (Assays by Demand). Assays to beta glucuronidase (*BGUS*), *CDKN2B*, *CDKAL1*, *IGF2BP2* and *HHEX* were custom-designed through the equivalent Assays by Design service. Primer and probe sequences are available from the authors upon request. Where genes coded for multiple isoforms, probes and primers were designed to anneal to common regions as far as possible. Custom assays only were validated using standard curve analysis. Real-time PCR reactions were carried out using the ABI Prism 7900 platform in a total volume of 10μl. Each sample was amplified in triplicate to ensure accuracy of quantification. PCR reactions contained 5μl TaqMan Fast Universal Mastermix (no AMPerase) (Applied Biosystems, Foster City, USA), 0.9μM each primer, 0.25μM probe and 2μl cDNA reverse transcribed as

above in a total volume of 10μl. PCR conditions were a single cycle of 95ºC for 20 seconds followed by 50 cycles of 95ºC for 1 second and 60ºC for 20 seconds.

Relative expression levels for each test gene were calculated by the ΔΔCT method, based on comparison of the average real-time PCR crossing point obtained for each test gene (ΔCT^{test}) relative to that obtained from an endogenous control (ΔCt^{ref}). To minimize variation in endogenous control transcript levels between tissues, we pooled results for two control genes, beta glucuronidase (*BGUS*) and beta 2 microglobulin (*B2M*). The expression level of each test gene was then determined as a test:*B2M*/*BGUS* ratio. Final expression profiles for each test gene were produced by normalization of the relative ratios for each tissue sample with the test transcript levels in human adult pancreas for tissue to tissue comparison.

Replication set-only combined effect size estimation

When we estimated combined ORs based on the UK, DGI and FUSION replication sets only, the p values for all three of the shared signals reported provided very strong support (*IGF2BP2* OR 1.14 (1.09-1.18), $p=7.5x10^{-10}$; *CDKAL1* 1.09 (1.05-1.14), $p=2.0x10^{-5}$; *CDKN2B* 1.17 (1.11-1.24), $p=4.3x10^{-9}$) for association. Notably, the effect size estimates from those replication sets are mostly very similar to those observed in the original GWA datasets (**Table 1**), which indicates that the extent of the effect size inflation is modest.

Haplotype-based analysis results

Haplotype=based analyses (using GENEBPM (*S15*) and UNPHASED (*S18*)) did not provide additional evidence for association in any of the genomic intervals examined, with the sole exception of the chr9 region (and in particular the 3' signal at 22.12Mb). For the signals in this region, haplotype-based approaches within the

WTCCC data provided overwhelming evidence for association with type 2 diabetes. GENEBPM (*S15*) analysis of 5 SNP sliding windows generated a 3' haplotype signal with Bayes' factors more than four orders of magnitude greater than the equivalent single-point analyses (high risk haplotypes shared alleles T and T at rs10811661 and rs10757283 [**Figure S8**]). Our complementary LD block-based analysis, using blocks as defined by Gabriel et al (*S16*), also revealed that a haplotype (consisting of rs2383208, rs10811661, rs10757283 and rs1333051) provided stronger association than single-point analysis ($p=5.8x10^{-8}$ for the comparison of the risk haplotype with all others). The two SNPs, rs10811661 (the SNP driving the single-point result) and rs10757283, which defined this haplotype were genotyped in RS1 and RS2. The association of the TT haplotype (frequency in controls ~24%) formed by rs10811661 and rs10757283 was replicated in RS1 and RS2 ($p=2.2x10^{-4}$) and combining this evidence with that from the WTCCC data provided very strong evidence of association (OR, 1.24 [1.16-1.31], $p=2.9x10^{-11}$). Analysis of CEU HapMap data has failed to identify any single SNP that could account for this haplotypic association, suggesting that the association may be explained by an as yet unidentified variant in the region. (A second possibility is that the two SNPs -or proxies thereof- have independent and synergistic effects on T2D risk). Comprehensive fine-mapping and extensive resequencing will be required to characterize the disease gene architecture of this region.

Overlap between association and linkage signals

We found little/no overlap between the location of previously reported type 2 diabetes linkage peaks (*S23*) and association signals arising from the genome-wide association scan. In a comparison of the total number of observed independent hits to that expected (assuming constant LD) within regions demarked by "one-LOD-

drop" intervals under T2D linkage peaks, we did not find convincing evidence of regional over-representation. In this analysis, the number of expected hits was derived from the total number of hits genome-wide ($p<10^{-3}$) and number of independent loci (r^2 was conservatively set at 0.8) in each linkage region. Although we observed nominal evidence of an excess of significant association p values within the chr10 linkage region (in which the *HHEX* signal resides) (*S2*), these results are hard to interpret given that none of the association signals detected in our work or in others (including *TCF7L2*) has characteristics which could generate a detectable linkage signal (as can be shown by estimating the lambda(s) values attributable to each signal). It is possible that rarer, more penetrant variants mapping to the *HHEX* association region are responsible for the linkage signal on chr10, but establishing this will require extensive resequencing efforts.

Quantitative trait analysis results

FTO showed consistent patterns of association with BMI and waist circumference, and the type 2 diabetes signal was attenuated by adjustment for these covariates (**Tables S5-S7**). None of the other SNPs studied recapitulated this pattern. There were some interesting nominal associations in the quantitative analysis of waist circumference and BMI in specific datasets, but none of these was consistent and in no case was the overall T2D effect attenuated or abolished by adjustment for BMI or waist circumference (**Tables S5-S7**).

Departures from additivity

We found no evidence for departures from additivity at any of the signals, except for *CDKAL1* SNP rs9465871 ($p=3.0\times10^{-3}$). Of the 1df tests carried out, the dominant model fits the data best ($p=1.1\times10^{-7}$), and the 2df general model provides a

marginally better fit overall ($p=4.2x10^{-8}$) (**Table S8**). This SNP is in partial LD with ($r^2=0.4$) with the *CDKAL1* SNP we report in Table 1 (rs10946398), for which no departure from additivity is evident. We have observed no significant two-way interactions between the T2D susceptibility variants examined (**Table S10**).

Data access details

Information on accessing data from the WTCCC scan is available at http://www.wtccc.org.uk/info/access_to_data_samples.shtml. Briefly, summary level data (association P values at each SNP) will be made widely available with publication of the primary report of the WTCCC study. Individual level data (that is, individual genotypes and case-control phenotypes) will be available shortly thereafter to approved users who apply through the Consortium Data Access Committee.

For the replication sets, genotype counts for all replication sets will be available from the authors on request. For the confirmed associations, individual genotypes and relevant phenotypic information) will be made available subject to a simple registration of interest (https://crs.dundee.ac.uk/).

Membership of WTCCC

Management Committee: Paul R Burton[1], David G Clayton[2], Lon R Cardon[3], Nick Craddock[4], Panos Deloukas[5], Audrey Duncanson[6], Dominic P Kwiatkowski[3,5], Mark I McCarthy[3,7], Willem H Ouwehand[8,9], Nilesh J Samani[10], John A Todd[2], Peter Donnelly (Chair)[11]

Analysis Committee: Jeffrey C Barrett[3], Paul R Burton[1], Dan Davison[11], Peter Donnelly[11], Doug Easton[12], David Evans[3], Hin-Tak Leung[2], Jonathan L Marchini[11], Andrew P Morris[3], Chris CA Spencer[11], Martin D Tobin[1], Lon R Cardon (Co-chair)[3], David G Clayton (Co-chair)[2]

UK Blood Services & University of Cambridge Controls: Antony P Attwood[5,8], James P Boorman[8,9], Barbara Cant[8], Ursula Everson[13], Judith M Hussey[14], Jennifer D Jolley[8], Alexandra S Knight[8], Kerstin Koch[8], Elizabeth Meech[15], Sarah Nutland[2], Christopher V Prowse[16], Helen E Stevens[2], Niall C Taylor[8], Graham R Walters[17], Neil M Walker[2], Nicholas A Watkins[8,9], Thilo Winzer[8], John A Todd[2], Willem H Ouwehand[8,9]

1958 Birth Cohort Controls: Richard W Jones[18], Wendy L McArdle[18], Susan M Ring[18], David P Strachan[19], Marcus Pembrey[18,20]

Bipolar Disorder (Aberdeen): Gerome Breen[21], David St Clair[21]; (**Birmingham):** Sian Caesar[22], Katherine Gordon-Smith[22,23], Lisa Jones[22]; **(Cardiff):** Christine Fraser[23], Elaine K Green[23], Detelina Grozeva[23], Marian L Hamshere[23], Peter A Holmans[23], Ian R Jones[23], George Kirov[23], Valentina Moskvina[23], Ivan Nikolov[23], Michael C O'Donovan[23], Michael J Owen[23], Nick Craddock[23]; **(London):** David A Collier[24], Amanda Elkin[24], Anne Farmer[24], Richard Williamson[24], Peter McGuffin[24]; **(Newcastle):** Allan H Young[25], I Nicol Ferrier[25]

Coronary Artery Disease (Leeds): Stephen G Ball[26], Anthony J Balmforth[26], Jennifer H Barrett[26], D Timothy Bishop[26], Mark M Iles[26], Azhar Maqbool[26], Nadira Yuldasheva[26], Alistair S Hall[26]; **(Leicester):** Peter S Braund[10], Paul R Burton[1],

Richard J Dixon[10], Massimo Mangino[10], Suzanne Stevens[10], Martin D Tobin[1], John R Thompson[1], Nilesh J Samani[10]

Crohn's Disease (Cambridge): Francesca Bredin[27], Mark Tremelling[27], Miles Parkes[27]; **(Edinburgh):** Hazel Drummond[28], Charles W Lees[28], Elaine R Nimmo[28], Jack Satsangi[28]; **(London):** Sheila A Fisher[29], Alastair Forbes[30], Cathryn M Lewis[29], Clive M Onnie[29], Natalie J Prescott[29], Jeremy Sanderson[31], Christopher G Mathew[29]; **(Newcastle):** Jamie Barbour[32], M Khalid Mohiuddin[32], Catherine E Todhunter[32], John C Mansfield[32]; **(Oxford):** Tariq Ahmad[33], Fraser R Cummings[33], Derek P Jewell[33]

Hypertension (Aberdeen): John Webster[34]; **(Cambridge):** Morris J Brown[35], David G Clayton[2]; **(Evry, France): G** Mark Lathrop[36]; **(Glasgow):** John Connell[37], Anna Dominiczak[37]; **(Leicester):** Nilesh J Samani[10]; **(London):** Carolina A Braga Marcano[38], Beverley Burke[38], Richard Dobson[38], Johannie Gungadoo[38], Kate L Lee[38], Patricia B Munroe[38], Stephen J Newhouse[38], Abiodun Onipinla[38], Chris Wallace[38], Mingzhan Xue[38], Mark Caulfield[38]; **(Oxford):** Martin Farrall[39]

Rheumatoid Arthritis: Anne Barton[40], The Biologics in RA Genetics and Genomics Study Syndicate (BRAGGS) Steering Committee*, Ian N Bruce[40], Hannah Donovan[40], Steve Eyre[40], Paul D Gilbert[40], Samantha L Hider[40], Anne M Hinks[40], Sally L John[40], Catherine Potter[40], Alan J Silman[40], Deborah PM Symmons[40], Wendy Thomson[40], Jane Worthington[40]

Type 1 Diabetes: David G Clayton[2], David B Dunger[2,41], Sarah Nutland[2], Helen E Stevens[2], Neil M Walker[2], Barry Widmer[2,41], John A Todd[2]

Type 2 Diabetes (Exeter): Timothy M Frayling[42,43], Rachel M Freathy[42,43], Hana Lango[42,43], John R B Perry[42,43], Beverley M Shields[43], Michael N Weedon[42,43], Andrew T Hattersley[42,43]; **(London):** Graham A Hitman[44]; **(Newcastle):** Mark Walker[45]; **(Oxford):** Kate S Elliott[3,7], Christopher J Groves[7], Cecilia M Lindgren[3,7], Nigel W Rayner[3,7], Nicholas J Timpson[3,46], Eleftheria Zeggini[3,7], Mark I McCarthy[3,7]

Tuberculosis (Gambia): Melanie Newport[47], Giorgio Sirugo[47]; **(Oxford):** Emily Lyons[3], Fredrik Vannberg[3], Adrian VS Hill[3]

Ankylosing Spondylitis: Linda A Bradbury[48], Claire Farrar[49], Jennifer J Pointon[48], Paul Wordsworth[49], Matthew A Brown[48,49]

AutoImmune Thyroid Disease: Jayne A Franklyn[50], Joanne M Heward[50], Matthew J Simmonds[50], Stephen CL Gough[50]

Breast Cancer: Sheila Seal[51], Breast Cancer Susceptibility Collaboration (UK)*, Michael R Stratton[51,52], Nazneen Rahman[51]

Multiple Sclerosis: Maria Ban[53], An Goris[53], Stephen J Sawcer[53], Alastair Compston[53]

Gambian Controls (Gambia): David Conway[47], Muminatou Jallow[47], Melanie Newport[47], Giorgio Sirugo[47]; **(Oxford):** Kirk A Rockett[3], Dominic P Kwiatkowski[3,5]

DNA, Genotyping, Data QC and Informatics (Wellcome Trust Sanger Institute, Hinxton): Suzannah J Bumpstead[5], Amy Chaney[5], Kate Downes[2,5], Mohammed JR Ghori[5], Rhian Gwilliam[5], Sarah E Hunt[5], Michael Inouye[5], Andrew Keniry[5], Emma King[5], Ralph McGinnis[5], Simon Potter[5], Rathi Ravindrarajah[5], Pamela Whittaker[5], Claire Widden[5], David Withers[5], Panos Deloukas[5]; **(Cambridge):** Hin-Tak Leung[2], Sarah Nutland[2], Helen E Stevens[2], Neil M Walker[2], John A Todd[2]

Statistics (Cambridge): Doug Easton[12], David G Clayton[2]; **(Leicester):** Paul R Burton[1], Martin D Tobin[1]; **(Oxford):** Jeffrey C Barrett[3], David Evans[3], Andrew P Morris[3], Lon R Cardon[3]; **(Oxford):** Niall J Cardin[11], Dan Davison[11], Teresa Ferreira[11], Joanne Pereira-Gale[11], Ingeleif B Hallgrimsdóttir[11], Bryan N Howie[11], Jonathan L Marchini[11], Chris CA Spencer[11], Zhan Su[11], Yik Ying Teo[3,11], Damjan Vukcevic[11], Peter Donnelly[11]

PIs: David Bentley[5,54], Matthew A Brown[48,49], Lon R Cardon[3], Mark Caulfield[38], David G Clayton[2], Alistair Compston[53], Nick Craddock[23], Panos Deloukas[5], Peter Donnelly[11],

Martin Farrall[39], Stephen CL Gough[50], Alistair S Hall[26], Andrew T Hattersley[42,43], Adrian VS Hill[3], Dominic P Kwiatkowski[3,5], Christopher G Mathew[29], Mark I McCarthy[3,7], Willem H Ouwehand[8,9], Miles Parkes[27], Marcus Pembrey[18,20], Nazneen Rahman[51], Nilesh J Samani[10], Michael R Stratton[51,52], John A Todd[2], Jane Worthington[40]

[1] Genetic Epidemiology Group, Department of Health Sciences, University of Leicester, Adrian Building, University Road, Leicester, LE1 7RH, UK; [2] Juvenile Diabetes Research Foundation/Wellcome Trust Diabetes and Inflammation Laboratory, Department of Medical Genetics, Cambridge Institute for Medical Research, University of Cambridge, Wellcome Trust/MRC Building, Cambridge, CB2 0XY, UK; [3] Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK; [4] Department of Psychological Medicine, Henry Wellcome Building, School of Medicine, Cardiff University, Heath Park, Cardiff CF14 4XN, UK; [5] The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK; [6] The Wellcome Trust, Gibbs Building, 215 Euston Road, London NW1 2BE, UK; [7] Oxford Centre for Diabetes, Endocrinology and Medicine, University of Oxford, Churchill Hospital, Oxford, OX3 7LJ, UK; [8] Department of Haematology, University of Cambridge, Long Road, Cambridge, CB2 2PT, UK; [9] National Health Service Blood and Transplant, Cambridge Centre, Long Road, Cambridge, CB2 2PT, UK; [10] Department of Cardiovascular Sciences, University of Leicester, Glenfield Hospital, Groby Road, Leicester, LE3 9QP, UK; [11] Department of Statistics, University of Oxford, 1 South Parks Road, Oxford OX1 3TG, UK; [12] Cancer Research UK Genetic Epidemiology Unit, Strangeways Research Laboratory, Worts Causeway, Cambridge CB1 8RN, UK; [13] National Health Service Blood and Transplant, Sheffield Centre, Longley Lane, Sheffield S5 7JN, UK; [14]

