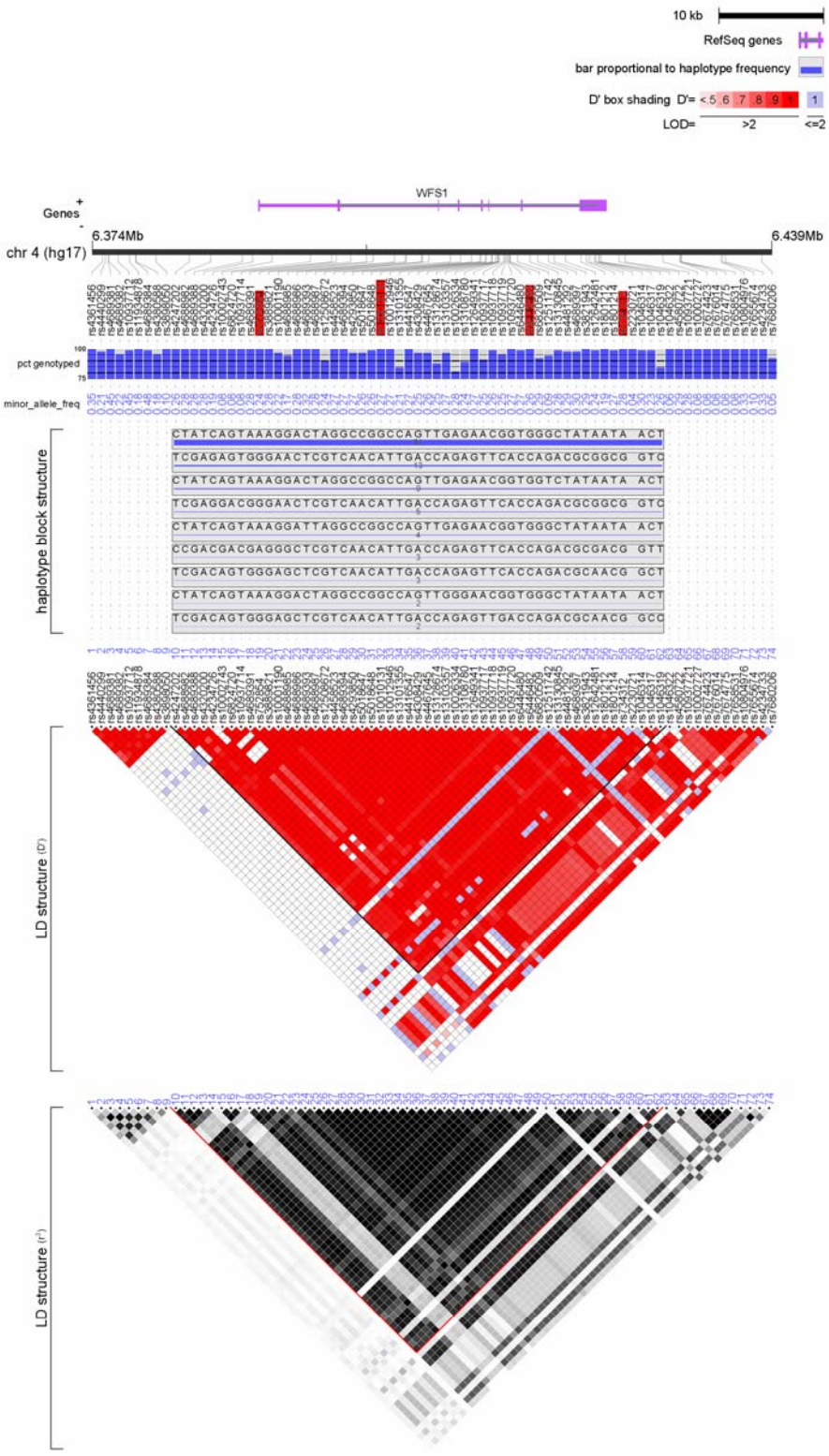
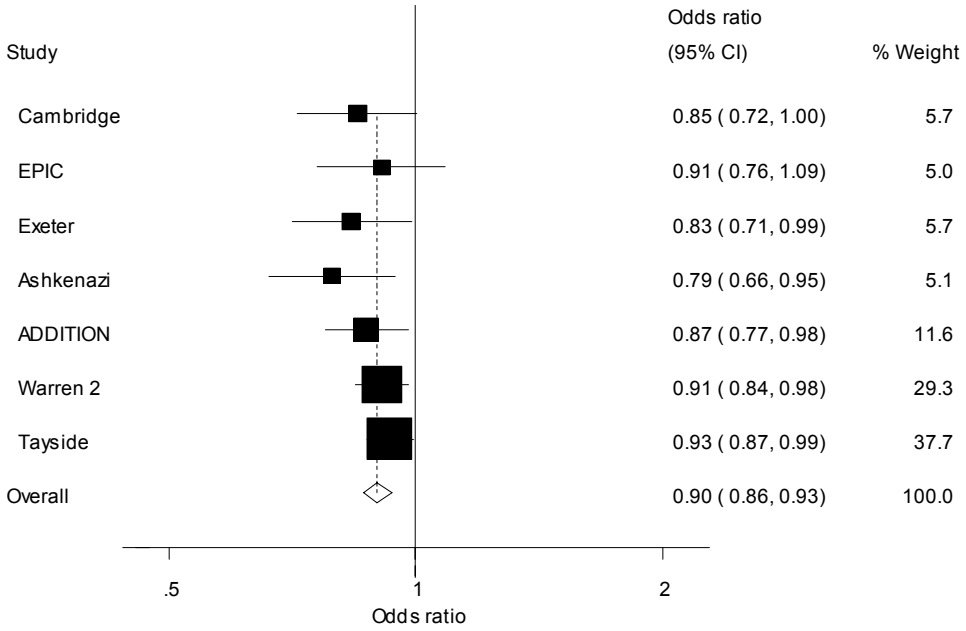
Supplementary figure 1



