Replication of Genome-Wide Association Signals in UK Samples Reveals Risk Loci for Type 2 Diabetes

Eleftheria Zeggini, *et al.*

Science **316**, 1336 (2007);
DOI: 10.1126/science.1142364

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005.

Study design: Richa Saxena[1–6] and Valeriya Lyssenko[7] (Team Leaders), Peter Almgren,[7] Paul I. W. de Bakker,[1–6] Noël P. Burtt,[1] Jose C. Florez,[1–6] Hong Chen,[8] Joanne Meyer,[8] Joel N. Hirschhorn,[1,6,9–11] Mark J. Daly,[1–3,5] Thomas E. Hughes,[8] Leif Groop,[7,12] David Altshuler[1–6] (Chair)

Clinical characterization and phenotypes: Valeriya Lyssenko[7] and Richa Saxena[1–6] (Team Leaders), Peter Almgren,[7] Kristin Ardlie,[1] Kristina Bengtsson Boström,[13] Noël P. Burtt,[1] Hong Chen,[8] Jose C. Florez,[1–6] Bo Isomaa,[14,15] Sekar Kathiresan,[1,3,5] Guillaume Lettre,[1,6,9–11] Ulf Lindblad,[16] Helen N. Lyon,[1,6,9–11] Olle Melander,[7] Christopher Newton-Cheh,[1–3,5] Peter Nilsson,[17] Marju Orho-Melander,[7] Lennart Råstam,[16] Elizabeth K. Speliotes,[1,3,6,9–11] Marja-Riitta Taskinen,[12] Tiinamaija Tuomi,[12,15] Benjamin F. Voight,[1–3,5] David Altshuler,[1–6] Joel N. Hirschhorn,[1,6,9–11] Thomas E. Hughes,[8] Leif Groop[7,12] (Chair)

DNA sample QC and diabetes replication genotyping: Candace Guiducci[1] and Valeriya Lyssenko[7] (Team Leaders), Anna Berglund,[7] Joyce Carlson,[18] Lauren Gianniny,[1] Rachel Hackett,[1] Liselotte Hall,[18] Johan Holmkvist,[7] Esa Laurila,[7] Marju Orho-Melander,[7] Marketa Sjögren,[7] Maria Sterner,[18] Aarti Surti[1] Margareta Svensson,[7] Malin Svensson,[7] Ryan Tewhey,[1] Noël P. Burtt[1] (Chair)

Whole genome scan genotyping: Brendan Blumenstiel[1] (Team Leader), Melissa Parkin,[1] Matthew DeFelice,[1] Candace Guiducci,[1] Ryan Tewhey,[1] Rachel Barry,[1] Wendy Brodeur,[1]Noël P. Burtt,[1] Jody Camarata,[1] Nancy Chia,[1] Mary Fava,[1]John Gibbons,[1] Bob Handsaker,[1] Claire Healy,[1] Kieu Nguyen,[1] Casey Gates,[1] Carrie Sougnez,[1] Diane Gage,[1] Marcia Nizzari,[1] David Altshuler,[1–6] Stacey B. Gabriel[1] (Chair)

GCKR replication genotyping and analysis (Malmö Diet and Cancer Study): Sekar Kathiresan[1,3,5] (Team Leader), Candace Guiducci,[1] Aarti Surti,[1] Noël P. Burtt,[1] Olle Melander,[7] Marju Orho-Melander[7] (Chair)

Statistical analysis: Benjamin F. Voight[1–3,5] and Paul I. W. de Bakker[1–6] (Team Leaders), Richa Saxena,[1–6] Valeriya Lyssenko,[7] Peter Almgren,[7] Noël P. Burtt,[1] Hong Chen,[8] Gung-Wei Chirn,[8] Qicheng Ma,[8] Hemang Parikh,[7] Delwood Richardson,[8] Darrell Ricke,[8] Jeffrey J. Roix,[8] Leif Groop,[7,12] Shaun Purcell,[1,2] David Altshuler,[1–6] Mark J. Daly[1–3,5] (Chair)

[1]Broad Institute of Harvard and Massachusetts Institute of Technology (MIT), Cambridge, MA 02142, USA. [2]Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA 02114, USA. [3]Department of Medicine, Massachusetts General Hospital, Boston, MA 02114, USA. [4]Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114, USA. [5]Department of Medicine, Harvard Medical School, Boston, MA 02115, USA. [6]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. [7]Department of Clinical Sciences, Diabetes and Endocrinology Research Unit, University Hospital Malmö, Lund University, Malmö, Sweden. [8]Diabetes and Metabolism Disease Area, Novartis Institutes for BioMedical Research, 100 Technology Square, Cambridge, MA 02139, USA. [9]Department of Pediatrics, Harvard Medical School, Boston, MA 02115, USA. [10]Division of Endocrinology, Children's Hospital, Boston, MA 02115, USA. [11]Division of Genetics, Children's Hospital, Boston, MA 02115, USA. [12]Department of Medicine, Helsinki University Hospital, University of Helsinki, Helsinki, Finland. [13]Skaraborg Institute, Skövde, Sweden. [14]Malmska Municipal Health Center and Hospital, Jakobstad, Finland. [15]Folkhälsan Research Center, Helsinki, Finland. [16]Department of Clinical Sciences, Community Medicine Research Unit, University Hospital Malmö, Lund University, Malmö, Sweden. [17]Department of Clinical Sciences, Medicine Research Unit, University Hospital Malmö, Lund University, Malmö, Sweden. [18]Clinical Chemistry, University Hospital Malmö, Lund University, Malmö, Sweden. [19]Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02115, USA.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1142358/DC1
Materials and Methods
Figs. S1 and S2
Tables S1 to S6
References

9 March 2007; accepted 20 April 2007
Published online 26 April 2007;
10.1126/science.1142358
Include this information when citing this paper.