National Health Service Blood and Transplant, Brentwood Centre, Crescent Drive, Brentwood, CM15 8DP, UK; [15] The Welsh Blood Service, Ely Valley Road, Talbot Green, Pontyclun, CF72 9WB, UK; [16] The Scottish National Blood Transfusion Service, Ellen's Glen Road, Edinburgh, EH17 7QT, UK; [17] National Health Service Blood and Transplant, Southampton Centre, Coxford Road, Southampton, SO16 5AF, UK; [18] Avon Longitudinal Study of Parents and Children, University of Bristol, 24 Tyndall Avenue, Bristol, BS8 1TQ, UK; [19] Division of Community Health Services, St George's University of London, Cranmer Terrace, London SW17 0RE, UK; [20] Institute of Child Health, University College London, 30 Guilford St, London WC1N 1EH, UK; [21] University of Aberdeen, Institute of Medical Sciences, Foresterhill, Aberdeen, AB25 2ZD, UK; [22] Department of Psychiatry, Division of Neuroscience, Birmingham University, Birmingham, B15 2QZ, UK; [23] Department of Psychological Medicine, Henry Wellcome Building, School of Medicine, Cardiff University, Heath Park, Cardiff CF14 4XN, UK; [24] SGDP, The Institute of Psychiatry, King's College London, De Crespigny Park Denmark Hill London SE5 8AF, UK; [25] School of Neurology, Neurobiology and Psychiatry, Royal Victoria Infirmary, Queen Victoria Road, Newcastle upon Tyne, NE1 4LP, UK; [26] LIGHT and LIMM Research Institutes, Faculty of Medicine and Health, University of Leeds, Leeds, LS1 3EX, UK; [27] IBD Research Group, Addenbrooke's Hospital, University of Cambridge, Cambridge, CB2 2QQ, UK; [28] Gastrointestinal Unit, School of Molecular and Clinical Medicine, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU UK; [29] Department of Medical & Molecular Genetics, King's College London School of Medicine, 8th Floor Guy's Tower, Guy's Hospital, London, SE1 9RT, UK; [30] Institute for Digestive Diseases, University College London Hospitals Trust, London, NW1 2BU, UK; [31] Department of Gastroenterology, Guy's and St Thomas' NHS Foundation Trust, London, SE1 7EH, UK; [32] Department of Gastroenterology & Hepatology, University of Newcastle upon Tyne, Royal Victoria Infirmary, Newcastle upon Tyne, NE1 4LP, UK;

[33] Gastroenterology Unit, Radcliffe Infirmary, University of Oxford, Oxford, OX2 6HE, UK; [34] Medicine and Therapeutics, Aberdeen Royal Infirmary, Foresterhill, Aberdeen, Grampian AB9 2ZB, UK; [35] Clinical Pharmacology Unit and the Diabetes and Inflammation Laboratory, University of Cambridge, Addenbrookes Hospital, Hills Road, Cambridge CB2 2QQ, UK; [36] Centre National de Genotypage, 2, Rue Gaston Cremieux, Evry, Paris 91057.; [37] BHF Glasgow Cardiovascular Research Centre, University of Glasgow, 126 University Place, Glasgow, G12 8TA, UK; [38] Clinical Pharmacology and Barts and The London Genome Centre, William Harvey Research Institute, Barts and The London, Queen Mary's School of Medicine, Charterhouse Square, London EC1M 6BQ, UK; [39] Cardiovascular Medicine, University of Oxford, Wellcome Trust Centre for Human Genetics, Roosevelt Drive, Oxford OX3 7BN, UK; [40]arc Epidemiology Research Unit, University of Manchester, Stopford Building, Oxford Rd, Manchester, M13 9PT, UK; [41] Department of Paediatrics, University of Cambridge, Addenbrooke's Hospital, Cambridge, CB2 2QQ, UK; [42] Genetics of Complex Traits, Institute of Biomedical and Clinical Science, Peninsula Medical School, Magdalen Road, Exeter EX1 2LU UK; [43] Diabetes Genetics, Institute of Biomedical and Clinical Science, Peninsula Medical School, Barrack Road, Exeter EX2 5DU UK; [44] Centre for Diabetes and Metabolic Medicine, Barts and The London, Royal London Hospital, Whitechapel, London, E1 1BB UK; [45] Diabetes Research Group, School of Clinical Medical Sciences, Newcastle University, Framlington Place, Newcastle upon Tyne NE2 4HH, UK; [46] The MRC Centre for Causal Analyses in Translational Epidemiology, Bristol University, Canynge Hall, Whiteladies Rd, Bristol BS2 8PR, UK; [47] MRC Laboratories, Fajara, The Gambia; [48] Diamantina Institute for Cancer, Immunology and Metabolic Medicine, Princess Alexandra Hospital, University of Queensland, Woolloongabba, Qld 4102, Australia; [49] Botnar Research Centre, University of Oxford, Headington, Oxford OX3 7BN, UK; [50] Department of Medicine, Division of Medical Sciences, Institute of Biomedical Research, University of

Birmingham, Edgbaston, Birmingham B15 2TT, UK; [51] Section of Cancer Genetics, Institute of Cancer Research, 15 Cotswold Road, Sutton, SM2 5NG, UK; [52] Cancer Genome Project, The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK; [53] Department of Clinical Neurosciences, University of Cambridge, Addenbrooke's Hospital, Hills Road, Cambridge CB2 2QQ, UK; [54] PRESENT ADDRESS: Illumina Cambridge, Chesterford Research Park, Little Chesterford, Nr Saffron Walden, Essex, CB10 1XL, UK.

Detailed acknowledgements

Collection of the type 2 diabetes cases was supported by Diabetes UK, BDA Research and the UK Medical Research Council (Biomedical Collections Strategic Grant G0000649). The UK Type 2 Diabetes Genetics Consortium collection was supported by the Wellcome Trust (Biomedical Collections Grant GR072960). The genome-wide association genotyping was supported by the Wellcome Trust (076113) and replication genotyping by the European Commission (EURODIA LSHG-CT-2004-518153), Wellcome Trust, Peninsula Medical School, and Diabetes UK. Personal funding comes from the Wellcome Trust (EZ: Research Career Development Fellow; ATH: Research Leave Fellow; LRC: Principal Research Fellow); UK Medical Research Council (JRBP); Diabetes UK (RMF); Throne-Holst Foundation (CML); Leverhulme Trust (APM); Research Councils UK (LWH); and the Scottish Executive Generations Scotland Initiative (ADM). MNW is Vandervell Foundation Research Fellow at the Peninsula Medical School. We acknowledge the assistance of Robert Lawrence and Dr Paul Emsley and the particular contribution of Dr Colin Palmer, Dr Michael Sampson, Prof Stephen O'Rahilly, and our team of research nurses as well as the continued support provided by Prof Simon Howell, Dr Moira Murphy and Dr Angela Wilson (Diabetes UK). We acknowledge the efforts of Jane Collier, Phil Robinson, Steven Asquith and others at Kbiosciences (http://www.kbioscience.co.uk/) for their rapid and accurate large scale genotyping. Sample preparation was helped by Bryan Bolton and colleagues at ECACC and Martin Yuille at CIGMR. Chemgenex Pharmaceuticals are acknowledged for allowing use of GeneSniffer. We are grateful to members of the DGI and FUSION teams for sharing data from their studies and to all the subjects who participated in the research.

Figure S1: Overview of *FTO* signal region

A Plot of -log(p values) for T2D (Cochran-Armitage test for trend) against chromosome position in Mb. Blue diamonds represent primary scan results and pink triangles denote meta-analysis results across all UK samples.

B Genomic location of genes showing intron and exon structure (NCBI Build 35). Pink triangles show position of replication SNPs relative to gene structure.

C MULTIZ (*S24*) vertebrate alignment of 17 species showing evolutionary conservation.

D GoldSurfer2 (*S19*) plot of linkage disequilibrium (r^2) for SNPs genotyped in WTCCC scan (passing T2D-specific quality control) in WTCCC T2D cases.

E Recombination rate given as cM/MB. Red boxes represent recombination hotspots (*S25*).

F GoldSurfer2 plot of linkage disequilibrium (r^2) for all HapMap SNPs across the region (HapMap CEU data).

Figure S1: Overview of *FTO* signal region

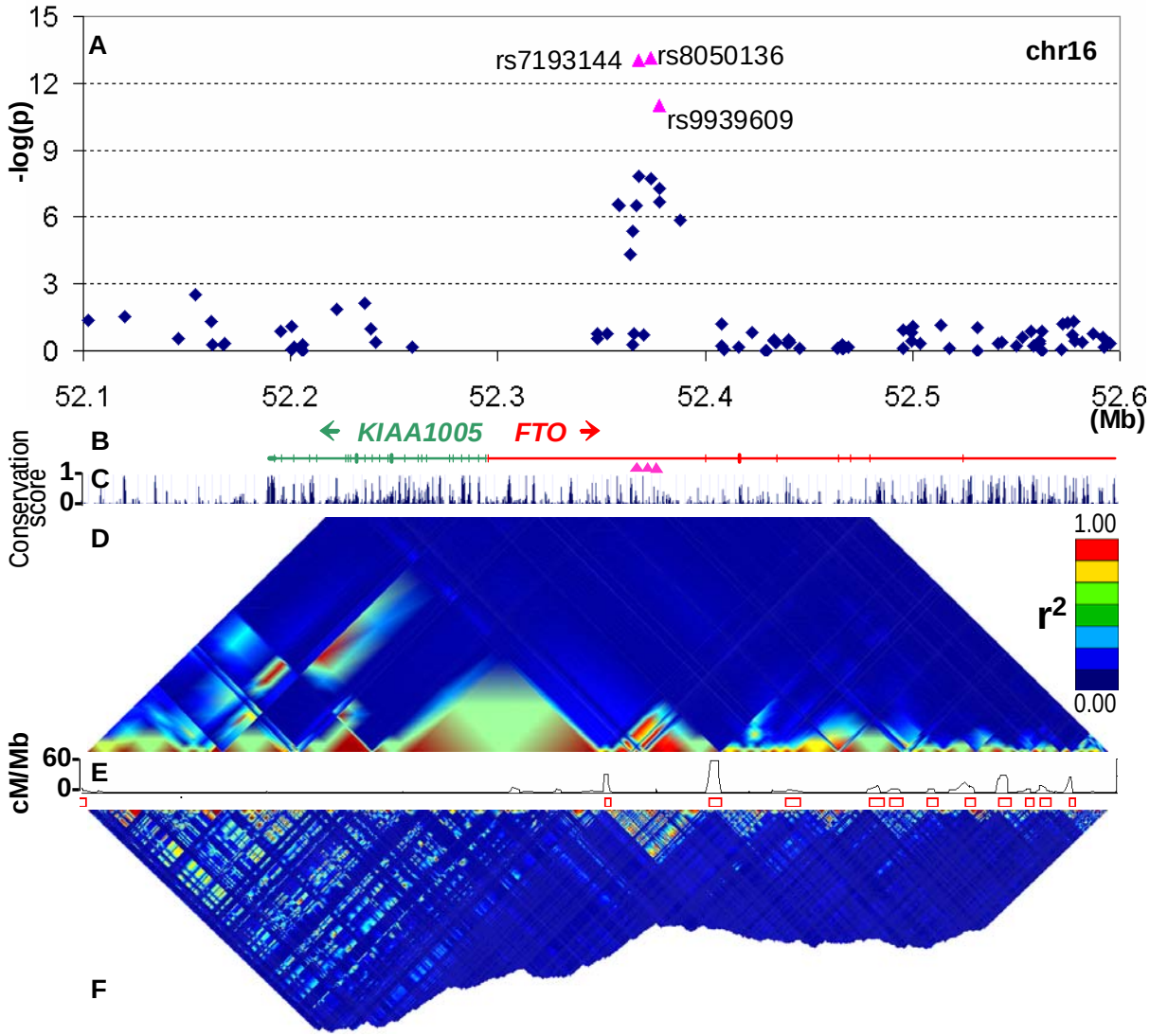
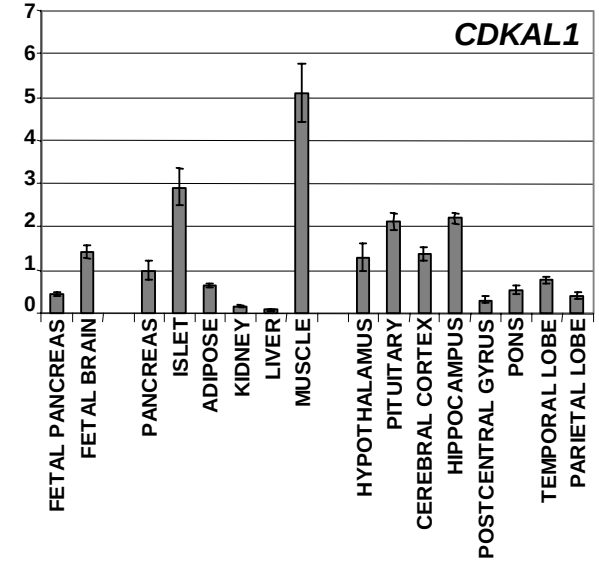
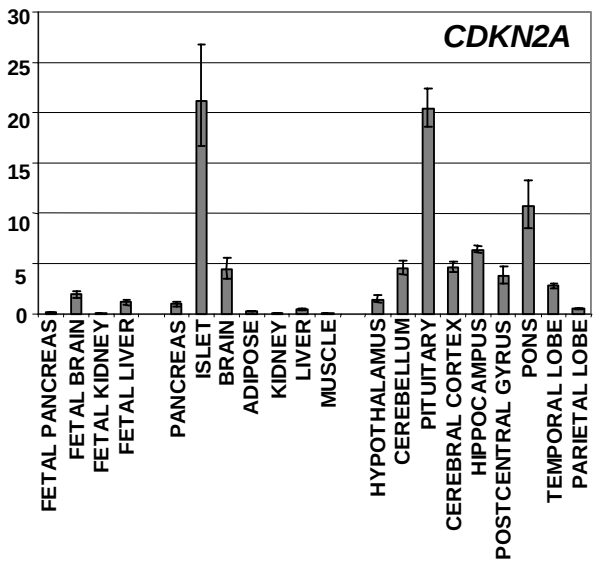
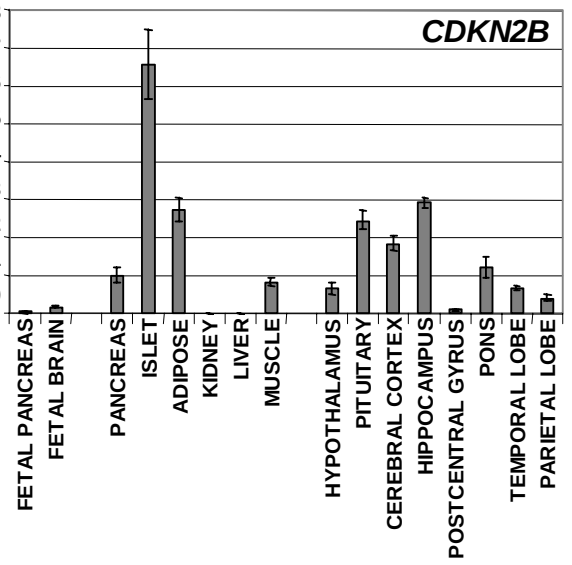


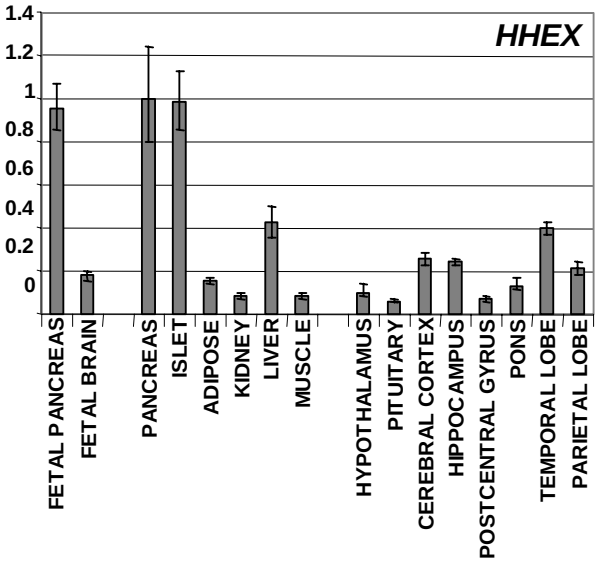
Figure S2: Expression patterns of *CDKAL1*, *CDKN2A*, *CDKN2B*, *HHEX*, *IGF2BP2*. Messenger RNA expression profiles are shown for the genes listed for a range of human tissues, as determined by RT-PCR. Figures on the y axes refer to the test transcript levels relative to two separate endogenous control genes (beta glucuronidase [*BGUS*] and beta 2 microglobulin [*B2M*]). Each test:control ratio was then normalized to that of adult human pancreas for tissue to tissue comparison.