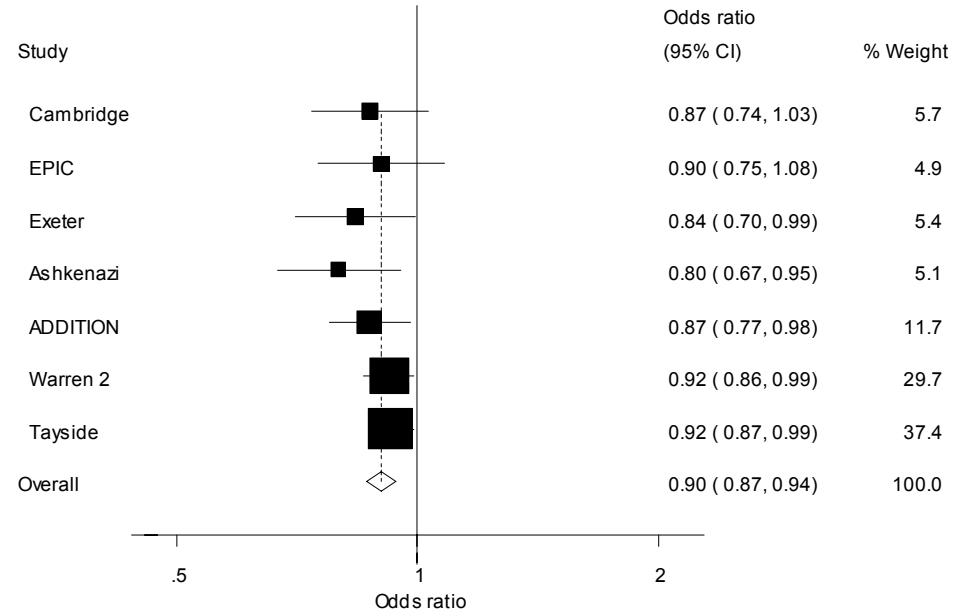
Supplementary Figure 1. Feature map of the *WFS1* gene and surrounding sequence. Positions of SNPs genotyped in HapMap (release 21/phaseII Jul 06, NCBI B35 assembly, dbSNP build 125, (CEPH Utah trios)) with a MAF≥ 0.01 are shown relative to the locus (purple) and chromosome 4 (black bar) (see text for details). The four SNPs typed in all studies are highlighted in red. Underlying each SNP is QC information and MAF in the CEPH samples. The middle part of the figure shows the nine common haplotypes for the *WFS1* gene with frequency ≥ 0.01 in the CEPH samples within the 39 kb LD block that encompasses 53 SNPs. The thickness of the blue lines is proportional to the haplotype frequencies. The bottom of the figure depicts two LD plots for the *WFS1* locus region with pairwise LD values presented for SNPs. The 39 kb LD block encompassing the *WFS1* locus is highlighted in black and red in the upper and lower plots respectively. The upper plot presents LD as D' - see figure key for details. The figure was generated using LocusView (T. Petryshen, A. Kirby, M. Ainscow, unpublished software, available from the Broad Institute, Cambridge, MA (http://www.broad.mit.edu/mpg/locusview/). In the lower plot, LD among SNPs is given as r^2. r^2 values of 1.0 are represented by black diamonds, intermediate r^2 values are shown in grey and r^2 values of 0 as white. This plot was generated using Haploview, available from the HapMap website (http://www.broad.mit.edu/mpg/haploview/index.php).