Replication of Genome-Wide Association Signals in UK Samples Reveals Risk Loci for Type 2 Diabetes

Eleftheria Zeggini,[1,2]* Michael N. Weedon,[3,4]* Cecilia M. Lindgren,[1,2]* Timothy M. Frayling,[3,4]* Katherine S. Elliott,[2] Hana Lango,[3,4] Nicholas J. Timpson,[2,5] John R. B. Perry,[3,4] Nigel W. Rayner,[1,2] Rachel M. Freathy,[3,4] Jeffrey C. Barrett,[2] Beverley Shields,[4] Andrew P. Morris,[2] Sian Ellard,[4,6] Christopher J. Groves,[1] Lorna W. Harries,[4] Jonathan L. Marchini,[7] Katharine R. Owen,[1] Beatrice Knight,[4] Lon R. Cardon,[2] Mark Walker,[8] Graham A. Hitman,[9] Andrew D. Morris,[10] Alex S. F. Doney,[10] The Wellcome Trust Case Control Consortium (WTCCC),† Mark I. McCarthy,[1,2]‡§ Andrew T. Hattersley[3,4]‡

The molecular mechanisms involved in the development of type 2 diabetes are poorly understood. Starting from genome-wide genotype data for 1924 diabetic cases and 2938 population controls generated by the Wellcome Trust Case Control Consortium, we set out to detect replicated diabetes association signals through analysis of 3757 additional cases and 5346 controls and by integration of our findings with equivalent data from other international consortia. We detected diabetes susceptibility loci in and around the genes *CDKAL1*, *CDKN2A/CDKN2B*, and *IGF2BP2* and confirmed the recently described associations at *HHEX/IDE* and *SLC30A8*. Our findings provide insight into the genetic architecture of type 2 diabetes, emphasizing the contribution of multiple variants of modest effect. The regions identified underscore the importance of pathways influencing pancreatic beta cell development and function in the etiology of type 2 diabetes.

The pathophysiological basis of type 2 diabetes (T2D) remains unclear despite its growing global importance (*1*). Candidate gene and positional cloning efforts have suggested many putative susceptibility variants, but unequivocal replications are so far limited to variants in just three genes: *PPARG*, *KCNJ11*, and *TCF7L2* (*2*–*4*).

Improved understanding of the correlation between genetic variants [linkage disequilibrium (LD)], allied to advances in genotyping technology, have enabled systematic searches for disease-associated common variants on a genome-wide scale. The Wellcome Trust Case Control Consortium (WTCCC) recently completed such a genome-wide association (GWA) scan in 1924 T2D cases and 2938 population controls from the United Kingdom, using the Affymetrix GeneChip Human Mapping 500 k Array Set (*5*). The strongest association signals genome-wide were observed for single-nucleotide polymorphisms (SNPs) in *TCF7L2*. [For example, for rs7901695, odds ratio (OR) = 1.37, 95% confidence interval (CI) = 1.25–1.49, and $P = 6.7 \times 10^{-13}$.] The other known T2D susceptibility variants were detected with effect sizes consistent with previous reports (*2*, *3*).

Here, we describe how integration of data from the WTCCC scan and our own replication studies with similar information generated by the Diabetes Genetics Initiative (DGI) (*6*) and the Finland–United States Investigation of NIDDM Genetics (FUSION) (*7*) has identified several additional susceptibility variants for T2D.

In the WTCCC study, analysis of 490,032 autosomal SNPs in 16,179 samples yielded 459,448 SNPs that passed initial quality control (*5*). We considered only the 393,453 autosomal SNPs with minor allele frequency (MAF) exceeding 1% in both cases and controls and no extreme departure from Hardy-Weinberg equilibrium ($P < 10^{-4}$ in cases or controls) (*8*). This T2D-specific data set shows no evidence of substantial confounding from population substructure and genotyping biases (*8*).

To distinguish true associations from those reflecting fluctuations under the null or residual errors arising from aberrant allele calling, we first submitted putative signals from the WTCCC study to additional quality control, including cluster-plot visualization and validation genotyping on a second platform (*8*). Next, we attempted replication of selected signals in up to 3757 additional cases and 5346 controls (replication sets RS1 to RS3). RS1 comprised 2022 cases and 2037 controls from the U.K. Type 2 Diabetes Genetics Consortium collection (UKT2DGC) (all from Tayside, Scotland). RS2 included 632 additional T2D cases and 1750 population controls from the Exeter Family Study of Child Health (EFSOCH). A subset of SNPs were typed in RS3, comprising a further 1103 cases and 1559 controls from the UKT2DGC (table S1).

The first wave of validated SNPs sent for replication was selected from the 30 SNPs, in nine distinct chromosomal regions (excluding *TCF7L2*), which had, in the WTCCC scan alone,

attained the most extreme ($P < 10^{-5}$) significance values on Cochran-Armitage tests of association. Genotyping of 21 representative SNPs generated evidence of replication ($P < 0.05$) for three of these nine regions (Table 1 and table S2).

Rs8050136 [mapping to the *FTO* (fat mass and obesity-associated) gene region on chr16] was among a cluster of SNPs generating the strongest evidence for association outside *TCF7L2* in the original scan [risk allele OR = 1.27 (1.16–1.37), $P = 2.0 \times 10^{-8}$] (fig. S1). This SNP showed strong replication [OR = 1.22 (1.12–1.32), $P = 5.4 \times 10^{-7}$]. As we recently reported (*9*), this effect on T2D risk is mediated through a primary effect on adiposity, and adjustment for body mass index (BMI) abolishes the T2D association.

Replication was also obtained for SNPs within the *CDKAL1* locus on chromosome 6, including rs9465871 and rs10946398. Although rs9465871 generated the stronger signal in the WTCCC scan, replication at this SNP was modest ($P = 0.023$). The replication signal at rs10946398 was more striking [OR = 1.14 (1.07–1.22), $P = 8.4 \times 10^{-5}$] (Table 1 and table S2). Consistent evidence of association is provided by the DGI ($P = 4.1 \times 10^{-4}$ at rs7754840) and FUSION groups ($P = 9.5 \times 10^{-3}$ at rs471253) (Table 1 and table S3) (*6*, *7*), both SNPs being strong ($r^2 > 0.99$) proxies for rs10946398. Across all studies, combined evidence for association at *CDKAL1* is compelling ($P \sim 4.1 \times 10^{-11}$).