Figure S2: Expression patterns of *CDKAL1*, *CDKN2A*, *CDKN2B*, *HHEX*, *IGF2BP2*.









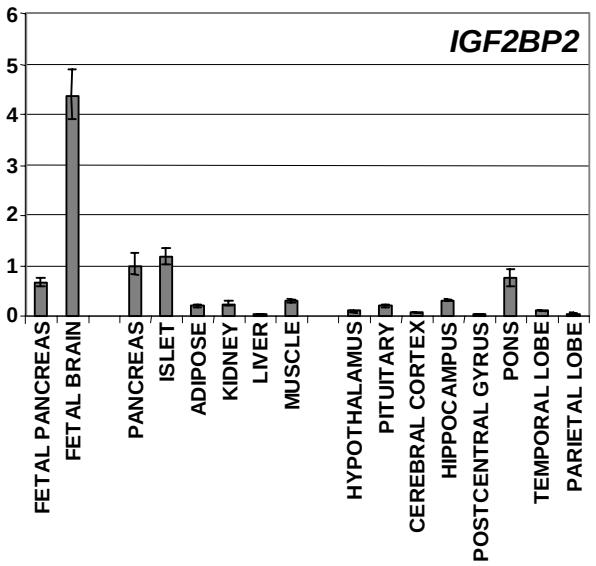


Figure S3: Overview of *HHEX* signal region

A Plot of -log(p values) for T2D (Cochran-Armitage test for trend) against chromosome position in Mb. Blue diamonds represent primary scan results and pink triangles denote meta-analysis results across all UK samples. Note that rs5015480 was typed in the WTCCC scan and rs1111875 in the replication set, so the meta-analysis result is based on a combined analysis of the two (r^2=1 in HapMap CEU) and the position of this signal denoted at both locations.

B Genomic location of genes showing intron and exon structure (NCBI Build 35). Pink triangles show position of replication SNPs relative to gene structure.

C MULTIZ (*S24*) vertebrate alignment of 17 species showing evolutionary conservation.

D GoldSurfer2 (*S19*) plot of linkage disequilibrium (r^2) for SNPs genotyped in WTCCC scan (passing T2D-specific quality control) in WTCCC T2D cases.

E Recombination rate given as cM/MB. Red boxes represent recombination hotspots (*S25*).

F GoldSurfer2 plot of linkage disequilibrium (r^2) for all HapMap SNPs across the region (HapMap CEU data).

Figure S3: Overview of *HHEX* signal region

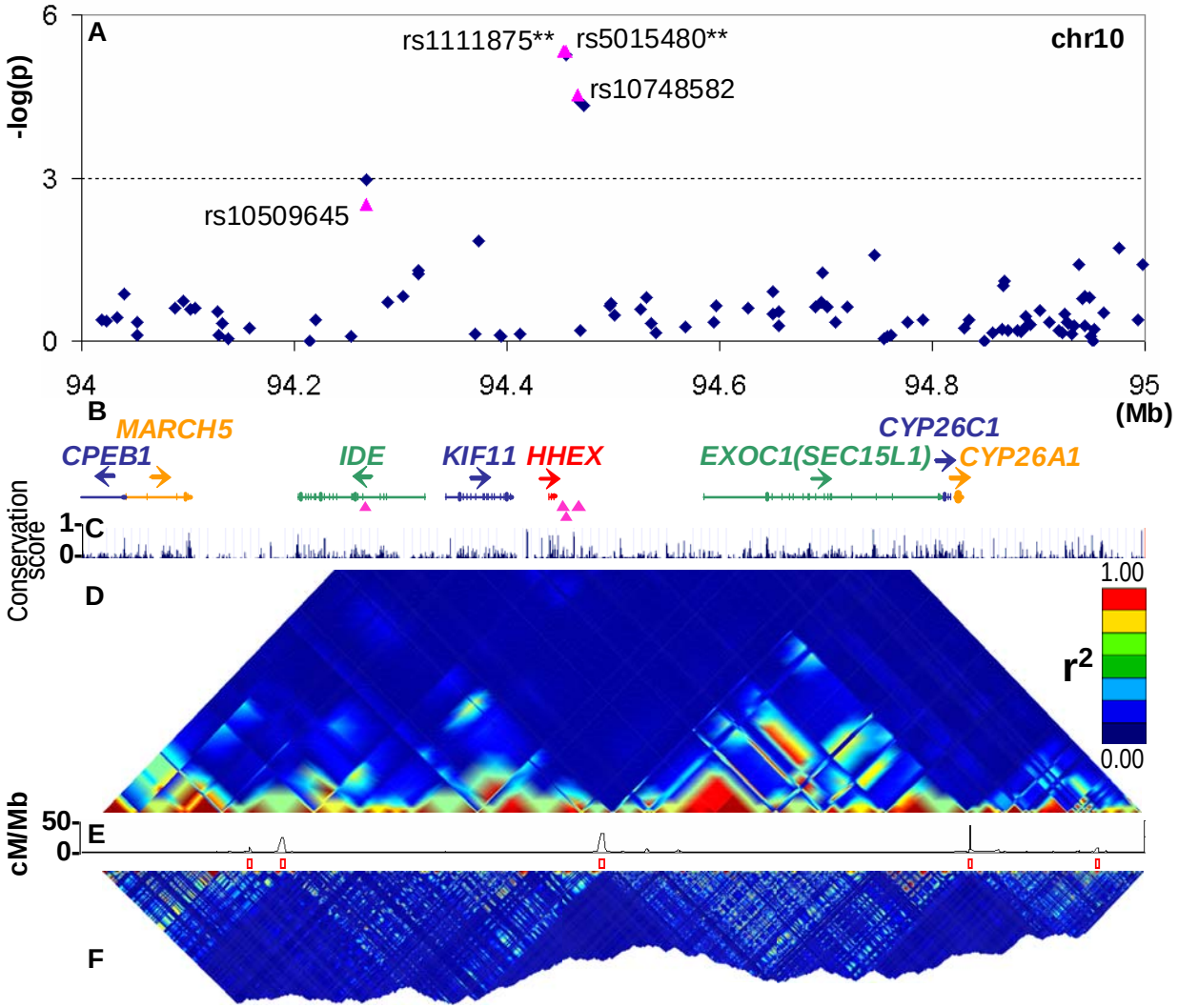


Figure S4: Overview of *IGF2BP2* signal region

A Plot of -log(p values) for T2D (Cochran-Armitage test for trend) against chromosome position in Mb. Blue diamonds represent primary scan results and pink triangles denote meta-analysis results across all UK samples.

B Genomic location of genes showing intron and exon structure (NCBI Build 35). Pink triangles show position of replication SNPs relative to gene structure.

C MULTIZ (*S24*) vertebrate alignment of 17 species showing evolutionary conservation.

D GoldSurfer2 (*S19*) plot of linkage disequilibrium (r^2) for SNPs genotyped in WTCCC scan (passing T2D-specific quality control) in WTCCC T2D cases.

E Recombination rate given as cM/MB. Red boxes represent recombination hotspots (*S25*).

F GoldSurfer2 plot of linkage disequilibrium (r^2) for all HapMap SNPs across the region (HapMap CEU data).

Figure S4: Overview of *IGF2BP2* signal region

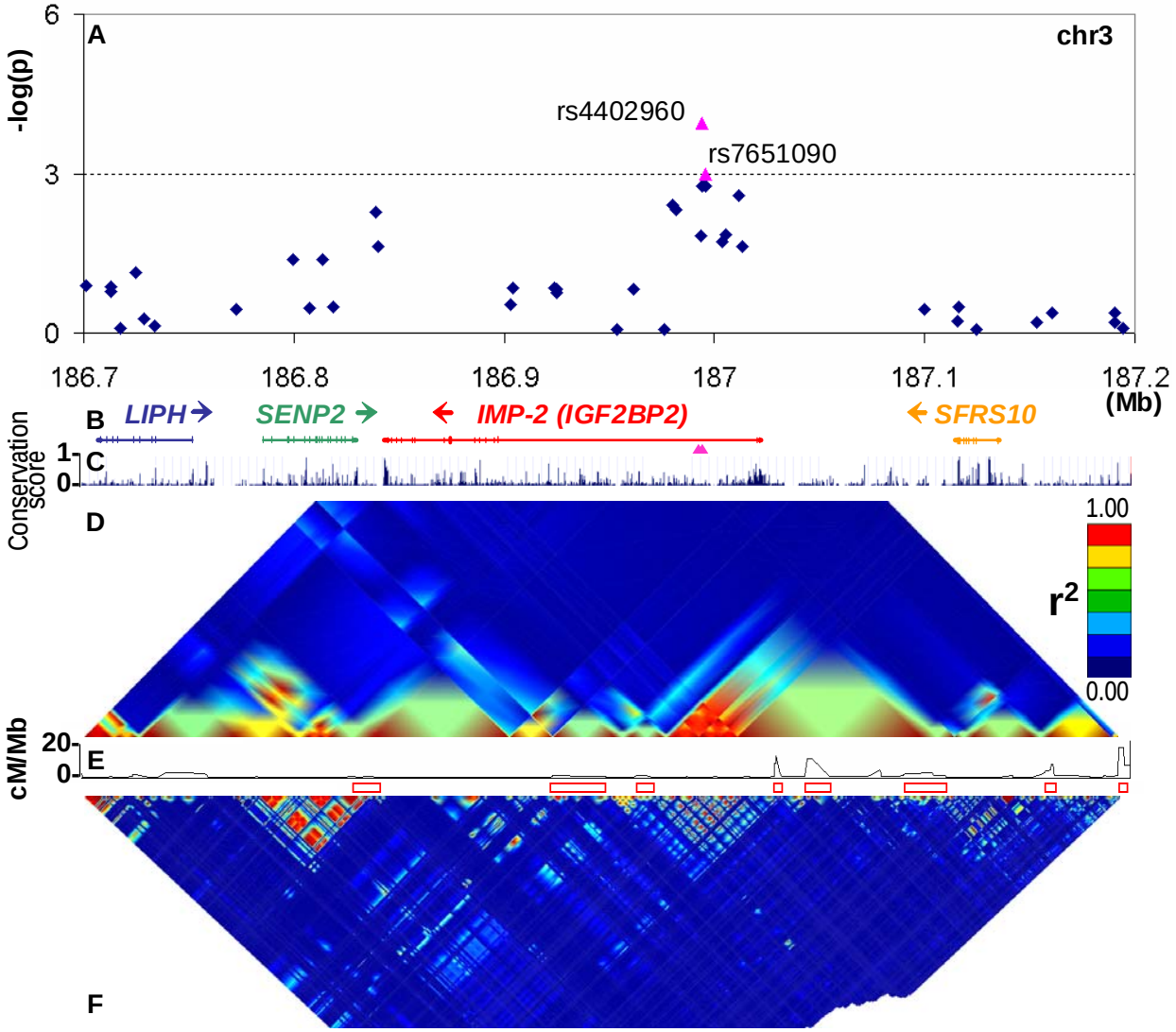


Figure S5: Overview of *VEGFA* signal region

A Plot of -log(p values) for T2D (Cochran-Armitage test for trend) against chromosome position in Mb. Blue diamonds represent primary scan results and pink triangles denote meta-analysis results across all UK samples.

B Genomic location of genes showing intron and exon structure (NCBI Build 35). Pink triangles show position of replication SNPs relative to gene structure.

C MULTIZ (*S24*) vertebrate alignment of 17 species showing evolutionary conservation.

D GoldSurfer2 (*S19*) plot of linkage disequilibrium (r^2) for SNPs genotyped in WTCCC scan (passing T2D-specific quality control) in WTCCC T2D cases.

E Recombination rate given as cM/MB. Red boxes represent recombination hotspots (*S25*).

F GoldSurfer2 plot of linkage disequilibrium (r^2) for all HapMap SNPs across the region (HapMap CEU data).

Figure S5: Overview of *VEGFA* signal region

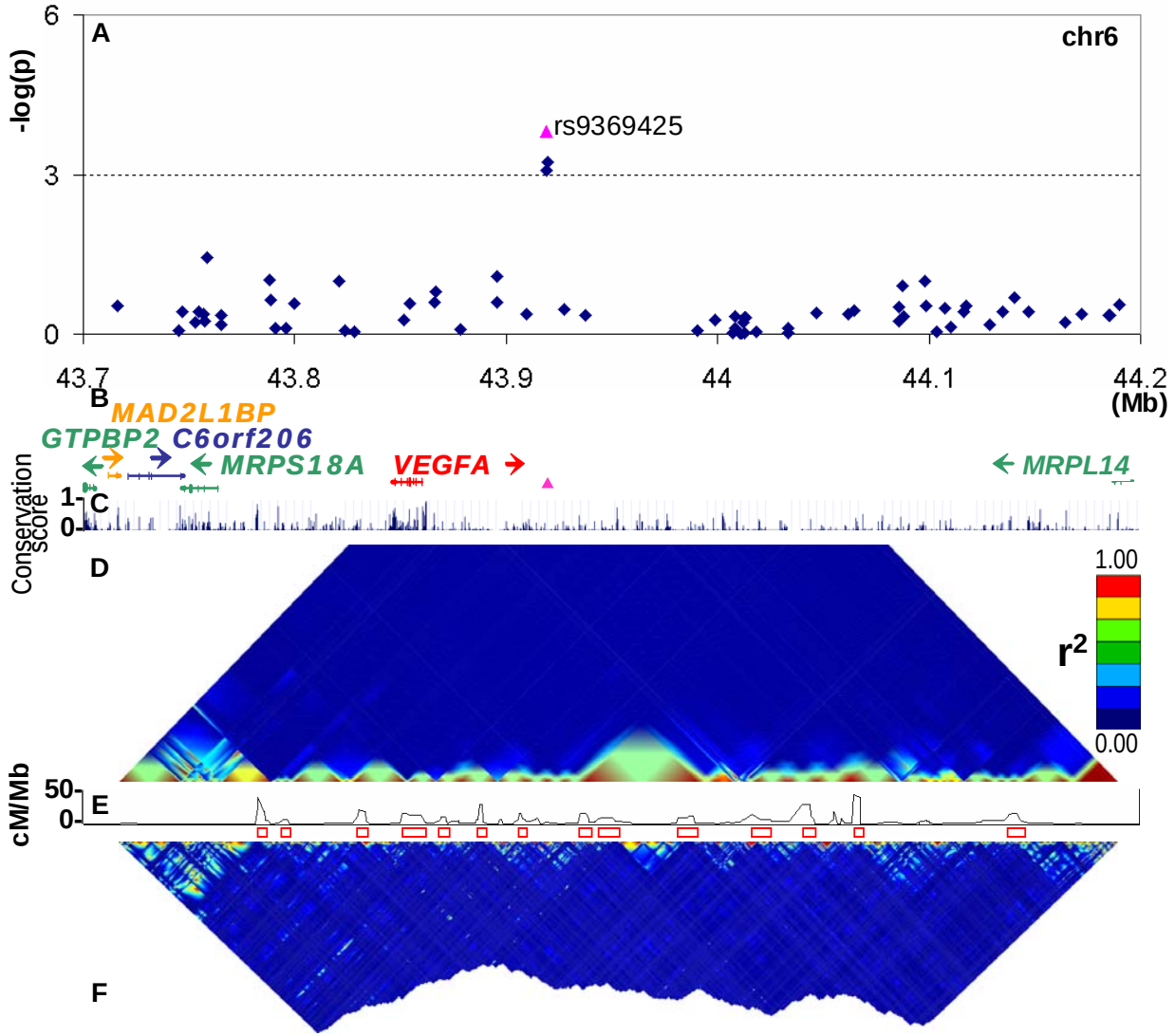


Figure S6: Quantile-Quantile plots for the 393,453 SNPs passing T2D-specific quality control. The plot compares the observed and expected distributions for the 1df chi-square statistics generated from single-point Cochran-Armitage tests.