Supplementary Figure 2 (a) Association between SNP rs10010131 and risk of type 2 diabetes in individual studies. (b) Association between SNP rs6446482 and risk of type 2 diabetes in individual studies.

(a)



(b)



Supplementary methods

Gene and SNP selection

Candidate genes were chosen based on their known or putative role in pancreatic β-cells. In total 84 candidate genes were selected belonging to one or more of the following categories: genes involved in pancreas development/transcription; genes involved in β-cell death/ apoptosis; genes involved in insulin secretion and signalling in the β-cell; enzymes and metabolism; inflammation (see table below). All SNPs present in HapMap in November 2004 that had been genotyped in CEPH and mapped to these 84 transcripts were selected for genotyping except those in complete linkage disequilibrium ($r^2 = 1$) with a SNP already selected. This procedure generated 1,536 SNPs that captured variation in 84 genes for genotyping.

Table. Candidate genes and number of SNPs genotyped according to functional group

Group	Hugo Symbol	Common Name	Number of SNPs
Pancreas development/ transcription	*ACVR1*	activin A receptor, type I	12
	ACVR1b	activin A receptor, type IB	4
	ACVR2	activin A receptor, type II ; ACTRII	9
	BTC	betacellulin	5
	CDH1	cadherin 1, type 1, E-cadherin (epithelial)	10
	CDH2	cadherin 2, type 1, N-cadherin (neuronal)	80
	CDK2	cyclin-dependent kinase 2	4
	EGF	epidermal growth factor (beta-urogastrone)	7
	EGFR	epidermal growth factor receptor	37
	FGF10	fibroblast growth factor	5
	FGF7	fibroblast growth factor 7 (keratinocyte growth factor)	9
	FGFR1	fibroblast growth factor receptor 1 (fms-related tyrosine kinase 2, Pfeiffer syndrome)	4
	FGFR2	fibroblast growth factor receptor 2 (bacteria-expressed kinase, keratinocyte growth factor receptor, craniofacial dysostosis 1, Crouzon syndrome, Pfeiffer syndrome, Jackson-Weiss syndrome)	113
	FOXA3	HNF3G	4
	FOXO1A	FKHR, FOXO1	10
	FOXO3A	FKHRL1	14
	HNF4G	HNF4G	9
	ISL1	ISL1 transcription factor, LIM/homeodomain, (islet-1)	1
	JAG1	jagged 1 (Alagille syndrome)	10
	JAG2	jagged 2	5
	LHX4	LIM homeobox 4	10
	LHX6	LHX6.1	8
	LMX1A	LIM homeobox transcription factor 1, alpha	29
	NEUROD1	neurod/Beta2	1