All associated SNPs map to a large (90 kb) intron within *CDKAL1* (Fig. 1). Flanking recombination hotspots define a 200-kb interval likely to contain the etiological variant(s). *CDKAL1* [cyclin-dependent kinase 5 (CDK5) regulatory subunit associated protein 1–like 1] encodes a 579-residue, 65-kD protein of unknown function. We have detected expression of *CDKAL1* mRNA in human pancreatic islet and skeletal muscle (fig. S2). CDKAL1 shares considerable protein domain and amino acid homology with CDK5 regulatory subunit associated protein 1 (CDK5RAP1), a known inhibitor of CDK5 activation. CDK5 has been implicated in the regulation of pancreatic beta cell function through formation of p35/CDK5 complexes that down-regulate insulin expression (*10*, *11*).

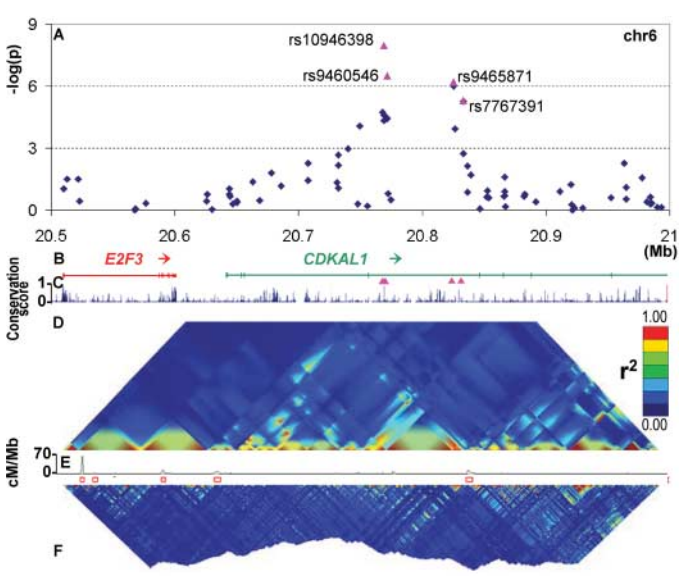
The third replicated association maps to the *HHEX* (homeobox, hematopoietically expressed) gene region on chromosome 10. This gene showed strong association in the WTCCC scan [rs5015480: risk allele OR = 1.22 (CI, 1.12–1.33), $P = 5.4 \times 10^{-6}$] and is a powerful biological candidate (*12*, *13*). We could not optimize a replication assay for rs5015480 but observed evidence for replication at a perfect proxy, rs1111875 [risk allele OR = 1.08 (CI, 1.01–1.15), $P = 0.02$] (Table 1, tables S2 and S3). Both DGI and FUSION studies showed modest but consistent association signals generating strong combined evidence ($P \sim 5.7 \times 10^{-10}$) for a role in T2D susceptibility (Table 1 and table S3). A fourth genome-wide association scan, in French subjects, recently yielded independent evidence for a T2D signal in this region (*14*). The signal resides within an extended (295 kb) region of LD containing not only *HHEX* [highly expressed in fetal and adult pancreas (fig. S2)] but also the genes encoding kinesin-interacting factor (*KIF11*) and insulin-degrading enzyme (*IDE*) (fig. S3). *IDE* represents a second strong biological candidate given postulated effects on both insulin signaling and islet function and data from rodent models (*15*–*17*).

Of the remaining regions selected in the first wave, none showed any evidence of replication in U.K. samples (table S2), and for none was there strong support from the DGI and FUSION scans.

The relatively strict thresholds imposed for SNP selection in the first wave (i.e., point-wise

[1]Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Churchill Hospital, Oxford, OX3 7LJ, UK. [2]The Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford, OX3 7BN, UK. [3]Genetics of Complex Traits, Institute of Biomedical and Clinical Science, Peninsula Medical School, Magdalen Road, Exeter EX1 2LU, UK. [4]Diabetes Genetics Group, Institute of Biomedical and Clinical Science, Peninsula Medical School, Barrack Road, Exeter EX2 5DW, UK. [5]Medical Research Council Centre for Causal Analyses in Translational Epidemiology, Bristol University, Canynge Hall, Whiteladies Rd, Bristol, BS2 8PR, UK. [6]The Molecular Genetics Laboratory, Royal Devon and Exeter NHS Foundation Trust, Old Pathology Building, Barrack Road, Exeter, EX2 5DW, UK. [7]Department of Statistics, University of Oxford, 1 South Parks Road, Oxford, OX1 3TG, UK. [8]Diabetes Research Group, School of Clinical Medical Sciences, Newcastle University, Framlington Place, Newcastle upon Tyne NE2 4HH, UK. [9]Centre for Diabetes and Metabolic Medicine, Barts and The London, Royal London Hospital, Whitechapel, London, E1 1BB, UK. [10]Diabetes Research Group, Division of Medicine and Therapeutics, Ninewells Hospital and Medical School, Dundee, DD1 9SY, UK.

*These authors contributed equally to this work.
†Membership of the WTCCC is listed at the end of this paper.
‡These authors contributed equally to this work.
§To whom correspondence should be addressed. E-mail: mark.mccarthy@drl.ox.ac.uk