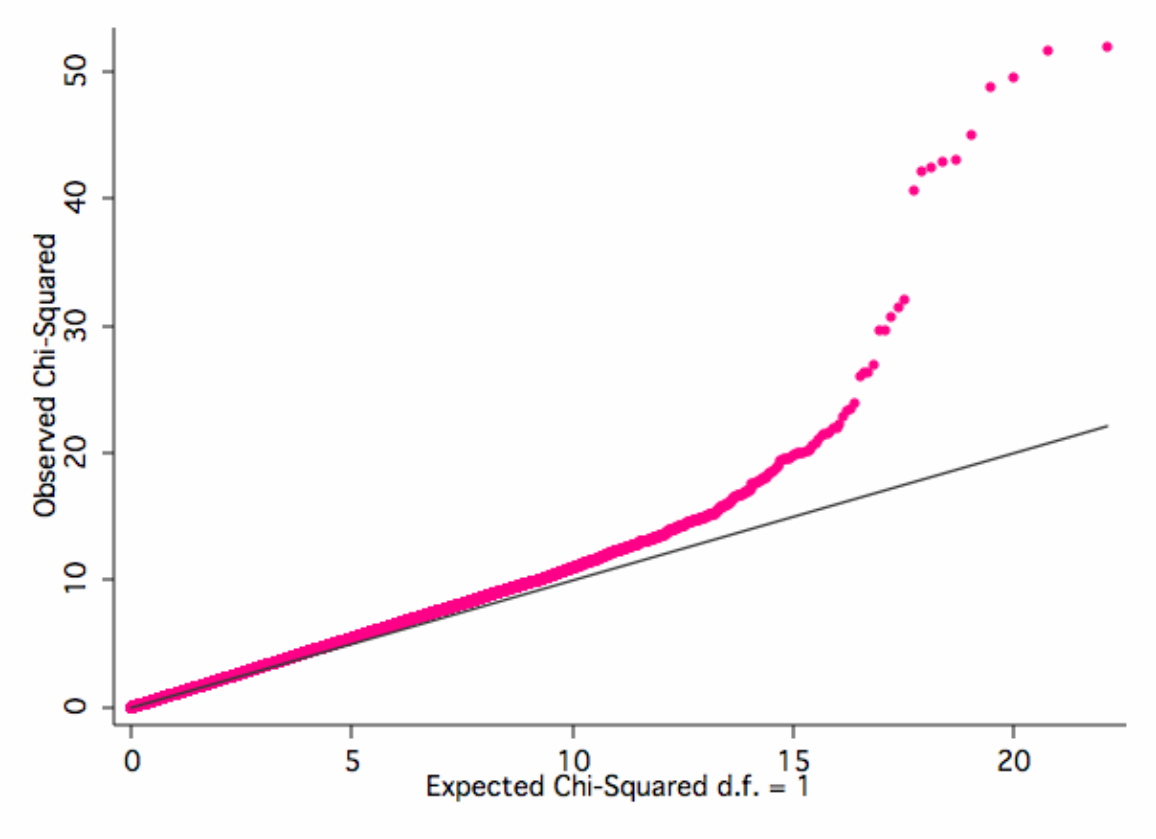


Figure S7: Correlation plot of association statistics for the WTCCC scan genotypes before and after adjusting for population structure. The plot compares the 1df χ^2 (single-point Cochran-Armitage) values obtained from a naïve trend test for the 393.453 SNPs passing T2D-specific quality control, with the equivalent statistics generated after correcting for population substructure using Eigenstrat (*S8*). The high correlation overall (r^2>0.99), and in particular the strong linearity for high χ^2 values (top right) indicates minimal confounding from population substructure after implementation of the various QC measures described.

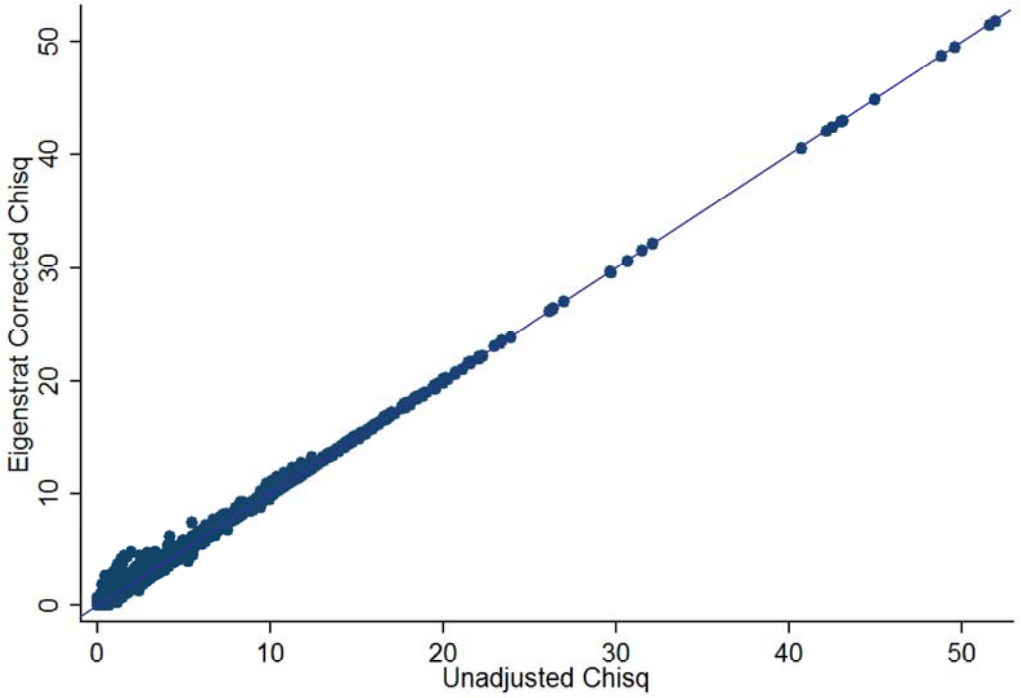


Figure S8: Single-point and haplotype-based analysis results for the chr9 signal region using GENEBPM. Circles denote single-point analysis results and the continuous line represents 5 SNP sliding window haplotype-based analyses using GENEBPM (*S15*). The two peaks of association are separated by a recombination hotspot (**Figure 2**). Multipoint analyses reveal much stronger evidence for association at the 3' peak in this region. Common (>0.01) high risk haplotypes in the 3' signal share alleles T and T at SNPs rs10811661 and rs10757283 respectively. In the WTCCC scan (*S1*), Cochran-Armitage p values for rs2891169 were similar to those of SNPs rs10811661 and rs10757283, which were selected for replication. In contrast, GENEBPM analysis indicated stronger evidence for single-point association for rs2891169.

Figure S8: Single-point and haplotype-based analysis results for the chr9 signal region using GENEBPM.

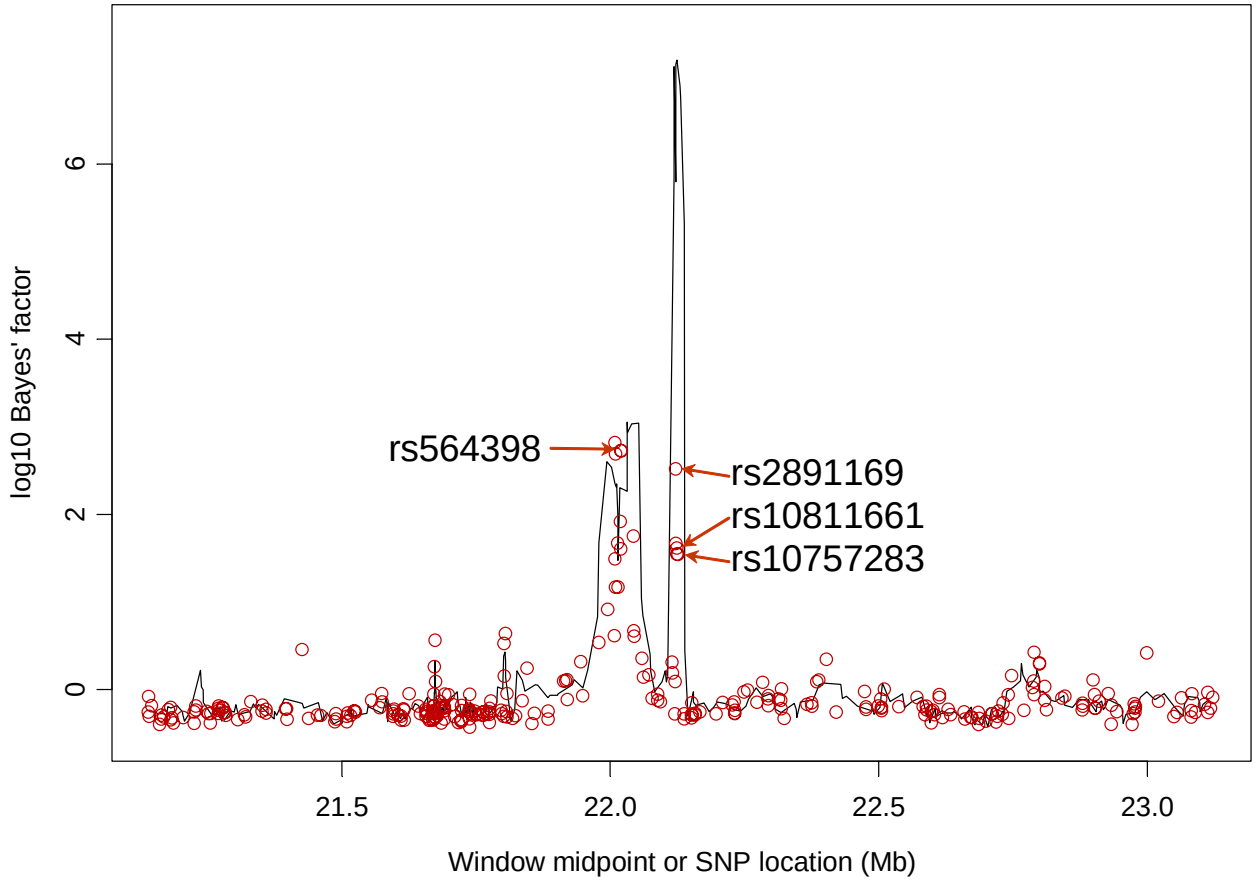


Table S1: Clinical characteristics of UK samples. WTCCC controls came from two sources. No data on age, waist circumference or BMI are available for the UK Blood Service controls. Control individuals from the 1958 Birth Cohort were last reviewed at age 41. Under the terms of access, waist circumference and BMI values from these controls are not available to WTCCC researchers. Only 46% (all male) of the RS2 control individuals had available waist circumference measures.

	Cohort	N	Male participants (%)	Age of diagnosis (mean, SD)	Age at study (mean, SD)	Waist circumference (cm) (mean, SD)		Body mass index (kg/m^2) (geometric mean, SD range)
						Male	Female	
T2D Cases								
	WTCCC cases	1924	58.1	50.3 (9.2)	58.6 (10.1)	106.2 (13.4)	102.5 (15.4)	30.7 (25.4, 37.1)
	RS1 cases	2022	57.9	55.7 (8.9)	64.2 (9.4)	107.1 (13.1)	101.1 (13.7)	31.0 (25.9, 37.2)
	RS2 cases	632	58.3	48.1 (9.1)	58.4 (9.0)	108.1 (13.6)	104.5 (14.6)	31.5 (26.1, 38.0)
	RS3 cases	1103	54.2	55.3 (9.1)	63.7 (9.9)	107.5 (13.3)	103.0 (15.3)	31.2 (25.6, 38.0)
Controls								
	WTCCC controls	2938	49.2	-	-	-	-	-
	RS1 controls	2037	51.2	-	58.7 (12.0)	97.6 (10.6)	85.4 (12.3)	26.3 (22.4, 30.9)
	RS2 controls	1750	48.9	-	31.6 (5.7)	92.0 (11.0)	-	25.0 (21.1, 29.5)
	RS3 controls	1559	52.6	-	59.0 (11.8)	98.2 (10.9)	87.5 (12.9)	26.6 (22.7, 31.2)

Table S2: Replication results for SNPs selected for Cochran-Armitage $p<10^{-5}$ on WTCCC scan. Alleles in this table are named alphabetically (as per the forward strand) with the ancestral allele underlined (where known). A2F denotes allele 2 frequency. For consistency, all ORs in this table are reported for allele 2 (and may therefore be the reciprocal of the ORs reported in the text and Table 1). It was not possible to design a working replication assay for rs5015480, and the UK meta-analysis for this signal combines data from rs5015480 and rs1111875 ($r^2=1$ in HapMap CEU).

Table S2: Replication results for SNPs selected for Cochran-Armitage $p<10^{-5}$ on WTCCC scan.