	Notch1	Notch homolog 1, translocation-associated (Drosophila)	9
	NOTCH2	Notch homolog 2 (Drosophila)	8
	NOTCH3	Notch homolog 3 (Drosophila)	8
	Notch4	Notch homolog 4 (Drosophila)	5
	NR5A2	Lrh1	30
	ONECUT1	HNF6	6
	PAX6	paired box gene 6 (aniridia, keratitis)	4
	PBX1	pre-B-cell leukemia transcription factor 1	56
	PSEN2	presenilin 2 (Alzheimer disease 4)	5
	RARB	retinoic acid receptor, beta	95
	RORB	nuclear receptor RZR-beta	19
	TCF1	HNF1A	5
	TCF2	HNF1B	12
	TCF3	E47	3
	TGFBR1	transforming growth factor, beta receptor I (activin A receptor type II-like kinase, 53kDa)	5
	VIPR1	vasoactive intestinal peptide receptor 1	9
	VIPR2	vasoactive intestinal peptide receptor 2	9
β-cell death/ apoptosis	*BAX*	BCL2-associated X protein	3
	BCL2	B-cell leukemia/lymphoma 2	53
	BID	BH3 interacting domain death agonist	16
	BIK	BCL2-interacting killer (apoptosis-inducing)	4
	CASP9	caspase 9	4
	EIF2AK3	PERK	7
	GSK3B	glycogen synthase kinase 3 beta	5
	WFS1	Wolfram syndrome 1 (wolframin)	6
Insulin secretion	*ATP2B1*	ATPase, Ca++ transporting	6
	CACNA1C	calcium channel, voltage-dependent, L type, alpha 1C subunit; CaV1.2	133
	CACNA1D	calcium channel, voltage-dependent, L type, alpha 1D subunit	69
	CHGA	chromogranin A (secretogranin 2)	5
	CPE	carboxypeptidase E	19
	EIF2S1	eukaryotic translation initiation factor 2, subunit 1 alpha, 35kDa	4
	GLP1R	glucagon-like peptide 1 receptor	10
	KCNJ6	potassium inwardly-rectifying channel, subfamily J, member 6	103
	MAPK8IP1	islet-brain 1	3
	PCSK1	prohormone convertase 1	8
	PCSK2	prohormone convertase 2	45
	SGNE1	secretory granule, neuroendocrine prot 1	9
	SNAP25	SNAP25 synaptosomal-associated protein	16
	VTI1B	VTI1B vesicle transport through interaction with t-SNAREs homolog 1B (yeast) or vSNARE	4
Insulin signalling	*FGF2*	fibroblast growth factor 2 (basic); bFGF	9
	INSR	insulin receptor	27
	IRS1	insulin receptor substrate 1	5
	IRS2	insulin receptor substrate 2	7
	PIK3R1	phosphatidylinositol 3-kinase, regulatory subunit, polypeptide 1 (p85 alpha)	15
	PRKCE	PKC epsilon	109
	PRKCZ	PKC zeta	5
	PSEN1	presenilin 1 (Alzheimer disease 3)	8
	RPS6KB1	ribosomal protein S6 kinase, 70kDa, polypeptide 1	7
Enzymes and metabolism	*CAMK2A*	CAMK2A	16
	CAMK2B	CAMK2B	8
	CAMK2D	CAMK2D	28
	CAMK2G	calcium/calmodulin-dependent protein kinase (CaM kinase) II gamma	7
	CAPN3	calpain 3	11
	GCK	glucokinase	6
	GCKR	glucokinase regulatory protein	5
Inflammation	*IKBKB*	IKKb	4
	IL6	interleukin 6	2
	NFATC1	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 1	25
	NFKB1	NFKbeta; NFkB	10

Genotyping methods and QC strategy

Genotyping

SNPs were assayed at the Wellcome Trust Sanger Institute with Illumina's Golden Gate assay as described elsewhere[1,2]. DNA samples from all UK populations were whole-genome amplified by OmniPlex™ at Rubicon Genomics, Inc (Ann Arbor, MI, USA). Genomic DNA was used for the Ashkenazi samples. All DNA samples were diluted to 100ng/ul and used at this concentration for genotyping. Samples with multiple SNP failures were repeated once and then excluded from clustering if their 50% locus Gene Call (GC) score was below 70%; these were thought to be samples of poor quality DNA. Clustering was performed on a per panel basis analysing no more than 500 samples at a time and used duplicate samples (two per 96-well plate) to improve clustering. We applied a locus cut off of 0.3 and cut off value in the genotype confidence score of 0.25. On a per panel basis we applied a call rate threshold ≥80% and removed markers that displayed more than 1 discrepancy per plate. Markers departing from Hardy Weinberg equilibrium ($\chi^2 \geq 10$) were flagged at this point.

Secondary QC strategy

Of the 1,536 SNPs with attempted genotypes in all populations only those SNPs that obeyed the following QC criteria were included in the analysis: genotype call-rates greater than 90%; no significant difference in genotype call-rate between cases and controls; a P value ≥ 0.001 for tests of deviation from Hardy-Weinberg Equilibrium in controls. For those SNPs that survived QC where discordancy between replicate samples were found these samples were blanked out prior to analysis. Of the 1,536 SNPs selected from 84 candidate genes, 1,367 (90%) were successfully genotyped in ≥ 90% of the UK study samples, were polymorphic (minor allele frequency (MAF) of ≥ 0.1%), and were in Hardy-Weinberg equilibrium (P > 0.001 in controls).

The ADDITION and Tayside studies were genotyped using a custom TaqMan® SNP assay (Applied Biosystems, UK) at Strangeways Research Laboratory, University of Cambridge, and at the Biomedical Research Centre, University of Dundee, respectively. Allele calling was done on the ABI PRISM® 7900HT Sequence Detection System (Applied Biosystems, UK). The Warren 2 study was genotyped by Kbiosciences (Herts, U.K.) using a KASPar assay system (details of the methods used can be found at http://www.kbioscience.co.uk). For the rs734312 variant, the EPIC, Cambridge and Ashkenazim case control studies were genotyped using a custom TaqMan® SNP assay (Applied Biosystems, UK) at Strangeways Research Laboratory, University of Cambridge.