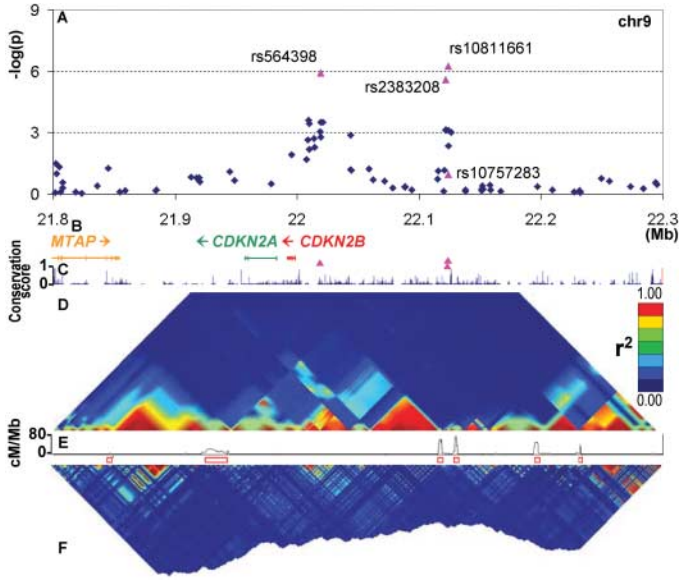


Fig. 1. (**left**) Overview of *CDKAL1* signal region. (**A**) Plot of –log(*P* values) for T2D (Cochran-Armitage test for trend) against chromosome position in Mb. Blue diamonds represent primary scan results and pink triangles denote meta-analysis results across all UK samples. (**B**) Genomic location of genes showing intron and exon structure (NCBI Build 35). Pink triangles show position of replication SNPs relative to gene structure. (**C**) MULTIZ (*24*) vertebrate alignment of 17 species showing evolutionary conservation. (**D**) GoldSurfer2 (*25*) 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 (*26*). (**F**) GoldSurfer2 plot of linkage disequilibrium (r^2) for all HapMap (haplotype map of the human genome) SNPs across the region (HapMap CEU data) (*27*). **Fig. 2.** (**right**) Overview of chr9 signal region. Panel layout as in Fig. 1.

Table 1. Confirmed T2D susceptibility variants. Representative SNPs are shown for each signal with ORs and 95% CIs reported (for the Cochran-Armitage 1 df test) with respect to the risk allele (denoted in bold, with the ancestral allele underlined where known). SNPs selected for inclusion are those with the strongest evidence for association in the U.K. data sets (except in the case of *TCF7L2*, where, to maximize consistency across the data sets, rs7901695 is presented). In the case of *HHEX*, the U.K. meta-analysis combines data from rs5015480 and rs1111875 (r^2 = 1 in HapMap CEU). Because DGI and FUSION had not typed the identical SNPs in all samples, results shown for those studies feature the SNP generating the strongest association: In all cases, these were SNPs in strong LD (minimum r^2 0.95, except *TCF7L2*) and with consistent direction of effect with the SNP reported in the U.K. data (see table S3 for details). The use of different SNPs may result in slightly different estimates of *P* values and OR between the three studies. Combined estimates of the ORs were calculated by weighting the logORs of each study by the inverse of their variance.

rs	chr	position	A1	A2	Region	WTCCC 1924 cases 2938 controls OR (95% CI)	P_{add}	Replication meta-analysis 3757 cases 5346 controls OR (95% CI)	P_{add}	All UK sample meta-analysis 5681 cases 8284 controls OR (95% CI)	P_{add}	DGI 6529 cases 7252 controls OR (95% CI)	P_{add}	FUSION 2376 cases 2432 controls OR (95% CI)	P_{add}	All combined 14,586 cases 17,968 controls OR (95% CI)	P_{add}
rs8050136	16	52373776	**A**	C	*FTO*	1.27 (1.16–1.37)	2.0×10^{-8}	1.22 (1.12–1.32)	5.4×10^{-7}	1.23 (1.18–1.32)	7.3×10^{-14}	1.03 (0.91–1.17)	0.25	1.11 (1.02–1.20)	0.017	1.17 (1.12–1.22)	1.3×10^{-12}
rs10946398	6	20769013	A	**C**	*CDKAL1*	1.20 (1.10–1.31)	2.5×10^{-5}	1.14 (1.07–1.22)	8.3×10^{-5}	1.16 (1.10–1.22)	1.3×10^{-8}	1.08 (1.03–1.14)	2.4×10^{-3}	1.12 (1.03–1.22)	9.5×10^{-3}	1.12 (1.08–1.16)	4.1×10^{-11}
rs5015480	10	94455539	**C**	T	*HHEX*	1.22 (1.12–1.33)	5.4×10^{-6}	–	–	1.13 (1.07–1.19)	4.6×10^{-6}	1.14 (1.06–1.22)	1.7×10^{-4}	1.10 (1.01–1.19)	0.025	1.13 (1.08–1.17)	5.7×10^{-10}
rs1111875	10	94452862	**C**	T	*HHEX*	–	–	1.08 (1.01–1.15)	0.020								
rs10811661	9	22124094	C	**T**	*CDKN2B*	1.22 (1.09–1.37)	7.6×10^{-4}	1.18 (1.08–1.28)	1.7×10^{-4}	1.19 (1.11–1.28)	4.9×10^{-7}	1.20 (1.12–1.28)	5.4×10^{-8}	1.20 (1.07–1.36)	2.2×10^{-3}	1.20 (1.14–1.25)	7.8×10^{-15}
rs564398	9	22019547	C	**T**	*CDKN2B*	1.16 (1.07–1.27)	3.2×10^{-4}	1.12 (1.05–1.19)	8.6×10^{-4}	1.13 (1.08–1.19)	1.3×10^{-6}	1.05 (0.94–1.17)	0.5	1.13 (1.01–1.27)	0.039	1.12 (1.07–1.17)	1.2×10^{-7}
rs4402960	3	186994389	G	**T**	*IGF2BP2*	1.15 (1.05–1.25)	1.7×10^{-3}	1.09 (1.01–1.16)	0.018	1.11 (1.05–1.16)	1.6×10^{-4}	1.17 (1.11–1.23)	1.7×10^{-9}	1.18 (1.08–1.28)	2.4×10^{-4}	1.14 (1.11–1.18)	8.6×10^{-16}
rs13266634	8	118253964	**C**	T	*SLC30A8*	1.12 (1.02–1.23)	0.020	1.12 (1.04–1.19)	1.2×10^{-3}	1.12 (1.05–1.18)	7.0×10^{-5}	1.07 (1.00–1.16)	0.047	1.18 (1.09–1.29)	7.0×10^{-5}	1.12 (1.07–1.16)	5.3×10^{-8}
rs7901695	10	114744078	**C**	T	*TCF7L2*	1.37 (1.25–1.49)	6.7×10^{-13}	–	–	–	–	1.38 (1.31–1.46)	2.3×10^{-31}	1.34 (1.21–1.49)	1.4×10^{-8}	1.37 (1.31–1.43)	1.0×10^{-48}
rs5215	11	17365206	**C**	T	*KCNJ11*	1.15 (1.05–1.25)	1.3×10^{-3}	–	–	–	–	1.15 (1.09–1.21)	1.0×10^{-7}	1.11 (1.02–1.20)	0.014	1.14 (1.10–1.19)	5.0×10^{-11}
rs1801282	3	12368125	**C**	G	*PPARG*	1.23 (1.09–1.41)	1.3×10^{-3}	–	–	–	–	1.09 (1.01–1.16)	0.019	1.20 (1.07–1.33)	1.4×10^{-3}	1.14 (1.08–1.20)	1.7×10^{-6}

$P < 10^{-5}$) help to limit false discovery, but many genuine susceptibility variants will fail to reach them. We initiated a second wave of replication based around SNPs for which the WTCCC scan generated more modest evidence for association (Cochran-Armitage $P \sim 10^{-2}$ to 10^{-5}). We prioritized the 5367 SNPs in this range using additional criteria: (i) evidence of association in DGI and FUSION (*6*, *7*); (ii) presence of multiple, independent ($r^2 < 0.4$) associations within the same locus; and (iii) biological candidacy (*8*, *18*).