					WTCCC				RS1			RS2			RS3			Replication meta-analysis		All UK sample meta-analysis	
rs	chr	position	A1	A2	A2F cases	A2F controls	OR (95%CIs)	P_{add}	A2F cases	A2F controls	OR (95%CIs)	A2F cases	A2F controls	OR (95%CIs)	A2F cases	A2F controls	OR (95%CIs)	OR (95%CIs)	P_{add}	OR (95%CIs)	P_{add}
rs4655595	1	66270780	A	G	0.108	0.080	1.39 (1.21-1.60)	$2.7x10^{-6}$	0.078	0.073	1.08 (0.91-1.27)	0.101	0.091	1.13 (0.90-1.41)	-	-	-	1.09 (0.96-1.25)	0.19	1.23 (1.12-1.36)	$2.8x10^{-5}$
rs6718526	2	161039682	C	T	0.171	0.209	0.78 (0.70-0.86)	$2.4x10^{-6}$	0.207	0.207	1.00 (0.90-1.12)	0.202	0.192	1.06 (0.90-1.25)	-	-	-	1.02 (0.93-1.12)	0.69	0.91 (0.85-0.97)	$4.8x10^{-3}$
rs10946398	6	20769013	A	C	0.361	0.319	1.20 (1.10-1.31)	$2.5x10^{-5}$	0.348	0.328	1.10 (1.00-1.20)	0.354	0.309	1.22 (1.07-1.41)	0.351	0.32	1.15 (1.02-1.30)	1.14 (1.07-1.22)	$8.4x10^{-5}$	1.16 (1.10-1.22)	$1.3x10^{-8}$
rs9460546	6	20771611	G	T	0.640	0.680	0.84 (0.77-0.91)	$3.7x10^{-5}$	0.653	0.672	0.92 (0.84-1.01)	0.645	0.692	0.81 (0.71-0.93)	-	-	-	0.88 (0.82-0.95)	$1.6x10^{-3}$	0.86 (0.81-0.91)	$3.1x10^{-7}$
rs9465871	6	20825234	C	T	0.782	0.822	0.78 (0.70-0.86)	$1.0x10^{-6}$	0.808	0.824	0.90 (0.80-1.01)	0.794	0.813	0.89 (0.76-1.05)	-	-	-	0.90 (0.82-0.99)	0.023	0.84 (0.78-0.90)	$6.4x10^{-7}$
rs7767391	6	20833219	C	T	0.781	0.818	0.79 (0.71-0.88)	$5.5x10^{-6}$	0.809	0.823	0.91 (0.81-1.02)	0.794	0.810	0.90 (0.77-1.06)	-	-	-	0.91 (0.83-1.00)	0.042	0.85 (0.80-0.91)	$4.8x10^{-6}$
rs9326506	10	43388564	A	C	0.538	0.492	1.21 (1.11-1.31)	$7.8x10^{-6}$	0.512	0.514	0.99 (0.91-1.09)	0.502	0.518	0.94 (0.82-1.07)				0.98 (0.90-1.05)	0.50	1.07 (1.02-1.14)	0.012
rs10509645	10	94267846	A	C	0.681	0.650	1.16 (1.06-1.27)	$1.1x10^{-3}$	0.665	0.664	1.01 (0.92-1.11)	0.670	0.644	1.12 (0.98-1.30)	-	-	-	1.04 (0.96-1.12)	0.29	1.09 (1.03-1.16)	$3.0x10^{-3}$
rs1111875	10	94452862	C	T	-	-	-	-	0.381	0.384	0.99 (0.90-1.08)	0.373	0.412	0.85 (0.75-0.97)	0.371	0.396	0.90 (0.80-1.01)	0.93 (0.87-0.99)	0.020	0.89 (0.84-0.93)	$4.6x10^{-6}$
rs5015480	10	94455539	C	T	0.379	0.425	0.82 (0.75-0.89)	$5.4x10^{-6}$	-	-	-	-	-	-	-	-	-	-	-		
rs10748582	10	94467199	A	T	0.660	0.619	1.20 (1.10-1.31)	$3.7x10^{-5}$	0.633	0.629	1.02 (0.93-1.12)	0.658	0.612	1.21 (1.06-1.39)	-	-	-	1.08 (1.00-1.16)	0.057	1.13 (1.07-1.20)	$3.0x10^{-5}$
rs11178531	12	69694957	A	G	0.500	0.546	0.83 (0.77-0.90)	$1.0x10^{-5}$	0.532	0.552	0.92 (0.84-1.01)	0.546	0.533	1.05 (0.93-1.20)	-	-	-	0.96 (0.89-1.04)	0.31	0.90 (0.85-0.95)	$2.0x10^{-4}$
rs10506625	12	69808664	C	G	0.313	0.352	0.84 (0.77-0.91)	$7.2x10^{-5}$	0.341	0.345	0.98 (0.90-1.08)	0.326	0.345	0.92 (0.80-1.05)	-	-	-	0.96 (0.89-1.04)	0.33	1.06 (1.00-1.12)	0.057
rs7972188	12	69853021	C	T	0.632	0.588	1.21 (1.11-1.32)	$9.8x10^{-6}$	0.601	0.599	1.01 (0.92-1.10)	0.628	0.606	1.10 (0.96-1.25)	-	-	-	1.03 (0.96-1.12)	0.38	1.11 (1.05-1.17)	$3.2x10^{-4}$
rs1495377	12	69863368	C	G	0.547	0.497	1.23 (1.13-1.33)	$1.3x10^{-6}$	0.505	0.499	1.03 (0.94-1.12)	0.530	0.511	1.08 (0.94-1.23)	-	-	-	1.04 (0.97-1.12)	0.28	1.12 (1.06-1.18)	$5.2x10^{-5}$
rs7961581	12	69949369	C	T	0.681	0.725	0.81 (0.74-0.89)	$3.6x10^{-6}$	0.707	0.721	0.93 (0.85-1.03)	0.703	0.715	0.94 (0.82-1.09)	-	-	-	0.94 (0.86-1.02)	0.12	0.88 (0.83-0.93)	$1.9x10^{-5}$
rs2930291	15	72391887	A	G	0.668	0.623	1.22 (1.12-1.33)	$7.7x10^{-6}$	0.625	0.634	0.96 (0.88-1.06)	0.620	0.636	0.93 (0.81-1.08)	-	-	-	0.96 (0.89-1.03)	0.25	1.06 (1.01-1.13)	0.034
rs2903265	15	78200439	A	G	0.757	0.716	1.23 (1.12-1.35)	$9.6x10^{-6}$	0.739	0.729	1.05 (0.95-1.16)	0.733	0.736	0.98 (0.85-1.14)	-	-	-	1.03 (0.95-1.12)	0.50	1.11 (1.05-1.19)	$5.2x10^{-4}$
rs7193144	16	52368187	C	T	0.546	0.603	0.79 (0.72-0.85)	$1.4x10^{-8}$	0.568	0.615	0.82 (0.75-0.90)	0.570	0.613	0.83 (0.72-0.96)	-	-	-	0.83 (0.76-0.89)	$9.5x10^{-7}$	0.81 (0.76-0.85)	$9.6x10^{-14}$
rs8050136	16	52373776	A	C	0.545	0.602	0.79 (0.73-0.86)	$2.0x10^{-8}$	0.571	0.619	0.82 (0.75-0.90)	0.571	0.613	0.84 (0.73-0.96)	-	-	-	0.82 (0.76-0.89)	$5.4x10^{-7}$	0.81 (0.76-0.85)	$7.3x10^{-14}$
rs9939609	16	52378028	A	T	0.547	0.602	0.79 (0.73-0.86)	$5.4x10^{-8}$	0.569	0.615	0.82 (0.75-0.90)	0.564	0.613	0.81 (0.71-0.93)	0.589	0.592	0.99 (0.88-1.10)	0.87 (0.81-0.92)	$9.4x10^{-6}$	0.84 (0.80-0.88)	$9.0x10^{-12}$

Table S3: Confirmed T2D susceptibility signals: SNPs reported for the DGI and FUSION studies. As DGI and FUSION did not always type the same SNPs as the UK study in all their samples, results in Table 1 include data from the SNPs generating the strongest association in their respective studies. Table S3 gives details of the SNPs reported for DGI and FUSION, and their LD relationships (based on HapMap CEU and/or genome-wide or imputation data as available) with the UK index SNP. In all cases these proxies were SNPs in strong LD (r^2>0.95, except *TCF7L2*) and showed consistent direction of effect with the SNP reported in the UK data.

UK				DGI				FUSION			
				SNP	r^2	DGI dataset Table 1 results		**SNP**	r^2	FUSION dataset Table 1 results	
rs	chr	position	region		with UK SNP	GWA scan	replication data		with UK SNP	stage 1	stage 2
rs8050136	16	52373776	*FTO*	rs8050136	identical	√	not typed	rs8050136	identical	√	√
rs10946398	6	20769013	*CDKAL1*	rs7754840	>0.99 (WTCCC GWA and DGI GWA)	√	√	rs7754840	>0.99 (WTCCC GWA) 1.00 (FUSION GWA)	√	√
rs5015480	10	94455539	*HHEX*	rs1111875	1.00 (HapMap CEU)	√	√	rs1111875	1.00 (HapMap CEU and FUSION imputation)	√	√
rs1111875	10	94452862	*HHEX*	rs1111875	identical			rs1111875	identical		
rs10811661	9	22124094	*CDKN2B*	rs10811661	identical	√	√	rs10811661	identical	√	√
rs564398	9	22019547	*CDKN2B*	rs564398	identical	√	not typed	rs1412829	0.97 (WTCCC GWA)	√	not typed
rs4402960	3	186994389	*IGF2BP2*	rs4402960	identical	√	√	rs4402960	identical	√	√
rs13266634	8	118254964	*SLC30A8*	rs13266634	identical	√	√	rs13266634	identical	√	√
rs7901695	10	114744078	*TCF7L2*	rs4506565	0.98 (WTCCC GWA) 0.99 (DGI GWA)	√	√	rs7903146	0.80 (HapMap CEU) 0.85 (FUSION GWA)	√	√
rs5215	11	17365206	*KCNJ11*	rs5215	identical	√	√	rs5219	>0.99 (FUSION GWA)	√	√
rs1801282	3	12368125	*PPARG*	rs1801282	identical	√	√	rs1801282	identical	√	√

Table S4: Replication results for SNPs selected for the "second wave" of replication. Alleles in this table are named alphabetically (as per the forward strand) with the ancestral allele underlined (where known). A2F denotes allele 2 frequency. For consistency, all ORs in this table are reported for allele 2 (and may therefore be the reciprocal of the ORs reported in the text). These signals were (with the exception of rs13266634 in *SLC30A8*) selected on the basis of Cochran-Armitage test p values between 10^{-2} and 10^{-5} on the WTCCC scan, prioritized on the basis of biological candidacy, multiple independent associations and/or support from DGI and/or FUSION scans. It was not possible to design a working replication assay for rs11140802, and the UK meta-analysis for this signal combines data from rs11140802 and rs12346884 (r^2=1 in HapMap CEU). Rs13266634 in *SLC30A8* was not captured on the Affymetrix chip but was selected for replication on the basis of the associations in French (*S18*) and FUSION (*S17*) subjects.

Table S4: Replication results for SNPs selected for the "second wave" of replication

					WTCCC				RS1				RS2				RS3			Replication meta-analysis		All UK sample meta-analysis	
rs	chr	position	A1	A2	A2F cases	A2F controls	OR (95%CIs)	P_{add}	A2F Cases	A2F controls	OR (95%CIs)		A2F cases	A2F controls	OR (95%CIs)		A2F cases	A2F controls	OR (95%CIs)	OR (95%CIs)	P_{add}	OR (95%CIs)	P_{add}
rs10889039	1	57540566	C	T	0.538	0.507	1.12 (1.04-1.22)	$3.4x10^{-3}$	0.510	0.497	1.05 (0.96-1.15)		0.516	0.516	1.00 (0.87-1.15)		-	-	-	1.03 (0.96-1.11)	0.38	1.08 (1.02-1.14)	$8.0x10^{-3}$
rs480075	1	183187993	A	G	0.821	0.849	0.82 (0.73-0.91)	$3.2x10^{-4}$	0.846	0.839	1.05 (0.93-1.19)		0.854	0.838	1.12 (0.94-1.35)		-	-	-	1.08 (0.97-1.19)	0.16	0.95 (0.88-1.02)	0.16
rs2383529	1	183400749	A	G	0.208	0.236	0.85 (0.77-0.94)	$1.2x10^{-3}$	0.216	0.226	0.94 (0.85-1.05)		0.225	0.222	1.02 (0.87-1.19)		-	-	-	0.97 (0.88-1.06)	0.46	0.91 (0.85-0.98)	$6.2x10^{-3}$
rs11679606	2	16717046	A	G	0.376	0.415	0.85 (0.78-0.93)	$1.3x10^{-4}$	0.409	0.406	1.01 (0.93-1.11)		0.401	0.393	1.04 (0.91-1.19)		-	-	-	1.02 (0.94-1.10)	0.58	0.94 (0.89-0.99)	0.032
rs1022377	2	21968138	A	G	0.545	0.517	1.12 (1.03-1.21)	$7.2x10^{-3}$	0.524	0.507	1.07 (0.98-1.17)		0.499	0.515	0.94 (0.81-1.07)		-	-	-	1.03 (0.95-1.11)	0.47	1.07 (1.01-1.13)	0.019
rs4140806	2	37742813	C	T	0.474	0.501	0.90 (0.83-0.97)	$8.5x10^{-3}$	0.510	0.479	1.13 (1.04-1.24)		0.511	0.472	1.17 (1.02-1.34)		-	-	-	1.15 (1.06-1.24)	$3.8x10^{-4}$	1.02 (0.97-1.08)	0.40
rs3814045	2	46650808	A	G	0.369	0.401	0.88 (0.81-0.95)	$2.4x10^{-3}$	0.403	0.381	1.10 (1.00-1.20)		0.408	0.386	1.10 (0.96-1.25)		-	-	-	1.10 (1.02-1.19)	0.015	0.99 (0.93-1.05)	0.81
rs314890	2	125700943	C	T	0.548	0.516	1.14 (1.05-1.24)	$1.6x10^{-3}$	0.514	0.514	1.00 (0.91-1.09)		0.526	0.535	0.96 (0.85-1.10)		-	-	-	0.99 (0.92-1.06)	0.74	1.05 (1.00-1.11)	0.06
rs2217616	2	182825154	G	T	0.322	0.291	1.16 (1.06-1.27)	$1.1x10^{-3}$	0.299	0.305	0.97 (0.88-1.08)		0.302	0.316	0.93 (0.81-1.08)		-	-	-	0.96 (1.04-0.88)	0.32	1.04 (0.99-1.11)	0.14
rs11688935	2	189001411	A	G	0.336	0.295	1.21 (1.11-1.32)	$2.1x10^{-5}$	0.319	0.305	1.07 (0.97-1.18)		0.308	0.306	1.01 (0.87-1.16)		-	-	-	1.05 (0.97-1.14)	0.24	1.12 (1.05-1.19)	$2.0x10^{-4}$
rs17248501	2	205864551	A	C	0.348	0.392	0.83 (0.76-0.90)	$1.4x10^{-5}$	0.374	0.369	1.02 (0.93-1.12)		0.350	0.362	0.95 (0.83-1.09)		-	-	-	1.00 (0.92-1.08)	0.94	0.92 (0.87-0.97)	$2.8x10^{-3}$
rs440646	3	11266171	G	T	0.566	0.607	0.84 (0.77-0.91)	$4.4x10^{-5}$	0.600	0.598	1.01 (0.92-1.11)		0.588	0.596	0.96 (0.84-1.10)		-	-	-	1.00 (0.92-1.07)	0.91	0.92 (0.87-0.98)	$4.7x10^{-3}$
rs4402960	3	186994389	G	T	0.351	0.320	1.15 (1.05-1.25)	$1.7x10^{-3}$	0.338	0.321	1.08 (0.98-1.19)		0.329	0.322	1.03 (0.90-1.19)		0.333	0.307	1.12 (1.00-1.27)	1.09 (1.01-1.16)	0.018	1.11 (1.05-1.16)	$1.6x10^{-4}$
rs7651090	3	186996094	A	G	0.350	0.320	1.15 (1.05-1.25)	$1.8x10^{-3}$	0.338	0.321	1.08 (0.98-1.19)		0.329	0.322	1.03 (0.90-1.19)		-	-	-	1.06 (0.98-1.15)	0.12	1.10 (1.04-1.16)	$1.1x10^{-3}$
rs3118617	4	2307648	C	T	0.52	0.556	0.87 (0.80-0.94)	$5.0x10^{-4}$	0.551	0.537	1.06 (0.97-1.16)		0.559	0.555	1.02 (0.89-1.16)		-	-	-	1.05 (0.97-1.13)	0.24	0.96 (0.91-1.01)	0.14
rs10516948	4	95525785	A	C	0.840	0.817	1.17 (1.05-1.30)	$4.5x10^{-3}$	0.821	0.815	1.04 (0.93-1.17)		0.807	0.825	0.88 (0.74-1.05)		-	-	-	0.99 (0.90-1.09)	0.85	1.07 (0.99-1.15)	0.078
rs7698608	4	104295578	A	C	0.546	0.510	1.16 (1.06-1.25)	$4.6x10^{-4}$	0.526	0.514	1.05 (0.96-1.15)		0.538	0.511	1.11 (0.98-1.27)		0.522	0.529	0.97 (0.87-1.09)	1.04 (0.98-1.10)	0.23	1.08 (1.03-1.14)	$2.2x10^{-3}$
rs2139980	4	157974803	A	G	0.665	0.629	1.17 (1.07-1.27)	$3.1x10^{-4}$	0.635	0.641	0.98 (0.89-1.07)		0.657	0.629	1.13 (0.98-1.29)		-	-	-	1.02 (0.95-1.10)	0.58	1.09 (1.03-1.15)	$4.5x10^{-3}$
rs152189	5	61716205	C	T	0.723	0.751	0.86 (0.79-0.95)	$2.2x10^{-3}$	0.755	0.761	0.97 (0.87-1.08)		0.732	0.740	0.96 (0.83-1.11)		-	-	-	0.96 (0.88-1.05)	0.44	0.92 (0.86-0.98)	$8.0x10^{-3}$
rs1029328	6	28555894	C	T	0.887	0.867	1.20 (1.06-1.36)	$4.8x10^{-3}$	0.858	0.859	0.99 (0.87-1.13)		0.891	0.885	1.07 (0.87-1.32)		-	-	-	1.01 (0.91-1.13)	0.81	1.09 (1.00-1.18)	0.041
rs2242655	6	31735428	C	G	0.881	0.859	1.21 (1.07-1.37)	$2.0x10^{-3}$	0.884	0.874	1.10 (0.96-1.27)		0.871	0.872	0.99 (0.82-1.20)		-	-	-	1.06 (0.95-1.19)	0.27	0.95 (0.87-1.03)	0.20
rs9369425	6	43918952	A	G	0.319	0.287	1.16 (1.06-1.27)	$8.6x10^{-4}$	0.311	0.287	1.12 (1.01-1.23)		0.298	0.290	1.04 (0.90-1.20)		0.293	0.285	1.04 (0.92-1.18)	1.08 (1.01-1.15)	0.033	1.11 (1.05-1.17)	$2.0x10^{-4}$
rs1665901	6	107540093	A	T	0.316	0.357	0.83 (0.76-0.90)	$2.5x10^{-5}$	0.338	0.341	0.99 (0.90-1.09)		0.325	0.348	0.90 (0.78-1.04)		-	-	-	0.96 (0.89-1.04)	0.31	0.90 (0.85-0.95)	$3.5x10^{-4}$
rs7755651	6	148766018	C	T	0.145	0.123	1.21 (1.07-1.36)	$2.1x10^{-3}$	0.126	0.127	0.99 (0.86-1.13)		0.118	0.134	0.86 (0.66-1.13)		-	-	-	0.96 (0.85-1.08)	0.50	1.08 (0.99-1.17)	0.087