Populations

Cambridgeshire case–control

This is a population based case-control study in which a total of 552 patients aged 45–76 years with T2D were randomly selected from general practitioner diabetes registers in Cambridgeshire, UK[3]. Presence of T2D was based on clinical criteria; onset of diabetes after the age of 30 years without treatment with insulin in the first year after diagnosis. The controls were recruited at random from the same population sampling frames, and were individually matched to cases for age, sex and GP practice. Diabetes was excluded in controls by medical record search and by a glycated haemoglobin measurement of less than 6%. The study received ethical approval from the Cambridge Local Research Ethics Committee, and participants provided informed consent.

EPIC- Norfolk case-control study

The EPIC case-control study is nested within the EPIC - Norfolk Study, a population based cohort study of men and women aged 40-78 years of European descent. Both the case-control[4] and full cohort[5] study have been previously described in detail. Briefly, the case-control study consists of 417 incident type 2 diabetes cases and two sets of 417 controls, each matched in terms of age, sex, general practice, recruitment date, with one

set additionally matched for BMI. A case was defined by a physician's diagnosis of type 2 diabetes, with no insulin prescribed within the first year following diagnosis, and/or HbA_{1c} > 7% at the health check. Controls were randomly selected from the EPIC-Norfolk cohort from among those without diabetes, cancer, stroke, or myocardial infarction at baseline and who had not developed diabetes by the time of selection. Potential controls with measured HbA_{1c} levels > 6% were excluded. The EPIC-Norfolk study was approved by the Norfolk Local Research Ethics Committtee.

Exeter case-control study

The diabetic subjects from Exeter came from two sources (i) a consecutive-case series of patients with T2D diagnosed before 45 years from North and East Devon[6]. The patients were unrelated and recruited through questionnaires distributed though general practitioners (97% agreed to send out questionnaires, >70% return rate and >90% recruitment of those identified through the questionnaires). Validation of the diagnosis of diabetes was based on either current prescribed treatment with sulphonylureas, biguanides and/or insulin, or, in the case of individuals treated with diet alone, historical or contemporary laboratory evidence of hyperglycaemia (as defined by present WHO guidelines). All patients were off insulin for at least 1 year after diagnosis, and patients were excluded if they had pancreatic autoantibodies (GAD), first degree history of type 1 diabetes or clinical features (or DNA test results) suggestive of monogenic diabetes[6]. (ii) Probands from a collection of type 2 diabetes families that had either both parents available, or one parent and at least two siblings[7]. Only subjects collected in Exeter were used in this study. The sex matched controls are taken from the parents in the Exeter Family Study, a cohort study of newly born babies and both their parents[8]. This study recruits from central Exeter so the controls come from a similar geographical region as the cases. Diabetes and hyperglycaemia were excluded by measuring fasting glucose and HbA1c. In total 601 cases and 610 controls were included in this study. Informed consent was obtained from all participants.

Ashkenazi case-control study

Of the cases, 303 are from the multiplex-affected sibships that were ascertained for the genome scan described by Permutt et al, 2001[9]. The cases were of Ashkenazi Jewish origin, defined as having all four grandparents born in Northern or Eastern Europe. Subjects with known or suspected Sephardic Jewish or non-Jewish ancestry were excluded. T2D was initially defined according to World Health Organization criteria (fasting glucose 140 mg/dl on two or more occasions, or random glucose 200 mg/dl on two or more occasions). Their average age at ascertainment was 60 years. Average age at diagnosis was 47 years and average duration of diabetes was 13 years (range 0-47). In this population, the incidence of type 1 diabetes is relatively low therefore anti-GAD or anti-islet cell antibody titers were not routinely measured. The additional 627 cases were ascertained as part of a study with the dual aim of looking for diabetes related genes and for genes related to the risk of developing diabetic complications. This group has an average age at ascertainment of 65.8 years, age of diagnosis of 46.8 years and duration of diabetes of 19.1 years. The Ashkenazi control samples consist of 149 elderly subjects (average age 76 years) with no personal or first-degree family history of T2D. The remaining 312 samples were obtained from The National Laboratory for the Genetics of Israeli Populations at Tel Aviv University, Israel. The institutional review boards of Washington University (St. Louis, MO) and Hadassah University Hospital (Jerusalem, Israel) approved the study.