Analysis of the 56 SNPs, representing 49 putative signals, selected for this "second wave" of replication (table S4) yielded two further regions implicated in T2D susceptibility. A cluster of SNPs on chromosome 9 (represented by rs10811661) generated a promising signal in all three scans. Replication was observed in UK samples [rs10811661: OR = 1.18 (CI, 1.08–1.28), $P = 1.7 \times 10^{-4}$), as well as DGI ($P = 2.2 \times 10^{-5}$) and FUSION follow-up studies (rs2383208, $P = 9.7 \times 10^{-3}$). A second signal from the WTCCC scan located ~100 kb 5′ [rs564398, OR = 1.16 (CI, 1.07–1.27), $P = 3.2 \times 10^{-4}$], was weakly supported in the FUSION scan but not the DGI scan (Table 1 and table S3), and was replicated in the U.K. RS samples [OR = 1.12 (CI, 1.05–1.19), $P = 8.6 \times 10^{-4}$] (Table 1 and table S3).

These two association signals are separated by a recombination hotspot (D′ between rs10811661 and rs564398 is 0.057, $r^2 < 0.001$) (Fig. 2). Across all studies, the combined evidence for association is stronger for the 3′ ($P \sim 7.8 \times 10^{-15}$) than for the 5′ ($P \sim 1.2 \times 10^{-7}$) peak (Table 1). The 3′ signal maps to sequence with no characterized genes, whereas the recombination interval enclosing the 5′ signal includes the full coding sequences of *CDKN2B* and *CDKN2A* (encoding $p15^{INK4b}$ and $p16^{INK4a}$, respectively). *CDKN2A* is a known tumor suppressor and its product, $p16^{INK4a}$, inhibits CDK4 (cyclin-dependent kinase 4), a powerful regulator of pancreatic beta cell replication (*19*–*21*). Overexpression of *Cdkn2a* leads to decreased islet proliferation in aging mice (*22*). *Cdkn2b* overexpression is also causally related to islet hypoplasia and diabetes in murine models (*23*). Both *CDKN2B* and *CDKN2A* display high levels of expression in pancreatic islets and pituitary (fig. S2).

A fifth replicated association lies within the *IGF2BP2* gene on chromosome 3. We observed some evidence of association for SNPs in this region in the WTCCC scan (*5*) [e.g., rs4402960: OR = 1.15 (CI, 1.05–1.25), $P = 1.7 \times 10^{-3}$]. Consistent associations in the DGI and FUSION scans (*6*, *7*) and the biological candidacy of the gene [a known regulator of insulin-like growth factor 2 (IGF2) translation] prompted replication. We obtained only modest evidence for replication at rs4402960 [OR = 1.09 (CI, 1.01–1.16), $P = 0.018$] (Table 1 and table S4), but combined evidence across all studies ($P \sim 8.6 \times 10^{-16}$) establishes this as a genuine T2D signal (Table 1 and table S3). The associated SNPs map to a 57-kb region spanning the promoter and first 2 exons of *IGF2BP2* (fig. S4).

Most of the remaining 50 "second-wave" SNPs can be discounted as susceptibility variants based on their failure to replicate (table S4), although some merit further consideration. One such example is rs9369425, located 57-kb downstream of the *VEGFA* (vascular endothelial growth factor A) gene on chromosome 6 (fig. S5). Evidence for association in the WTCCC scan [OR = 1.16 (CI, 1.06–1.27), $P = 8.6 \times 10^{-4}$] is supported by nominal replication in U.K. samples [OR = 1.08 (CI, 1.01–1.15), $P = 0.03$] and by DGI scan results [1.17 (1.04–1.32), $P = 4.4 \times 10^{-3}$]. Although no signal is apparent in the FUSION study, this does not allow us to reject the association. For 80% power to detect an OR of 1.11 ($\alpha = 0.05$), more than 3000 case-control pairs are needed.

In the French genome-wide scan (*14*), variants in both the *HHEX* and *SLC30A8* genes were implicated in T2D susceptibility. Because the associated SNPs in *SLC30A8* are poorly captured on the Affymetrix chip ($r^2 < 0.01$), the WTCCC scan was not informative for this locus. However, we genotyped rs13266634 independently and obtained replication of the finding [risk allele OR = 1.12 (CI, 1.05–1.18), $P = 7.0 \times 10^{-5}$ in all UK data] and across all three studies ($P \sim 5.3 \times 10^{-8}$) (Table 1 and table S4).

The present analysis has contributed to identification of several confirmed T2D susceptibility loci. One of these (*FTO*) exerts its primary effect on T2D risk through an impact on adiposity (*9*): None of the other signals was attenuated by adjustment for BMI or waist circumference (tables S5 to S7). One of the remaining four loci (*HHEX/IDE*) represents a strong replication of findings recently reported (*14*). The other three loci (near *CDKAL1*, *IGF2BP2*, and *CDKN2A*), all showing extensive replication across the three studies, represent previously unknown T2D susceptibility loci.

Across the four T2D scans completed (*5*–*7*, *14*), *TCF7L2* clearly emerges as the largest association signal. On current evidence, all other confirmed loci display more modest effect sizes (between 1.10 and 1.25 per allele). Extensive resequencing and fine-mapping will be required to define the full spectrum of etiological variation at each locus, and these may yet identify variants with greater impact. Our findings offer clear lessons for the design of future studies. Robust identification of variants with such effect sizes is only feasible with large-scale sample sets (13,965 individuals were typed in the present study). Further, the exchange of data between groups (providing data on up to 32,554 samples) was key to the rapid and unequivocal identification of the signals we report.