rs864745	7	27953796	C	T	0.529	0.490	1.16 (1.08-1.27)	$1.8x10^{-4}$	0.510	0.499	1.04 (0.95-1.14)	0.547	0.498	1.22 (1.06-1.41)	0.485	0.491	0.98 (0.88-1.10)	1.05 (0.99-1.12)	0.074	1.10 (1.04-1.15)	$2.3x10^{-4}$
rs12702146	7	45294288	C	T	0.523	0.487	1.16 (1.07-1.25)	$5.2x10^{-4}$	0.495	0.501	0.98 (0.89-1.07)	0.535	0.497	1.17 (1.02-1.33)	-	-	-	1.03 (0.96-1.11)	0.40	1.09 (1.03-1.15)	$3.1x10^{-3}$
rs4493865	7	54672611	A	T	0.213	0.246	0.84 (0.76-0.92)	$2.6x10^{-4}$	0.249	0.238	1.06 (0.96-1.18)	0.227	0.225	1.01 (0.87-1.18)	-	-	-	1.05 (0.96-1.14)	0.30	0.95 (0.89-1.01)	0.09
rs7007919	8	20691655	C	T	0.613	0.580	1.14 (1.05-1.24)	$1.6x10^{-3}$	0.596	0.584	1.05 (0.96-1.15)	0.600	0.589	1.05 (0.92-1.19)	-	-	-	1.05 (0.97-1.13)	0.21	1.09 (1.03-1.15)	$2.2x10^{-3}$
rs16879809	8	32692415	C	T	0.135	0.154	0.86 (0.77-0.97)	$1.3x10^{-2}$	0.163	0.162	1.00 (0.89-1.13)	0.152	0.148	1.03 (0.86-1.25)	-	-	-	1.01 (0.91-1.12)	0.82	0.95 (0.88-1.02)	0.15
rs16875331	8	108008872	A	G	0.891	0.868	1.24 (1.09-1.41)	$8.7x10^{-4}$	0.871	0.876	0.96 (0.84-1.10)	0.894	0.874	1.23 (0.99-1.54)	-	-	-	1.03 (0.91-1.15)	0.67	1.12 (1.03-1.22)	0.011
rs13266634	8	118253964	C	T	0.282	0.306	0.89 (0.81-0.98)	0.020	0.283	0.303	0.91 (0.82-1.00)	0.285	0.317	0.86 (0.74-1.00)	0.308	0.331	0.90 (0.80-1.01)	0.89 (0.84-0.96)	$1.2x10^{-3}$	0.89 (0.85-0.95)	$7.0x10^{-5}$
rs939024	8	134655895	A	G	0.726	0.755	0.86 (0.79-0.94)	$1.8x10^{-3}$	0.738	0.739	0.99 (0.89-1.10)	0.757	0.738	1.10 (0.94-1.28)	-	-	-	1.02 (0.94-1.11)	0.56	0.94 (0.89-1.01)	0.093
rs10511667	9	18979696	A	G	0.147	0.128	1.18 (1.04-1.32)	$6.9x10^{-3}$	0.135	0.143	0.93 (0.82-1.06)	0.124	0.137	0.89 (0.74-1.09)	-	-	-	0.92 (0.83-1.03)	0.14	1.03 (0.95-1.11)	0.45
rs564398	9	22019547	C	T	0.595	0.558	1.16 (1.07-1.27)	$3.2x10^{-4}$	0.603	0.569	1.16 (1.05-1.28)	0.588	0.576	1.05 (0.92-1.20)	0.597	0.573	1.10 (0.99-1.24)	1.12 (1.05-1.19)	$8.6x10^{-4}$	1.13 (1.08-1.19)	$1.3x10^{-6}$
rs2383208	9	22122076	A	G	0.141	0.166	0.82 (0.73-0.92)	$8.0x10^{-4}$	0.144	0.174	0.80 (0.71-0.90)	0.146	0.153	0.95 (0.79-1.14)	-	-	-	0.84 (0.76-0.93)	$9.4x10^{-4}$	0.83 (0.77-0.90)	$2.6x10^{-6}$
rs10811661	9	22124094	C	T	0.859	0.834	1.22 (1.09-1.37)	$7.6x10^{-4}$	0.861	0.830	1.26 (1.11-1.42)	0.859	0.848	1.09 (0.91-1.31)	0.855	0.841	1.12 (0.96-1.31)	1.18 (1.08-1.28)	$1.7x10^{-4}$	1.19 (1.11-1.28)	$4.9x10^{-7}$
rs10757283	9	22124172	C	T	0.430	0.401	1.13 (1.04-1.23)	$4.4x10^{-3}$	0.422	0.415	1.03 (0.94-1.12)	0.413	0.403	1.04 (0.91-1.19)	-	-	-	1.03 (0.96-1.11)	0.44	1.07 (1.02-1.14)	0.013
rs10780690	9	84707726	G	T	0.545	0.511	1.15 (1.06-1.24)	$1.1x10^{-3}$	0.524	0.525	1.00 (0.91-1.09)	0.523	0.518	1.02 (0.89-1.16)	-	-	-	1.00 (0.93-1.08)	0.96	1.07 (1.01-1.13)	0.025
rs12346884	9	84739293	A	C	-	-	-	-	0.136	0.144	0.94 (0.83-1.07)	0.149	0.136	1.11 (0.93-1.34)	-	-	-	0.99 (0.89-1.10)	0.89	1.09 (1.00-1.18)	0.046
rs11140802	9	84741930	C	G	0.140	0.118	1.23 (1.08-1.39)	$1.2x10^{-3}$	-	-	-	-	-	-	-	-	-	-	-		
rs4743611	9	96011963	C	T	0.447	0.416	1.14 (1.05-1.23)	$2.6x10^{-3}$	0.439	0.441	0.99 (0.90-1.09)	0.447	0.442	1.02 (0.89-1.16)	-	-	-	1.00 (0.93-1.08)	0.99	1.06 (1.00-1.12)	0.044
rs10120268	9	108627659	C	T	0.316	0.354	0.84 (0.77-0.92)	$1.0x10^{-4}$	0.347	0.338	1.04 (0.95-1.15)	0.338	0.343	0.98 (0.85-1.12)	-	-	-	1.02 (0.94-1.10)	0.60	0.94 (0.88-0.99)	0.027
rs10817674	9	114602134	A	C	0.667	0.626	1.20 (1.10-1.30)	$4.2x10^{-5}$	0.638	0.634	1.02 (0.93-1.11)	0.651	0.652	0.99 (0.87-1.14)	-	-	-	1.01 (0.93-1.09)	0.83	1.09 (1.03-1.15)	$3.7x10^{-3}$
rs10818930	9	124070739	G	T	0.197	0.172	1.18 (1.06-1.31)	$1.9x10^{-3}$	0.174	0.166	1.06 (0.94-1.19)	0.191	0.180	1.07 (0.91-1.27)	-	-	-	1.06 (0.97-1.17)	0.22	1.12 (1.04-1.20)	$2.6x10^{-3}$
rs7023712	9	126328516	A	G	0.809	0.830	0.86 (0.78-0.96)	$6.7x10^{-3}$	0.831	0.836	0.96 (0.85-1.09)	0.794	0.814	0.88 (0.74-1.04)	-	-	-	0.93 (0.85-1.03)	0.17	0.90 (0.84-0.97)	$4.5x10^{-3}$
rs505699	10	99764066	C	T	0.351	0.316	1.18 (1.08-1.28)	$2.9x10^{-4}$	0.326	0.322	1.02 (0.93-1.12)	0.328	0.309	1.09 (0.95-1.25)	-	-	-	1.04 (0.96-1.12)	0.31	1.10 (1.04-1.16)	$1.5x10^{-3}$
rs11597086	10	101943695	A	C	0.410	0.447	0.85 (0.79-0.93)	$2.1x10^{-4}$	0.423	0.449	0.90 (0.82-0.98)	0.426	0.448	0.91 (0.80-1.04)	-	-	-	0.90 (0.84-0.97)	$7.7x10^{-3}$	0.88 (0.83-0.93)	$8.0x10^{-6}$
rs3802678	10	104105385	A	T	0.237	0.204	1.22 (1.10-1.34)	$9.6x10^{-5}$	0.211	0.207	1.03 (0.92-1.14)	0.228	0.219	1.05 (0.90-1.23)	-	-	-	1.03 (0.95-1.13)	0.46	1.11 (1.04-1.19)	$1.6x10^{-3}$
rs1153188	12	53385263	A	T	0.240	0.265	0.88 (0.80-0.97)	$6.7x10^{-3}$	0.258	0.269	0.95 (0.85-1.05)	0.242	0.253	0.94 (0.80-1.10)	-	-	-	0.94 (0.87-1.03)	0.19	0.91 (0.86-0.97)	$5.0x10^{-3}$
rs4426172	12	126483613	C	T	0.377	0.416	0.85 (0.78-0.92)	$1.3x10^{-4}$	0.409	0.400	1.04 (0.95-1.14)	0.395	0.387	1.04 (0.90-1.19)	-	-	-	1.04 (0.96-1.12)	0.31	0.95 (0.90-1.00)	0.069
rs7323267	13	40102015	C	T	0.816	0.789	1.18 (1.06-1.30)	$1.4x10^{-3}$	0.782	0.787	0.97 (0.87-1.09)	0.789	0.807	0.89 (0.76-1.05)	-	-	-	0.94 (0.86-1.04)	0.24	1.04 (0.97-1.11)	0.21
rs2268974	14	68492917	A	G	0.278	0.251	1.15 (1.05-1.27)	$3.1x10^{-3}$	0.259	0.256	1.01 (0.92-1.12)	0.267	0.259	1.04 (0.89-1.20)	-	-	-	1.02 (0.94-1.11)	0.60	1.08 (1.01-1.15)	0.018

rs4343209	14	98286712	A	G	0.375	0.415	0.85 (0.78-0.92)	$9.4x10^{-5}$	0.403	0.409	0.98 (0.89-1.07)	0.382	0.395	0.94 (0.82-1.08)	-	-	-	0.97 (0.90-1.04)	0.38	0.91 (0.86-0.96)	$1.0x10^{-3}$
rs2099106	16	9288209	C	T	0.275	0.236	1.23 (1.12-1.34)	$1.7x10^{-5}$	0.254	0.248	1.03 (0.93-1.14)	0.247	0.254	0.97 (0.83-1.12)	-	-	-	1.01 (0.93-1.10)	0.84	1.10 (1.04-1.17)	$2.3x10^{-3}$
rs6038317	20	5813636	A	C	0.512	0.544	0.88 (0.81-0.95)	$1.7x10^{-3}$	0.544	0.534	1.05 (0.96-1.14)	0.532	0.534	0.99 (0.86-1.13)	-	-	-	1.03 (0.95-1.11)	0.48	0.96 (0.91-1.01)	0.11
rs2154490	21	29837833	A	G	0.745	0.778	0.83 (0.76-0.92)	$1.9x10^{-4}$	0.728	0.751	0.88 (0.80-0.98)	0.763	0.764	0.99 (0.85-1.16)	0.730	0.752	0.89 (0.79-1.02)	0.91 (0.85-0.98)	0.01	0.88 (0.83-0.93)	$2.0x10^{-5}$
rs440544	21	35817449	C	T	0.614	0.582	1.15 (1.05-1.25)	$1.6x10^{-3}$	0.577	0.574	1.02 (0.93-1.11)	0.603	0.602	1.01 (0.88-1.15)	-	-	-	1.01 (0.94-1.09)	0.75	1.07 (1.01-1.13)	0.02

Table S5: Associations between T2D susceptibility variants and (a) BMI, (b) waist circumference in cases and controls. Analyses were performed using linear regression on (a) $\log_{10}$-transformed BMI values and (b) on waist circumference values (in cm) using gender as a covariate. Beta values, 95% CIs and asymptotic p values (t statistic) are reported. Fixed-effects meta-analyses are shown. BMI and waist circumference information was not available for WTCCC controls. In the case of *HHEX*, the UK meta-analysis combines data from rs5015480 and rs1111875 (r^2=1 in HapMap CEU). Rs8050136 was not typed in RS3. At rs13266634, in the WTCCC cases, the common allele (C; T2D risk allele) homozygotes have a waist circumference that is on average 3.2 cm and a BMI that is 2.0 kg/m^2 less than the rare allele (TT) homozygotes.

Table S5: Associations between T2D susceptibility variants and (a) BMI, (b) waist circumference in cases and controls.

(a) BMI

		Rs8050136	rs10946398	rs5015480	rs1111875	rs10811661	rs564398	rs4402960	rs13266634
		Chr16; 52373776 bp	Chr 6; 20769013 bp	Chr 10; 94455539 bp	Chr 10; 94452862 bp	Chr 9; 22124094 bp	Chr 9; 22019547 bp	Chr 3; 186994389 bp	Chr 8; 118253964 bp
	Region	*FTO*	*CDKAL1*	*HHEX*	*HHEX*	*CDKN2B*	*CDKN2B*	*IGF2BP2*	*SLC30A8*
WTCCC cases	Beta value 95% CI (p value)	0.012 (0.007, 0.017) $p=8.7x10^{-6}$	-0.002 (-0.008, 0.003) p=0.38	0.001 (-0.005, 0.006) p=0.82	-	0.002 (-0.006, 0.009) p=0.63	0.000 (-0.005, 0.006) p=0.93	0.002 (-0.003, 0.008) p=0.39	0.009 (0.003, 0.015) $p=5.4x10^{-3}$
RS1 cases	Beta value 95% CI (p value)	0.010 (0.005, 0.015) $p=1.6 x10^{-4}$	-0.001 (-0.007, 0.004) p=0.57	-	0.003 (-0.002, 0.009) p=0.20	0.001 (-0.006, 0.008) p=0.76	0.001 (-0.004, 0.007) p=0.62	-0.001 (-0.006, 0.004) p=0.71	0.004 (-0.001, 0.010) p=0.11
RS2 cases	Beta value 95% CI (p value)	0.013 (0.004, 0.022) $p=4.8 x10^{-4}$	-0.006 (-0.015, 0.004) p=0.22	-	-0.002 (-0.011, 0.007) p=0.68	0.007 (-0.006, 0.019) p=0.31	0.003 (-0.007, 0.012) p=0.59	-0.007 (-0.017, 0.003) p=0.15	-0.001 (-0.012, 0.009) p=0.84
RS3 cases	Beta value 95% CI (p value)	-	-0.002 (-0.009, 0.006) p=0.68	-	0.002 (-0.005, 0.010) p=0.55	0.009 (-0.002, 0.019) p=0.10	-0.008 (-0.015, -0.001) p=0.03	-0.001 (-0.009, 0.007) p=0.84	0.002 (-0.006, 0.009) p=0.70
Meta-analysis cases	Beta value 95% CI (p value)	0.011 (0.008, 0.0141) $p=1.1 x10^{-10}$	-0.002 (-0.005, 0.001) p=0.15	0.002 (-0.001, 0.005) p=0.24		0.003 (-0.001, 0.008) p=0.13	0.000 (-0.003, 0.003) p=0.93	0.000 (-0.004, 0.003) p=0.78	0.005 (0.001, 0.008) $p=9.0 x10^{-3}$
RS1 controls	Beta value 95% CI (p value)	0.004 (-0.001, 0.008) p=0.09	0.001 (-0.003, 0.006) p= 0.54	-	0.002 (-0.006, 0.003) p=0.51	-0.002 (-0.007, 0.004) p=0.61	-0.001 (-0.006, 0.004) p=0.74	-0.004 (-0.009, 0.001) p=0.10	0.003 (-0.002, 0.007) p=0.29
RS2 controls	Beta value 95% CI (p value)	0.009 (0.004, 0.014) $p=5.8x10^{-4}$	-0.002 (-0.007, 0.003) p=0.42	-	-0.007 (-0.012, -0.002) $p=7.5 x10^{-3}$	0.001 (-0.006, 0.008) P=0.70	0.004 (-0.001, 0.009) p=0.14	0.000 (-0.005, 0.005) p=0.95	0.001 (-0.004, 0.006) p=0.73
RS3 controls	Beta value 95% CI (p value)	-	0.002 (-0.003, 0.007) p=0.42	-	0.003 (-0.002, 0.008) p=0.26	0.005 (-0.002, 0.012) p=0.17	0.003 (-0.002, 0.008) p=0.32	0.006 (0.000, 0.011) p=0.04	-0.004 (-0.009, 0.002) p=0.24
Meta-analysis controls	Beta value 95% CI (p value)	0.006 (0.003, 0.010) $p=3.7 x10^{-4}$	0.001 (-0.002, 0.003) p=0.36	-	-0.002 (-0.005, 0.001) p=0.18	0.001 (-0.003, 0.005) p=0.54	0.002 (-0.001, 0.005) p=0.24	0.000 (-0.003, 0.003) p=0.94	0.000 (-0.003, 0.003) p=0.80