Warren 2 case-control study

Informed consent was obtained from all participants. The subjects have been described in detail previously [10]. Briefly, all type 2 diabetes subjects were unrelated and of white UK origin who had diabetes defined either by WHO criteria [11] or by being treated with medication for diabetes, and were recruited from 3 sources: i) probands from type 2 diabetic sibships from the Warren 2 sibling pairs described previously [12,13]; ii) a collection of type 2 diabetes cases (Warren 2 cases) diagnosed between 35-65 years, but not selected on family history; iii) and probands from a collection of

families that had either both parents available, or one parent and at least two siblings [7]. None of the subjects overlapped with those used in stage 1.

Population control subjects were all from the UK and of European descent. These were recruited from 3 sources: i) the remaining parents from a consecutive birth cohort (Exeter Family Study) with normal (<6.0mmol/l) fasting glucose and/or normal $HbA1_c$ levels (< 6%; Diabetes Control and Complications trial corrected) [12]; and ii) a nationally recruited population control sample of UK individuals of European descent obtained from the European cell culture collection (ECACC), and iii) a follow-up study, that is ongoing, of all people born in Great Britain during one week in 1958 (http://www.cls.ioe.ac.uk/Cohort/Ncds/mainncds.htm). Cases and families where the proband had high GAD autoantibody levels (>99th percentile of the normal population) were excluded from the study. Known subtypes of diabetes (e.g. MODY) were excluded by clinical criteria and/or genetic testing.

ADDITION case–control study

Cases were participants from the UK Cambridge arm of the ADDITION trial, which aims to evaluate whether screening for prevalent undiagnosed Type 2 diabetes is feasible, and whether subsequent optimised intensive treatment of diabetes is feasible and beneficial[14]. All cases were aged 40–69 and screen detected using OGTT and WHO diagnostic criteria. We used participants from the Ely study as controls—a population of white European men and women aged 35 to 79 years without diagnosed diabetes and from a similar population sampling frame as the Cambridge arm of the ADDITION study. The Medical Research Council (MRC) Ely Study is a population-based cohort study of the aetiology and pathogenesis of type 2 diabetes and related metabolic disorders in the UK[15]. All ELY participants underwent a standard anthropometric measurements and a 75-g OGTT and were defined as cases or controls based on their OGTT (WHO diagnostic criteria). For this analysis, the ADDITION case-control study comprised 926 cases and 1497 controls. The Cambridge Research Ethics Committee approved both studies.

Tayside case–control study

The Tayside case–control study comprises 3,745 individuals with type 2 diabetes and 3,786 controls from the Wellcome Trust UK type 2 diabetes case-control collection (Go-DARTS2) which is a sub-study of Diabetes Audit and Research Tayside (DARTS)[16-22]. All cases were of European descent with physician-diagnosed type 2 diabetes. They were recruited at primary or secondary care diabetes clinics or invited to participate from primary care registers from throughout the Tayside region of Scotland. The cases have not been characterised for GAD anti-bodies or MODY gene mutations. The controls were invited to participate through the primary care physicians or through their workplace occupational health departments. Controls did not have a previous physician-based diagnosis of diabetes. Control individuals with an HbA1c at recruitment of >6.2% or a fasting glucose of >=7 were removed from the analysis. All individuals in this ongoing study were recruited in Tayside between 1st October 2004 and 1st July 2006. For the purposes of this analysis, we excluded all participants below the age of 35 years, leaving 3,728 cases and 3,732 controls for analysis. The Tayside Medical Ethics Committee approved the study. Informed consent was obtained from all participants.

Statistical Analysis

Hardy-Weinberg was assessed using the χ^2 statistic (1 df). We used standard log likelihood ratio tests to assess the contribution of individual SNPs (under a log additive model (1 df)) to risk of type 2 diabetes in logistic regression analysis. We also used log likelihood ratio tests to assess whether statistically associated SNPs independently contributed to risk of type 2 diabetes and to determine the source of any association signal. Specifically, we compared the log likelihood of a nested model (2 df) with that of the full model (3 df) by consecutively adding the other SNPs (in a log additive form) to a model containing SNP rs10010131 (general inheritance (2df) form). We also conducted a reciprocal analysis, adding SNP rs10010131 (1 df form) to a model with each of the other SNPs (2 df form). We used pairwise correlation (r^2) to assess the extent of linkage disequilibrium among co-located SNPs. Haplotype analysis was

done using log-linear modelling embedded within an EM algorithm. The log-linear model was based on iterative proportional fitting[23].

For the pooled analysis, we used logistic regression with study as categorical covariate (equivalent to a Mantel-Haenszel model weighted by study). Heterogeneity among studies was assessed using the Q statistic. Single locus and haplotype analyses were done using Stata 8.2. For the calculation of excess familial risk, we assumed a familial risk of two.