As a result of the four GWA studies reported to date (*5*–*7*, *14*), the number of genuine, replicated T2D susceptibility signals has climbed from three to nine (adding *HHEX/IDE*, *SLC30A8*, *CDKAL1*, *CDKN2A*, *IGF2BP2*, and *FTO*). However, these loci explain only a small proportion of the observed familiality (the sibling relative risk, λ_s, attributable to all loci in the U.K. samples, is only ~1.07). We expect additional loci to be revealed by further rounds of replication initiated by more systematic meta-analysis of these and other scans. Our study provides an important validation of the genome-wide indirect association mapping approach and a demonstration of the value of aggressive data-sharing efforts. It also generates insights into T2D pathogenesis, emphasizing the likely importance of pathways involved in pancreatic beta cell development, regeneration, and function. In-depth physiological and functional studies are now needed to establish the precise mechanisms involved.

References and Notes

1. M. Stumvoll, B. J. Goldstein, T. W. van Haeften, *Lancet* **365**, 1333 (2005).
2. D. Altshuler *et al.*, *Nat. Genet.* **26**, 76 (2000).
3. A. L. Gloyn *et al.*, *Diabetes* **52**, 568 (2003).
4. S. F. Grant *et al.*, *Nat. Genet.* **38**, 320 (2006).
5. Wellcome Trust Case Control Consortium, *Nature*, in press.
6. Diabetes Genetics Initiative, *Science* **316**, 1331 (2007); published online 26 April 2007 (10.1126/science.1142358).
7. L. J. Scott *et al.*, *Science* **316**, 1341 (2007); published online 26 April 2007 (10.1126/science.1142382).
8. Materials and methods are available as supporting material on *Science* Online.
9. T. M. Frayling *et al.*, *Science* **316**, 889; 10.1126/science.1141634 (2007).
10. M. Ubeda, J. M. Rukstalis, J. F. Habener, *J. Biol. Chem.* **281**, 28858 (2006).
11. F. Y. Wei *et al.*, *Nat. Med.* **11**, 1104 (2005).
12. A. C. Foley, M. Mercola, *Genes Dev.* **19**, 387 (2005).
13. R. Bort, J. P. Martinez-Barbera, R. S. Beddington, K. S. Zaret, *Development* **131**, 797 (2004).
14. R. Sladek *et al.*, *Nature* **445**, 881 (2007).
15. H. Fakhrai-Rad *et al.*, *Hum. Mol. Genet.* **9**, 2149 (2000).
16. K. A. Seta, R. A. Roth, *Biochem. Biophys. Res. Commun.* **231**, 167 (1997).
17. W. Farris *et al.*, *Am. J. Pathol.* **164**, 1425 (2004).
18. GeneSniffer (www.genesniffer.org) was used to prioritize genes for biological candidacy by integrating information from diverse online databases (including Entrez Gene, Online Mendelian Inheritance in Man, PubMed, and Mouse Genome Informatics).
19. S. G. Rane *et al.*, *Nat. Genet.* **22**, 44 (1999).
20. R. V. Mettus, S. G. Rane, *Oncogene* **22**, 8413 (2003).
21. N. Marzo *et al.*, *Diabetologia* **47**, 686 (2004).
22. J. Krishnamurthy *et al.*, *Nature* **443**, 453 (2006).
23. M. Moritani *et al.*, *Mol. Cell. Endocrinol.* **229**, 175 (2005).
24. M. Blanchette *et al.*, *Genome Res.* **14**, 708 (2004).
25. F. Pettersson, O. Jonsson, L. R. Cardon, *Bioinformatics* **20**, 3241 (2004).
26. S. Myers, L. Bottolo, C. Freeman, G. McVean, P. Donnelly, *Science* **310**, 321 (2005).
27. The International HapMap Consortium, *Nature* **437**, 1299 (2005).

28. We are grateful to all the study participants. We acknowledge the support of Diabetes UK, British Diabetic Association Research, the UK Medical Research Council, the Wellcome Trust, European Commission (EURODIA LSHG-CT-2004-518153), and Peninsula Medical School. Personal funding comes from the Wellcome Trust (E.Z., A.T.H., and L.R.C.); UK Medical Research Council (J.R.B.P.); Diabetes UK (R.M.F.); Throne-Holst Foundation (C.M.L.); Leverhulme Trust (A.P.M.); Research Councils UK (L.W.H.); and the Scottish Executive Generations Scotland Initiative (C.N.A.P., A.D.M.). M.N.W. is Vandervell Foundation Research Fellow at the Peninsula Medical School. We are grateful to members of the DGI and FUSION teams for sharing data.

Membership of Wellcome Trust Case Control Consortium
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[1,1] Andrew P. Morris,[3] Chris C. A. Spencer,[11] Martin D. Tobin,[1] Lon R. Cardon (Co-chair),[3] David G. Clayton (Co-chair)[2]

UK Blood Services and 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 P. M. 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 V. S. 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 C. L. 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 J. R. 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 C. A. Spencer,[11] Zhan Su,[11] Yik Ying Teo,[3,11] Damjan Vukcevic,[11] Peter Donnelly[11]

Principal investigators: David Bentley,[5]‡ 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 C. L. Gough,[50] Alistair S. Hall,[26] Andrew T. Hattersley,[42,43] Adrian V. S. 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]The 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 Street, 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]Social, Genetic and Developmental Psychiatry Centre, 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]Leeds Institute of Genetics, Health and Therapeutics, and Leeds Institute of Molecular Medicine, Faculty of Medicine and Health, University of Leeds, Leeds, LS1 3EX, UK. [27]Inflammatory Bowel Disease 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' National Health Service Foundation Trust, London, SE1 7EH, UK. [32]Department of Gastroenterology and 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]British Heart Foundation, 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 Arthritis Research Campaign, Epidemiology Research Unit, University of Manchester, Stopford Building, Oxford Road, 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]Medical Research Council Centre for Causal Analyses in Translational Epidemiology, Bristol University, Canynge Hall, Whiteladies Road, Bristol BS2 8PR, UK. [47]Medical Research Council Laboratories, Fajara, The Gambia. [48]Diamantina Institute for Cancer, Immunology and Metabolic Medicine, Princess Alexandra Hospital, University of Queensland, Woolloongabba, Queensland 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.