(b) waist circumference

		rs8050136	rs10946398	rs5015480	rs1111875	rs10811661	rs564398	rs4402960	rs13266634
		Chr16; 52373776 bp	Chr 6; 20769013 bp	Chr 10; 94455539 bp	Chr 10; 94452862 bp	Chr 9; 22124094 bp	Chr 9; 22019547 bp	Chr 3; 186994389 bp	Chr 8; 118253964bp
	Region	*FTO*	*CDKAL1*	*HHEX*	*HHEX*	*CDKN2B*	*CDKN2B*	*IGF2BP2*	*SLC30A8*
WTCCC cases	Beta value 95% CI (p value)	1.584 (0.654, 2.514) $p=8.6x10^{-4}$	-0.645 (-1.581, 0.291) p=0.18	0.403 (-0.573, 1.380) p=0.42	-	-0.203 (-1.516, 1.11) p=0.76	-0.155 (-1.103, 0.793) p=0.75	0.119 (-0.840, 1.079) p=0.81	1.609 (0.472, 2.746) $p=5.6x10^{-3}$
RS1 cases	Beta value 95% CI (p value)	1.666 (0.812, 2.52) $p=1.4\ x10^{-4}$	-0.471 (-1.347, 0.406) p=0.29	-	0.474 (-0.408, 1.355) p=0.29	0.334 (-0.882, 1.550) p=0.59	-0.145 (-1.055, 0.765) p=0.76	-0.107 (-1.014, 0.799) p=0.82	0.894 (-0.043, 1.831) p=0.06
RS2 cases	Beta value 95% CI (p value)	2.174 (0.581, 3.767) $p=7.7\ x10^{-3}$	-0.380 (-2.071, 1.310) p=0.66	-	-1.306 (-2.865, 0.253) p=0.10	1.264 (-1.002, 3.529) p=0.27	0.691 (-0.946, 2.328) p=0.41	-1.817 (-3.548, -0.0861) p=0.04	-0.326 (-2.211, 1.559) p=0.73
RS3 cases	Beta value 95% CI (p value)	-	-0.615 (-1.913, 0.684) p=0.35	-	0.578 (-0.682, 1.837) p=0.40	1.624 (-0.112, 3.361) p=0.07	-1.601 (-2.833, -0.369) p=0.01	0.219 (-1.104, 1.543) p=0.75	0.457 (-0.868, 1.783) p=0.50
Meta-analysis cases	Beta value 95% CI (p value)	1.702 (1.117, 2.287) $p=1.2x10^{-8}$	-0.545 (-1.089, -0.002) p=0.05	0.254 (-0.290, 0.799) p=0.36		0.501 (-0.248, 1.250) p=0.19	-0.341 (-0.887, 0.205) p=0.22	-0.150 (-0.709, 0.408) p=0.60	0.880 (0.278, 1.482) $p=4.0\ x10^{-3}$
RS1 controls	Beta value 95% CI (p value)	0.3712 (-0.361, 1.103) p=0.32	0.0734 (-0.702, 0.849) p=0.85	-	-0.398 (-1.143, 0.347) p=0.30	0.190 (-0.756, 1.136) p=0.69	-0.337 (-1.103, 0.428) p=0.39	-0.767 (-1.529, -0.006) p=0.05	-0.141 (-0.928, 0.647) p=0.73
RS2 controls	Beta value 95% CI (p value)	1.228 (0.0420, 2.414) p=0.04	-1.013 (-2.200, 0.173) p=0.10	-	-1.279 (-2.442, -0.116) p=0.03	0.683 (-0.855, 2.222) p= 0.3844	0.668 (-0.527, 1.862) p= 0.27	1.021 (-0.229, 2.27) p= 0.11	-0.457 (-1.749, 0.835) p=0.49
RS3 controls	Beta value 95% CI (p value)	-	0.686 (-0.227, 1.598) p=0.14	-	0.244 (-0.656, 1.144) p=0.59	1.638 (0.482, 2.795) $p=5.6\ x10^{-3}$	0.751 (-0.111, 1.614) p=0.09	0.223 (-0.697, 1.143) P=0.63	0.518 (-0.390, 1.425) p=0.26
Meta-analysis controls	Beta value 95% CI (p value)	0.607 (-0.015, 1.230) p=0.06	0.063 (-0.466, 0.592) p=0.82	-	-0.360 (-0.875, 0.154) p=0.17	0.910 (0.309, 1.510) $p=3.0\ x10^{-3}$	0.240 (-0.276, 0.757) p= 0.36	-0.114 (-0.645, 0.417) p=0.67	0.038 (-0.502, 0.578) p=0.89

Table S6: Effects of adjusting T2D associations for BMI and waist circumference. In this table, ORs and 95% CIs are reported with respect to the risk allele (denoted in bold, with the ancestral allele underlined where known). Analyses report ORs and CIs before and after adjustment for log_{10}BMI or waist circumference and gender, by logistic regression, and fixed-effects meta-analysis. These analyses are only possible for the replication sets, since BMI and waist circumference values were not available in the WTCCC controls. Only the T2D associations at *FTO* are attenuated by adjustment for BMI and waist circumference.

						Replication meta analysis (unadjusted)		Replication meta analysis (adjusted for BMI)		Replication meta analysis (adjusted for WC)	
rs	**chr**	**Position**	**A1**	**A2**	**Region**	**OR (95%Cis)**	**P_{add}(unadjusted)**	**OR (95%Cis)**	**P_{add}(adjusted)**	**OR (95%Cis)**	**P_{add}(adjusted)**
rs8050136	16	52373776	**A**	C	*FTO*	1.22 (1.12-1.32)	$5.4x10^{-7}$	1.09 (0.99-1.18)	0.079	1.09 (0.99-1.19)	0.069
rs10946398	6	20769013	A	**C**	*CDKAL1*	1.14 (1.07-1.22)	$8.4x10^{-5}$	1.15 (1.07-1.23)	$2.8x10^{-4}$	1.15 (1.06-1.24)	$4.6x10^{-4}$
rs1111875	10	94452862	**A**	G	*HHEX*	1.08 (1.01-1.15)	0.020	1.09 (1.01-1.17)	0.026	1.08 (1.00-1.16)	0.057
rs10811661	9	22124094	C	**T**	*CDKN2B*	1.18 (1.08-1.28)	$1.7x10^{-4}$	1.22 (1.10-1.34)	$8.0x\ 10^{-5}$	1.26 (1.14-1.40)	$4.7x10^{-6}$
rs564398	9	22019547	C	**T**	*CDKN2B*	1.12 (1.05-1.19)	$8.6x10^{-4}$	1.12 (1.04-1.20)	$2.5x\ 10^{-3}$	1.12 (1.04-1.20)	$3.3x\ 10^{-3}$
rs4402960	3	186994389	G	**T**	*IGF2BP2*	1.09 (1.01-1.16)	0.018	1.10 (1.02-1.18)	0.014	1.12 (1.03-1.21)	$5.3x\ 10^{-3}$
rs13266634	8	118253964	**C**	T	*SLC30A8*	1.12 (1.04-1.19)	$1.2x10^{-3}$	1.14 (1.05-1.23)	$1.1\ x10^{-3}$	1.14 (1.06-1.24)	$9.5\ x\ 10^{-4}$

Table S7: Associations between T2D susceptibility variants and age of diagnosis in cases. Analyses were performed using linear regression on square root-transformed age of diagnosis values using gender as a covariate. Beta values, 95% CIs and asymptotic p values (t statistic) are reported. Fixed-effects meta-analyses are shown. In the case of *HHEX*, the UK meta-analysis combines data from rs5015480 and rs1111875 (r^2=1 in HapMap CEU). Rs8050136 were not typed in RS3. T2D- and adiposity-predisposing variants at *FTO* are associated with earlier age of diagnosis. At rs8050136, rare allele (A; T2D risk allele) homozygotes have an age of diagnosis that is on average 1.7 years earlier than the common allele (CC) homozygotes.

		rs8050136 **Chr16; 52373776 bp**	**rs10946398** **Chr 6; 20769013 bp**	**rs5015480** **Chr 10; 94455539 bp**	**rs1111875** **Chr 10; 94452862 bp**	**rs10811661** **Chr 9; 22124094 bp**	**rs564398** **Chr 9; 22019547 bp**	**rs4402960** **Chr 3; 186994389 bp**	**rs13266634** **Chr 8; 118253964 bp**
	Region	***FTO***	***CDKAL1***	***HHEX***	***HHEX***	***CDKN2B***	***CDKN2B***	***IGF2BP2***	***SLC30A8***
WTCCC cases	Beta value 95% CI (p value)	-0.032 (-0.075, 0.011) p=0.15	-0.029 (-0.072, 0.015) p=0.19	0.012 (-0.033, 0.056) p=0.61	-	0.012 (-0.049, 0.073) p=0.70	0.005 (-0.039, 0.049) p=0.82	-0.036 (-0.081, 0.008) p=0.11	-0.013 (-0.066, 0.040) p=0.64
RS1 cases	Beta value 95% CI (p value)	-0.048 (-0.087, -0.008) p=0.02	-0.003 (-0.043, 0.037) p=0.89	-	0.026 (-0.015, 0.067) p=0.21	0.004 (-0.052, 0.059) p=0.89	-0.023 (-0.065, 0.019) p=0.28	-0.021 (-0.063, 0.021) p=0.32	0.046 (0.003, 0.089) p=0.04
RS2 cases	Beta value 95% CI (p value)	-0.069 (-0.145, 0.007) p=0.08	-0.076 (-0.157, 0.005) p=0.07	-	-0.033 (-0.108, 0.042) p=0.40	-0.017 (-0.125, 0.091) p=0.76	-0.017 (-0.094, 0.061) p=0.68	-0.033 (-0.115, 0.049) p=0.43	0.061 (-0.028, 0.150) p=0.18
RS3 cases	Beta value 95% CI (p value)	-	-0.007 (-0.065, 0.050) p=0.80	-	-0.032 (-0.088, 0.024) p=0.26	-0.003 (-0.081, 0.074) p=0.93	0.075 (0.020, 0.130) p=7.60 x10^{-3}	0.023 (-0.036, 0.083) p=0.44	-0.050 (-0.109, 0.009) p=0.09
Meta-analysis all UK case samples	Beta value 95% CI (p value)	-0.044 (-0.071, -0.017) p=1.0x10^{-3}	-0.019 (-0.044, 0.006) p=0.13	0.006 (-0.020, 0.032) p=0.66		0.003 (-0.031, 0.037) p=0.87	0.007 (-0.018, 0.033) p=0.56	-0.019 (-0.045, 0.007) p=0.15	0.010 (-0.017, 0.038) p=0.46

Table S8: Genotype counts, association p values under different genetic models and test of departure from additivity for robustly replicating signals. The most significant model for each SNP is shown in bold. In the case of *HHEX*, the UK meta-analysis combines data from rs5015480 and rs1111875 (r^2=1 in HapMap CEU).

rs	**rs7193144**	**rs8050136**	**rs10946398**	**rs9465871**	**rs5015480**	**rs1111875**	**rs10811661**	**rs564398**	**rs4402960**	**rs13266634**
chr	16	16	6	6	10	10	9	9	3	8
position	52368187	52373776	20769013	20825234	94455539	94452862	22124094	22019547	186994389	118253964
region	*FTO*	*FTO*	*CDKAL1*	*CDKAL1*	*HHEX*	*HHEX*	*CDKN2B*	*CDKN2B*	*IGF2BP2*	*SLC30A8*
allele1	C	A	A	C	C	C	C	C	G	C
allele2	T	C	C	T	T	T	T	T	T	T
WTCCC cases 11	375	378	795	101	704	-	36	298	807	794
WTCCC cases 12	992	987	865	638	936	-	470	962	881	637
WTCCC cases 22	551	550	261	1184	248	-	1418	661	234	119
WTCCC controls 11	458	464	1364	77	915	-	74	583	1354	1393
WTCCC controls 12	1405	1407	1271	893	1448	-	827	1428	1278	1192
WTCCC controls 22	1064	1063	303	1964	488	-	2035	921	300	281
WTCCC p_general	4.8×10^{-8}	7.0×10^{-8}	1.1×10^{-4}	3.3×10^{-7}	2.9×10^{-5}	-	3.3×10^{-3}	5.2×10^{-4}	7.1×10^{-3}	0.039
WTCCC p_additive	1.4×10^{-8}	2.0×10^{-8}	2.5×10^{-5}	1.0×10^{-6}	5.4×10^{-6}	-	7.6×10^{-4}	3.2×10^{-4}	1.7×10^{-3}	0.020
WTCCC p_dominant_allele2	4.3×10^{-4}	4.2×10^{-4}	5.4×10^{-4}	1.9×10^{-6}	2.2×10^{-4}	-	0.14	1.1×10^{-4}	4.0×10^{-3}	0.096
WTCCC p_recessive_allele2	3.7×10^{-8}	5.8×10^{-8}	5.0×10^{-4}	1.3×10^{-4}	2.1×10^{-4}	-	9.7×10^{-4}	0.029	0.034	0.019
deviation from additivity in WTCCC p value	0.18	0.20	0.54	0.019	0.60	-	0.84	0.078	0.87	0.27
RS1 cases 11	347	351	829	83	-	724	38	264	831	992
RS1 cases 12	956	962	844	577	-	917	464	925	884	776
RS1 cases 22	606	625	247	1279	-	269	1434	642	207	156
RS1 controls 11	288	283	853	50	-	715	59	318	886	928
RS1 controls 12	902	915	868	580	-	917	534	915	839	829
RS1 controls 22	728	744	194	1305	-	273	1330	566	199	168
RS1 p_general	1.1×10^{-4}	8.2×10^{-5}	0.030	0.015	-	0.96	1.3×10^{-3}	8.4×10^{-3}	0.21	0.11
RS1 p_additive	2.9×10^{-5}	1.6×10^{-5}	0.058	0.072	-	0.78	2.8×10^{-4}	2.4×10^{-3}	0.13	0.055
RS1 p_dominant_allele2	8.6×10^{-3}	2.9×10^{-3}	0.39	3.7×10^{-3}	-	0.81	0.028	7.5×10^{-3}	0.67	0.038
RS1 p_recessive_allele2	5.5×10^{-5}	7.8×10^{-5}	7.9×10^{-3}	0.33	-	0.83	7.2×10^{-4}	0.021	0.079	0.49
deviation from additivity in RS1 p value	0.40	0.66	0.063	0.021	-	1.00	0.73	0.55	0.38	0.43