References

1. Ke, X. et al. The impact of SNP density on fine-scale patterns of linkage disequilibrium. *Hum Mol Genet* **13**, 577-88 (2004).
2. Fan, J.B. et al. Highly parallel SNP genotyping. *Cold Spring Harb Symp Quant Biol* **68**, 69-78 (2003).
3. Rathmann, W. et al. Low faecal elastase 1 concentrations in type 2 diabetes mellitus. *Scand J Gastroenterol* **36**, 1056-61 (2001).
4. Harding, A.H. et al. Dietary fat and the risk of clinical type 2 diabetes: the European prospective investigation of Cancer-Norfolk study. *Am J Epidemiol* **159**, 73-82 (2004).
5. Day, N. et al. EPIC-Norfolk: study design and characteristics of the cohort. European Prospective Investigation of Cancer. *Br J Cancer* **80 Suppl 1**, 95-103 (1999).
6. Owen, K.R., Stride, A., Ellard, S. & Hattersley, A.T. Etiological investigation of diabetes in young adults presenting with apparent type 2 diabetes. *Diabetes Care* **26**, 2088-93 (2003).
7. Frayling, T. et al. Parent-offspring Trios: a resource to facilitate the identification of Type 2 diabetes genes. *Diabetes* **48**, 2475-9 (1999).
8. Knight, B., Shields, B.M. & Hattersley, A.T. The Exeter Family Study of Childhood Health (EFSOCH): study protocol and methodology. *Paediatric And Perinatal Epidemiology* **20**, 172-179 (2006).
9. Permutt, M.A. et al. A genome scan for type 2 diabetes susceptibility loci in a genetically isolated population. *Diabetes* **50**, 681-5 (2001).
10. Weedon, M. et al. Common variants of the HNF4alpha P2 promoter are associated with type 2 diabetes in the UK population. *Diabetes* **53**, 3002-6 (2004).
11. WHO Study Group. Report of a WHO Consultation: Part 1: Diagnosis and Classification of Diabetes Mellitus. (World Health Organisation, Geneva, 1999).
12. Minton, J.A. et al. Association studies of genetic variation in the WFS1 gene and type 2 diabetes in U.K. populations. *Diabetes* **51**, 1287-90 (2002).
13. Wiltshire, S. et al. A genomewide scan for loci predisposing to type 2 diabetes in a U.K. population (the Diabetes UK Warren 2 Repository): analysis of 573 pedigrees provides independent replication of a susceptibility locus on chromosome 1q. *Am J Hum Genet* **69**, 553-69. (2001).
14. Lauritzen, T. et al. The ADDITION study: proposed trial of the cost-effectiveness of an intensive multifactorial intervention on morbidity and mortality among people with Type 2 diabetes detected by screening. *Int J Obes Relat Metab Disord* **24 Suppl 3**, S6-11 (2000).
15. Wareham, N.J. et al. A quantitative analysis of the relationship between habitual energy expenditure, fitness and the metabolic cardiovascular syndrome. *Br J Nutr* **80**, 235-41 (1998).
16. Doney, A. et al. Haplotype analysis of the PPARgamma Pro12Ala and C1431T variants reveals opposing associations with body weight. *BMC Genet* **3**, 21 (2002).

17. Doney, A.S. et al. Association of the Pro12Ala and C1431T variants of PPARG and their haplotypes with susceptibility to Type 2 diabetes. *Diabetologia* **47**, 555-8 (2004).
18. Doney, A.S. et al. Male preponderance in early diagnosed type 2 diabetes is associated with the ARE insertion/deletion polymorphism in the PPP1R3A locus. *BMC Genet* **4**, 11 (2003).
19. Doney, A.S. et al. Association of common variation in the PPARA gene with incident myocardial infarction in individuals with type 2 diabetes: A Go-DARTS study. *Nucl Recept* **3**, 4 (2005).
20. Doney, A.S., Fischer, B., Leese, G., Morris, A.D. & Palmer, C.N. Cardiovascular risk in type 2 diabetes is associated with variation at the PPARG locus: a Go-DARTS study. *Arterioscler Thromb Vasc Biol* **24**, 2403-7 (2004).
21. Doney, A.S., Lee, S., Leese, G.P., Morris, A.D. & Palmer, C.N. Increased cardiovascular morbidity and mortality in type 2 diabetes is associated with the glutathione S transferase theta-null genotype: a Go-DARTS study. *Circulation* **111**, 2927-34 (2005).
22. Morris, A.D. et al. The diabetes audit and research in Tayside Scotland (DARTS) study: electronic record linkage to create a diabetes register. DARTS/MEMO Collaboration. *Bmj* **315**, 524-8 (1997).
23. Mander, A.P. Haplotype analysis in population-based association studies. *The Stata Journal* **1**, 58-75 (2001).