*Members of the Biologics in RA Genetics and Genomics Study Syndicate (BRAGGS) Steering Committee are listed after the WTCCC list.

†Members of the Breast Cancer Susceptibility Collaboration (UK) are listed after the BRAGGS list.

‡Present address: Illumina Cambridge, Chesterford Research Park, Little Chesterford, Near Saffron Walden, Essex, CB10 1XL, UK.

Biologics in RA Genetics and Genomics Study Syndicate (BRAGGS) Steering Committee Anne Barton,[1] John D Isaacs,[2] Ann W Morgan,[3] Gerry D. Wilson[4]

[1]arc Epidemiology Unit, University of Manchester, Oxford Road, Manchester, M13 9PT, UK. [2]Department of Rheumatology, University of Newcastle-Upon-Tyne, Framlington Place, Newcastle-Upon-Tyne NE2 4HH, UK. [3]Leeds Institute of Molecular Medicine, Section of Academic Unit of Musculoskeletal Disease Wellcome Trust Brenner Building, St. James's University Hospital, Beckett Street, Leeds LS9 7TF, UK. [4]Genomic Medicine, The University of Sheffield, Western Bank, Sheffield, S10 2TN, UK.

Breast Cancer Susceptibility Collaboration (UK) A. Ardern-Jones,[1] J. Berg,[2] A. Brady,[3] N. Bradshaw,[4] C. Brewer,[5]

G. Brice,[6] B. Bullman,[7] J. Campbell,[8] B. Castle,[9] R. Cetnarsyj,[8] C. Chapman,[10] C. Chu,[11] N. Coates,[12] T. Cole,[10] R. Davidson,[4] A. Donaldson,[13] H. Dorkins,[3] F. Douglas,[2] D. Eccles,[9] R. Eeles,[1] F. Elmslie,[6] D. G. Evans,[7] S. Goff,[6] S. Goodman,[5] D. Goudie,[2] J. Gray,[15] L. Greenhalgh,[16] H. Gregory,[17] S. V. Hodgson,[6] T. Homfray,[6] R. S. Houlston,[1] L. Izatt,[18] L. Jackson,[18] L. Jeffers,[19] V. Johnson-Roffey,[12] F. Kavalier,[18] C. Kirk,[19] F. Lalloo,[7] C. Langman,[18] I. Locke,[1] M. Longmuir,[4] J. Mackay,[20] A. Magee,[19] S. Mansour,[6] Z. Miedzybrodzka,[17] J. Miller,[11] P. Morrison,[19] V. Murday,[4] J. Paterson,[21] G. Pichert,[18] M. Porteous,[8] N. Rahman,[6] M. Rogers,[15] S. Rowe,[22] S. Shanley,[1] A. Saggar,[6] G. Scott,[2] L. Side,[23] L. Snadden,[4] M. Steel,[2] M. Thomas,[5] S. Thomas,[1]

[1]Clinical Genetics Service, Royal Marsden Hospital, Downs Road, Sutton, Surrey, SM2 5PT, UK. [2]Department of Clinical Genetics, Ninewells Hospital, Dundee, DD1 9SY, UK. [3]Medical and Community Genetics, Kennedy-Galton Centre, Level 8V, Northwick Park and St. Mark's NHS Trust, Watford Rd, Harrow, HA1 3UJ, UK. [4]Institute of Medical Genetics, Yorkhill NHS Trust, Dalnair Street, Glasgow, G3 8SJ, UK. [5]Clinical Genetics Department, Royal Devon and Exeter Hospital (Heavitree), Gladstone Road, Exeter, EX1 2ED, UK. [6]Department of Clinical Genetics, St. George's Hospital Medical School, Jenner Wing, Cranmer Terrace, London, SW17 0RE, UK. [7]Department of Medical Genetics, St. Mary's Hospital, Hathersage Road, Manchester, M13 0JH, UK. [8]South East of Scotland Clinical Genetics Service, Western General Hospital, Crewe Road, Edinburgh, EH4 2XU, UK. [9]Department of Medical Genetics, The Princess Anne Hospital, Coxford Road, Southampton, SO16 5YA, UK. [10]Clinical Genetics Unit, Birmingham Women's Hospital, Metchley Park Road, Edgbaston, Birmingham, B15 2TG, UK. [11]Yorkshire Regional Genetic Service, Department of Clinical Genetics, Cancer Genetics Building, St. James University Hospital, Beckett Street, Leeds, LS9 7TF, UK. [12]Department of Clinical Genetics, Leicester Royal Infirmary, Leicester, LE1 5WW, UK. [13]Department of Clinical Genetics, St Michael's Hospital, Southwell Street, Bristol, BS2 8EG, UK. [14]Institute of Human Genetics, International Centre for Life, Central Parkway, Newcastle upon Tyne, NE1 3BZ, UK. [15]Institute of Medical Genetics, University Hospital of Wales, Heath Park, Cardiff, CF14 4XW, UK. [16]Department of Clinical Genetics, Alder Hey Children's Hospital, Eaton Road, Liverpool L12 2AP, UK. [17]Clinical Genetics Centre, Argyll House, Foresterhill, Aberdeen, AB25 2ZR, UK. [18]Clinical Genetics, 7th Floor New Guy's House, Guy's Hospital, St. Thomas Street, London, SE1 9RT, UK. [19]Clinical Genetics Service, Belfast City Hospital Trust, Belvoir Park Hospital, Lisburn Road, Belfast, BT9 7AB, UK. [20]Clinical and Medical Genetics Unit, Institute of Child Health, 30 Guildford Street, London, WC1N 1EH, UK. [21]Department of Clinical Genetics, Addenbrooke's NHS Trust, Box 134, Hills Road, Cambridge, CB2 2QQ, UK. [22]Department of Clinical Genetics, Moston Lodge, Countess of Chester Hospital, Liverpool Road, Chester, CH2 1UL, UK. [23]Department of Clinical Genetics, Churchill Hospital, Old Road, Headington, Oxford OX3 7LJ, UK.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1142364/DC1
Materials and Methods
Figs. S1 to S8
Tables S1 to S10
References

9 March 2007; accepted 20 April 2007
Published online 26 April 2007;
10.1126/science.1142364
Include this information when citing this paper.