RS2 cases 11	111	118	254	26	-	258	15	106	274	293
RS2 cases 12	286	292	294	204	-	254	145	281	285	245
RS2 cases 22	194	206	73	392	-	102	462	211	62	43
RS2 controls 11	195	222	785	57	-	560	43	269	740	620
RS2 controls 12	717	816	678	500	-	780	412	658	736	573
RS2 controls 22	518	589	164	1081	-	277	1183	483	158	135
RS2 p_general	0.011	5.0×10^{-3}	7.5×10^{-3}	0.38	-	3.7×10^{-3}	0.62	0.75	0.88	0.10
RS2 p_additive	0.010	0.010	4.0×10^{-3}	0.16	-	0.018	0.35	0.49	0.63	0.044
RS2 p_dominant_allele2	3.3×10^{-3}	1.2×10^{-3}	1.8×10^{-3}	0.43	-	1.2×10^{-3}	0.77	0.48	0.82	0.13
RS2 p_recessive_allele2	0.15	0.22	0.25	0.19	-	0.77	0.33	0.66	0.62	0.056
deviation from additivity in RS2 p value	0.15	0.056	0.21	0.98	-	0.019	0.79	0.77	0.88	0.42
RS3 cases 11	-	-	442	-	-	421	19	173	461	508
RS3 cases 12	-	-	508	-	-	507	273	524	496	466
RS3 cases 22	-	-	122	-	-	145	784	382	106	97
RS3 controls 11	-	-	697	-	-	528	37	274	727	673
RS3 controls 12	-	-	663	-	-	770	408	742	643	675
RS3 controls 22	-	-	153	-	-	213	1068	495	144	163
RS3 p_general	-	-	0.049	-	-	0.080	0.30	0.23	0.062	0.20
RS3 p_additive	-	-	0.020	-	-	0.070	0.14	0.087	0.050	0.078
RS3 p_dominant_allele2	-	-	0.015	-	-	0.026	0.24	0.16	0.70	0.15
RS3 p_recessive_allele2	-	-	0.30	-	-	0.67	0.21	0.16	0.020	0.15
deviation from additivity in RS3 p value	-	-	0.41	-	-	0.19	0.60	0.89	0.19	0.75
replication m/a p_general	6.0×10^{-6}	2.7×10^{-6}	4.2×10^{-4}	0.013	-	0.013	7.9×10^{-4}	3.3×10^{-3}	0.024	5.4×10^{-3}
replication m/a p_additive	9.5×10^{-7}	5.4×10^{-7}	8.3×10^{-5}	0.023	-	0.020	1.7×10^{-4}	8.6×10^{-4}	0.018	1.2×10^{-3}
replication m/a p_recessive_allele2	3.1×10^{-5}	7.7×10^{-5}	3.2×10^{-3}	0.12	-	0.60	4.9×10^{-4}	0.010	6.3×10^{-3}	0.030
replication m/a p_dominant_allele2	1.5×10^{-4}	1.9×10^{-5}	6.5×10^{-4}	4.7×10^{-3}	-	3.6×10^{-3}	0.022	3.5×10^{-3}	0.54	3.3×10^{-3}
deviation from additivity in replication sets p value	0.76	0.36	0.86	0.065	-	0.16	0.60	0.27	0.22	1.00
all UK samples m/a p_general	7.0×10^{-13}	6.7×10^{-13}	8.0×10^{-8}	**4.2×10^{-8}**	1.5×10^{-5}		3.2×10^{-6}	2.6×10^{-6}	3.9×10^{-4}	3.1×10^{-4}
all UK samples m/a p_additive	**9.6×10^{-14}**	**7.3×10^{-14}**	**1.3×10^{-8}**	6.4×10^{-7}	**4.6×10^{-6}**		**4.9×10^{-7}**	**1.3×10^{-6}**	**1.6×10^{-4}**	**7.0×10^{-5}**
all UK samples m/a p_recessive_allele2	1.2×10^{-11}	5.4×10^{-11}	8.0×10^{-6}	1.9×10^{-4}	7.7×10^{-3}		1.8×10^{-6}	7.8×10^{-4}	0.12	1.7×10^{-3}
all UK samples m/a p_dominant_allele2	2.3×10^{-7}	3.2×10^{-8}	1.6×10^{-6}	1.1×10^{-7}	5.8×10^{-6}		6.4×10^{-3}	2.8×10^{-6}	0.064	7.8×10^{-4}
deviation from additivity in all UK samples p value	0.52	0.89	0.61	3.0×10^{-3}	0.72		0.80	0.050	0.28	0.49

Table S9: Selection criteria used for "second wave" SNPs.

rs	chr	position (bp)	Selection criteria			
			Corroborating evidence from other T2D scans	WTCCC p value	Biological candidacy	Multiple independent association signals at the same locus
rs10889039	1	57540566	√	x	x	x
rs480075	1	183187993	x	x	√	√
rs2383529	1	183400749	x	x	√	√
rs11679606	2	16717046	√	√	x	x
rs1022377	2	21968138	√	x	x	x
rs4140806	2	37742813	√	x	x	x
rs3814045	2	46650808	x	x	√	x
rs314890	2	125700943	√	x	x	x
rs2217616	2	182825154	√	x	x	x
rs11688935	2	189001401	x	√	x	x
rs17248501	2	205864551	x	√	x	x
rs440646	3	11266171	√	√	x	x
rs4402960	3	186994389	√	x	x	x
rs7651090	3	186996094	√	x	x	x
rs3118617	4	2307648	√	x	x	x
rs10516948	4	95525785	√	x	x	x
rs7698608	4	104295578	√	x	x	x
rs2139980	4	157974803	√	x	x	x
rs152189	5	61716205	√	x	x	x
rs1029328	6	28555894	√	x	x	x
rs2242655	6	31735428	√	x	x	x
rs9369425	6	43918952	√	x	x	x
rs1665901	6	107540093	√	√	x	x
rs7755651	6	148766018	√	x	x	x
rs864745	7	27953796	√	x	x	x
rs12702146	7	45294288	√	x	x	x
rs4493865	7	54672611	√	x	x	x
rs7007919	8	20691655	√	x	x	x

rs16879809	8	32692415	√	x	x	x
rs16875331	8	108008872	√	x	x	x
rs13266634	8	118253964	√	x	x	x
rs939024	8	134655895	√	x	x	x
rs10511667	9	18979696	√	x	x	x
rs564398	9	22019547	x	x	√	√
rs2383208	9	22122076	√	x	x	x
rs10811661	9	22124094	√	x	x	x
rs10757283	9	22124172	x	√	x	x
rs10780690	9	84707726	x	x	√	√
rs12346884	9	84739293	x	x	√	x
rs11140802	9	84741930	x	x	√	√
rs4743611	9	96011963	√	x	x	x
rs10120268	9	108627659	x	√	x	x
rs10817674	9	114602134	√	x	x	x
rs10818930	9	124070739	√	x	x	x
rs7023712	9	126328516	√	x	x	x
rs505699	10	99764066	x	√	x	x
rs11597086	10	101943695	√	x	x	x
rs3802678	10	104105385	x	√	x	x
rs1153188	12	53385263	√	x	x	x
rs4426172	12	126483613	x	√	x	x
rs7323267	13	40102015	x	x	√	x
rs2268974	14	68492917	√	x	x	x
rs4343209	14	98286712	√	√	x	x
rs2099106	16	9288209	x	√	x	x
rs6038317	20	5813636	√	x	x	x
rs2154490	21	29837833	x	√	x	x
rs440544	21	35817449	√	x	x	x

Table S10: Pairwise interaction analyses of replicating SNPs, and known T2D susceptibility variants in *TCF7L2*, *KCNJ11* and *PPARG*. Odds ratios (for interactions under a log additive model) are reported to the risk allele at each SNP, as defined in Table 1. In the replication sets, the analyses were adjusted for the three strata. The results are ordered by interaction p value. Rs1801282, rs5215 and rs7901695 were not typed in the replication sets. Rs1111875 was typed as an r^2=1 proxy (HapMap CEU) for rs5015480 in the replication sets.

Table S10: Pairwise interaction analyses of replicating SNPs, and known T2D susceptibility variants in *TCF7L2*, *KCNJ11* and *PPARG*.

SNP1	Region1	SNP2	Region2	WTCCC OR_{int} (95% CI)	WTCCC p	Replication OR_{int} (95% CI)	Replication p	Combined OR_{int} (95% CI)	Combined p
rs10811661	*CDKN2B*	rs5015480	*HHEX*	1.11 (0.93-1.32)	0.24	1.14 (1.00-1.28)	0.06	1.12 (1.01-1.25)	0.025
rs564398	*CDKN2B*	rs13266634	*SLC30A8*	-	-	0.89 (0.81-0.99)	0.026	-	-
rs10811661	*CDKN2B*	rs4402960	*IGF2BP2*	1.05 (0.89-1.25)	0.53	0.81 (0.71-0.93)	$2x10^{-3}$	0.89 (0.81-0.99)	0.037
rs7901695	*TCF7L2*	rs10946398	*CDKAL1*	0.87 (0.77-0.99)	0.040	-	-	-	-
rs5215	*KCNJ11*	rs4402960	*IGF2BP2*	0.91 (0.80-1.03)	0.13	-	-	-	-
rs7901695	*TCF7L2*	rs5215	*KCNJ11*	0.89 (0.78-1.01)	0.096	-	-	-	-
rs564398	*CDKN2B*	rs5215	*KCNJ11*	1.05 (0.92-1.18)	0.49	-	-	-	-
rs5215	*KCNJ11*	rs10946398	*CDKAL1*	1.10 (0.97-1.25)	0.13	-	-	-	-
rs1801282	*PPARG*	rs4402960	*IGF2BP2*	1.15 (0.94-1.39)	0.18	-	-	-	-
rs4402960	*IGF2BP2*	rs8050136	*FTO*	1.08 (0.95-1.23)	0.22	0.85 (0.76-0.96)	$7x10^{-3}$	0.95 (0.88-1.04)	0.27
rs13266634	*SLC30A8*	rs5015480	*HHEX*	-	-	0.95 (0.86-1.05)	0.28	-	-
rs7901695	*TCF7L2*	rs1801282	*PPARG*	0.89 (0.73-1.10)	0.29	-	-	-	-
rs8050136	*FTO*	rs13266634	*SLC30A8*	-	-	0.94 (0.84-1.06)	0.32	-	-
rs5215	*KCNJ11*	rs1801282	*PPARG*	0.93 (0.77-1.13)	0.45	-	-	-	-
rs4402960	*IGF2BP2*	rs5015480	*HHEX*	0.94 (0.83-1.08)	0.36	0.98 (0.89-1.08)	0.68	0.97 (0.89-1.04)	0.39
rs1801282	*PPARG*	rs10946398	*CDKAL1*	0.88 (0.73-1.08)	0.23	-	-	-	-
rs10946398	*CDKAL1*	rs8050136	*FTO*	1.04 (0.93-1.18)	0.47	1.02 (0.91-1.15)	0.71	1.03 (1.12-0.95)	0.41
rs10946398	*CDKAL1*	rs4402960	*IGF2BP2*	1.00 (0.89-1.15)	0.89	0.96 (0.87-1.05)	0.39	0.97 (0.90-1.05)	0.45
rs564398	*CDKN2B*	rs10811661	*CDKN2B*	0.91 (0.77-1.08)	0.27	1.00 (0.88-1.13)	0.95	0.96 (0.87-1.07)	0.46
rs564398	*CDKN2B*	rs8050136	*FTO*	1.01 (0.89-1.14)	0.87	0.94 (0.84-1.05)	0.29	0.97 (0.89-1.05)	0.52
rs564398	*CDKN2B*	rs1801282	*PPARG*	0.98 (0.81-1.19)	0.82	-	-	-	-
rs4402960	*IGF2BP2*	rs13266634	*SLC30A8*	-	-	0.97 (0.87-1.07)	0.53	-	-

rs5215	*KCNJ11*	rs5015480	*HHEX*	0.91 (0.80-1.03	0.21	-	-	-	-
rs1801282	*PPARG*	rs5015480	*HHEX*	1.00 (0.82-1.22)	0.99	-	-	-	-
rs1801282	*PPARG*	rs10811661	*CDKN2B*	1.01 (0.77-1.33)	0.92	-	-	-	-
rs8050136	*FTO*	rs5015480	*HHEX*	0.98 (0.86-1.11)	0.71	0.95 (0.85-1.06)	0.38	0.98 (0.9-1.06)	0.64
rs7901695	*TCF7L2*	rs8050136	*FTO*	1.02 (0.9-1.16)	0.70	-	-	-	-
rs1801282	*PPARG*	rs13266634	*SLC30A8*	-	-	-	-	-	-
rs5215	*KCNJ11*	rs10811661	*CDKN2B*	0.96 (0.81-1.14)	0.64	-	-	-	-
rs564398	*CDKN2B*	rs10946398	*CDKAL1*	1.03 (0.91-1.16)	0.68	0.97 (0.88-1.06)	0.49	0.99 (0.92-1.07)	0.77
rs5215	*KCNJ11*	rs13266634	*SLC30A8*	-	-	-	-	-	-
rs10946398	*CDKAL1*	rs10811661	*CDKN2B*	1.03 (0.87-1.22)	0.76	0.97 (0.85-1.10)	0.60	0.99 (0.89-1.10)	0.81
rs1801282	*PPARG*	rs8050136	*FTO*	1.00 (0.83-1.20)	0.97	-	-	-	-
rs7901695	*TCF7L2*	rs4402960	*IGF2BP2*	0.99 (0.87-1.12)	0.83	-	-	-	-
rs10811661	*CDKN2B*	rs13266634	*SLC30A8*	-	-	1.01 (0.88-1.16)	0.84	-	-
rs7901695	*TCF7L2*	rs5015480	*HHEX*	0.99 (0.86-1.12)	0.85	-	-	-	-
rs10946398	*CDKAL1*	rs5015480	*HHEX*	0.98 (0.86-1.11)	0.71	1.02 (0.93-1.12)	0.64	1.01 (0.93-1.09)	0.87
rs7901695	*TCF7L2*	rs10811661	*CDKN2B*	1.01 (0.85-1.20)	0.88	-	-	-	-
rs564398	*CDKN2B*	rs4402960	*IGF2BP2*	1.05 (0.93-1.19)	0.43	0.98 (0.88-1.08)	0.67	1.01 (0.93-1.09)	0.89
rs10946398	*CDKAL1*	rs13266634	*SLC30A8*	-	-	1.01 (0.91-1.11)	0.91	-	-
rs564398	*CDKN2B*	rs7901695	*TCF7L2*	1.01 (0.88-1.14)	0.91	-	-	-	-
rs10811661	*CDKN2B*	rs8050136	*FTO*	0.97 (0.83-1.14)	0.72	1.02 (0.87-1.18)	0.84	0.99 (0.89-1.11)	0.92
rs564398	*CDKN2B*	rs5015480	*HHEX*	0.92 (0.81-1.04)	0.17	1.05 (0.96-1.16)	0.24	1.00 (0.93-1.08)	0.99
rs5215	*KCNJ11*	rs8050136	*FTO*	1.03 (0.91-1.17)	0.60	-	-	-	-

Supplementary Online Material: References

S1. P. Donnelly and the WTCCC, personal communication. Data from the Wellcome Trust Case Control Consortium scan.

S2. S. Wiltshire *et al*., *Am. J. of Hum. Gen.* **69**, 553-569 (2001).

S3. T. M. Frayling *et al*., *Diabetes* **48**, 2475-2479 (1999).

S4. C. Power, J. Elliott, *Int. J. Epidemiol.*, **35**, 34 (2006).

S5. K. R. Owen, A. Stride, S. Ellard, A. T. Hattersley, *Diabetes Care*, **26**, 2088-2093 (2003).

S6. C. J. Groves *et al*., *Diabetes,* **55**, 2640-2644 (2006).

S7. B. Knight, B. M. Shields, A. T. Hattersley, *Paediatr Perinat Epidemiol*., **20**, 172-179 (2006).

S8. A. L. Price, N. J. Patterson, R. M. Plenge, M. E. Weinblatt, N. A. Shadick, D. Reich, *Nat. Genet*., **38**, 904-909 (2006).

S9. Diabetes Genetics Initiative. Genome-wide association analysis identifies three novel loci for type 2 diabetes and one for triglyceride levels in the human population. (Submitted to *Science*).

S10. L. J.Scott *et al*. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. (Submitted to *Science*).

S11. R. Sladek *et al*., *Nature,* **445**, 881-885 (2007).

S12. PLINK v0.99p http://pngu.mgh.harvard.edu/~purcell/plink/.

S13. STATA 9.0, Stata Corporation: College Station, Texas, USA.

S14. StatXact, Cytel Corporation, Cambridge, MA

S15. A. P. Morris, *Gen. Epidemiol.,* **29**, 91-107 (2005).

S16. S. B. Gabriel *et al*., *Science*, **296**, 2225-2229 (2002).