Supplementary Table S1. Minor allele frequencies for all statistically associated SNPs (P < 0.01) in genes involved in β-cell development, growth, function and survival in UK populations (comprising up to 1,856 controls) and an Ashkenazi population (comprising up to 461 controls)

Gene	SNP	MAF in UK control populations	MAF in Ashkenazi control population
CHGA	rs941584	0.24	0.27
NFATC1	rs643705	0.13	0.30
PAX6	rs628224	0.17	0.12
NFKB1	rs1609798	0.41	0.32
NFKB1	rs11722146	0.25	0.20
NFKB1	rs230498	0.32	0.27
WFS1	rs10010131	0.37	0.31
CACNA1D	rs4687736	0.31	0.28
EGFR	rs2075112	0.40	0.36
PBX1	rs7535186	0.49	0.41
WFS1	rs6446482	0.42	0.33
NFKB1	rs230539	0.39	0.45
TCF2	rs7501939	0.06	0.16
CACNA1D	rs3796347	0.36	0.31
CAMK2A	rs3822607	0.32	0.45
NFATC1	rs3826567	0.41	0.30
FOXA3	rs11669442	0.40	0.43
FGF2	rs1048201	0.16	0.18

All SNPs were in Hardy-Weinberg equilibrium (P > 0.01 in controls)

Supplementary Table S2. Association between SNPs located in the *WFS1* gene and risk of type 2 diabetes: initial study populations comprising up to 2,414 cases and 2,317 controls

SNP	Odds ratio	95 % CI	P-value*
rs10010131	0.84	0.77–0.92	1.3×10^{-4}
rs6446482	0.85	0.78–0.93	2.7×10^{-4}
rs4689391	0.86	0.79–0.94	9.6×10^{-4}
rs3821943	0.89	0.81–0.96	5.0×10^{-3}
rs1801212	0.89	0.81–0.98	0.015
rs752854	0.94	0.85–1.03	0.164

CI, confidence interval

* Based on a single locus log additive model adjusted for study

All SNPs were in Hardy-Weinberg equilibrium in controls (P > 0.01)

Supplementary Table S3. Linkage disequilibrium (r^2) among *WFS1* SNPs genotyped in initial study populations: 2,317 controls

	rs10010131	rs6446482	rs4689391	rs3821943	rs1801212	rs752854
rs10010131	-					
rs6446482	0.98	-				
rs4689391	0.94	0.93	-			
rs3821943	0.88	0.87	0.92	-		
rs1801212	0.76	0.74	0.72	0.68	-	
rs752854	0.86	0.84	0.82	0.76	0.84	-

Correlations were not materially different in UK and Ashkenazi populations (data not shown)

Supplementary Table S4. Log likelihood ratio tests assessing the conditional association among SNPs in the *WFS1* gene with risk of type 2 diabetes: initial study populations comprising up to 2,414 cases and 2,317 controls

SNP	P value (1 df)*	P-value (1 df)**
rs10010131	–	–
rs6446482	0.254	0.053
rs4689391	0.524	0.047
rs3821943	0.290	5.1×10^{-3}
rs1801212	0.457	2.0×10^{-3}
rs752854	1.7×10^{-3}	4.3×10^{-6}

CI, confidence interval, all models include a study variable

* P value (log likelihood ratio test) for addition of SNP (log additive) to model containing SNP rs10010131 (2df)

** P value (log likelihood ratio test) for addition of SNP rs10010131 (log additive) to model containing SNP (2 df)

Supplementary Table S5. Genotype counts by case–control status and study for rs10010131 and rs6446482

	rs10010131*		rs6446482*	
Study	Controls	Cases	Controls	Cases
Cambridge	187/241/101	222/239/87	182/247/99	216/239/92
EPIC	275/332/133	137/165/52	264/343/133	133/169/52
EXETER	197/279/104	217/281/76	190/285/104	199/285/71
ASHKENAZI	205/205/45	443/358/59	199/203/45	439/356/63
ADDITION	540/691/254	376/402/137	527/704/259	369/410/141
WARREN 2	1314/1796/621	920/1086/369	1284/1823/641	898/1123/388
Tayside	1309/1809/597	1438/1690/594	1290/1825/599	1416/1712/585

* Each cell represents common homozygotes/heterozygotes/rare homozygotes, respectively