A Genome-Wide Association Study of Type 2 Diabetes in Finns Detects Multiple Susceptibility Variants

Laura J. Scott,[1] Karen L. Mohlke,[2] Lori L. Bonnycastle,[3] Cristen J. Willer,[1] Yun Li,[1] William L. Duren,[1] Michael R. Erdos,[3] Heather M. Stringham,[1] Peter S. Chines,[3] Anne U. Jackson,[1] Ludmila Prokunina-Olsson,[3] Chia-Jen Ding,[1] Amy J. Swift,[3] Narisu Narisu,[3] Tianle Hu,[1] Randall Pruim,[4] Rui Xiao,[1] Xiao-Yi Li,[1] Karen N. Conneely,[1] Nancy L. Riebow,[3] Andrew G. Sprau,[3] Maurine Tong,[3] Peggy P. White,[1] Kurt N. Hetrick,[5] Michael W. Barnhart,[5] Craig W. Bark,[5] Janet L. Goldstein,[5] Lee Watkins,[5] Fang Xiang,[1] Jouko Saramies,[6] Thomas A. Buchanan,[7] Richard M. Watanabe,[8,9] Timo T. Valle,[10] Leena Kinnunen,[10,11] Gonçalo R. Abecasis,[1] Elizabeth W. Pugh,[5] Kimberly F. Doheny,[5] Richard N. Bergman,[9] Jaakko Tuomilehto,[10,11,12] Francis S. Collins,[3]* Michael Boehnke[1]*

Identifying the genetic variants that increase the risk of type 2 diabetes (T2D) in humans has been a formidable challenge. Adopting a genome-wide association strategy, we genotyped 1161 Finnish T2D cases and 1174 Finnish normal glucose tolerant (NGT) controls with >315,000 single-nucleotide polymorphisms (SNPs) and imputed genotypes for an additional >2 million autosomal SNPs. We carried out association analysis with these SNPs to identify genetic variants that predispose to T2D, compared our T2D association results with the results of two similar studies, and genotyped 80 SNPs in an additional 1215 Finnish T2D cases and 1258 Finnish NGT controls. We identify T2D-associated variants in an intergenic region of chromosome 11p12, contribute to the identification of T2D-associated variants near the genes *IGF2BP2* and *CDKAL1* and the region of *CDKN2A* and *CDKN2B*, and confirm that variants near *TCF7L2*, *SLC30A8*, *HHEX*, *FTO*, *PPARG*, and *KCNJ11* are associated with T2D risk. This brings the number of T2D loci now confidently identified to at least 10.

Type 2 diabetes (T2D) is a disease characterized by insulin resistance and impaired pancreatic beta-cell function that affects >170 million people worldwide (*1*). With first-degree relatives having ~3.5 times as much risk as compared to individuals in the general middle-aged population (*2*), hereditary factors, together with lifestyle and behavioral factors, play an important role in determining T2D risk (*3*). To date, intense efforts to identify genetic risk factors in T2D have met with only limited success. This study, reports from our collaborators (*4*–*6*), and the recently published work of Sladek *et al.* (*7*) describe results of genome-wide association (GWA) studies that further define the genetic architecture of T2D and identify biological pathways involved in T2D pathogenesis.

We genotyped 1161 Finnish T2D cases and 1174 Finnish NGT controls on 317,503 SNPs on the Illumina HumanHap300 BeadChip in stage 1 of a two-stage GWA study of T2D (*8*). These samples are from the Finland–United States Investigation of Non-Insulin–Dependent Diabetes Mellitus Genetics (FUSION) (*9*, *10*) and Finrisk 2002 (*11*) studies (tables S1 and S2A). Among the 317,503 GWA SNPs, 315,635 had ≥10 copies of the less common allele [minor allele frequency (MAF) > 0.002] and passed quality-control criteria (*8*). We tested these 315,635 SNPs for association with T2D using a model that is additive on the log-odds scale (Table 1 and tables S3 and S4) (*8*). We observed a modest excess (41 observed versus 31.6 expected; $P = 0.19$) of SNPs with P values $< 10^{-4}$ (fig. S1). These results argue against the existence of multiple common SNPs with a large impact on T2D disease risk but are consistent with the presence of multiple common SNPs that each confer modest risk. The results also suggest that the matching of cases and controls by birth province, sex, and age (*8*) has been successful; in support of this conclusion, the genomic control (*12*) correction value is 1.026.

Analysis of our Illumina HumanHap300 data allowed us to query much of the known SNP variation in the genome. To increase this proportion, we developed an imputation method (*8*, *13*) that uses genotype data and linkage disequilibrium (LD) information from the HapMap Centre d'Etude du Polymorphisme Humain (Utah residents with ancestry from northern and

[1]Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, MI 48109, USA. [2]Department of Genetics, University of North Carolina, Chapel Hill, NC 27599, USA. [3]Genome Technology Branch, National Human Genome Research Institute, Bethesda, MD 20892, USA. [4]Department of Mathematics and Statistics, Calvin College, Grand Rapids, MI 49546, USA. [5]Center for Inherited Disease Research (CIDR), Institute of Genetic Medicine, Johns Hopkins School of Medicine, Baltimore, MD 21224, USA. [6]Savitaipale Health Center, 54800 Savitaipale, Finland. [7]Division of Endocrinology, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, USA. [8]Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA 90089, USA. [9]Department of Physiology and Biophysics, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, USA. [10]Diabetes Unit, Department of Epidemiology and Health Promotion, National Public Health Institute, 00300 Helsinki, Finland. [11]Department of Public Health, University of Helsinki, 00014 Helsinki, Finland. [12]South Ostrobothnia Central Hospital, 60220 Seinäjoki, Finland.

*To whom correspondence should be addressed. E-mail: boehnke@umich.edu (M.B.); francisc@mail.nih.gov (F.S.C.)