Detailed description of the 1q consortium samples (online appendix)

UK: The 449 UK probands analysed in the present study were taken from the 573 "core" pedigrees of the Warren 2 sibpair consortium collection. As detailed in Wiltshire *et al.* [4], the UK genome wide linkage scan identified a set of 573 pedigrees with the most robust T2D phenotype, the most complete linkage genotype data, and confirmed full-sib family relationships. All probands had diabetes diagnosed between ages 35 and 75 with a strong positive family history of the disease. T2D was confirmed through a combination of clinical, genetic and immunological criteria [4]. Sample number limits for the 1q consortium restricted inclusion to only 484 of 573 eligible UK probands: probands from sibpairs with no 1q sharing (i.e. IBD=0) were preferentially excluded to further maximize power [41]. The present analyses exclude genotypes from 35 probands which are currently quarantined pending repeat genotyping due to excessive QA failures for a subset of SNPs outside the *USF1* region. However, reanalyses including these 35 probands have no impact on the findings.

UK controls (n=450) were derived from the HRC control collection available from the European Centre for Cell Culture (ECACC, CAMR, Salisbury, UK). These are samples derived from UK (European) blood donors collected in centres in central England. Glycaemic status is not known. As with the cases, certain samples genotyped (n=16) are not included here for QC reasons related to SNPs outside the *USF1* region, but their inclusion makes no difference to the findings.

French: The French diabetic cases are composed of 146 probands (offsprings) from the T2D sib-pair families included in the whole genome scan by Vionnet et al. [3] ie. unrelated individuals from different T2D families (= one proband/family); 40 diabetic cases, which are related affected sibs from the family sample; 113 additional diabetic patients with similar characteristics (family history, absence of frank obesity) from the family collection recruited by the Pasteur Institute-CNRS Unit in Lille.

The French control subjects are composed of 13 normoglycemic individuals (fasting glucose < 5.6 mmol/l), which have familial relatedness with some diabetic cases (i.e. members of the sibship family sample) and 288 normoglycemic subjects (fasting glucose < 5.6 mmol/l), which are unrelated husbands or wives from the family collection recruited by the Pasteur Institute-CNRS Unit in Lille. These subjects are all of French Caucasian ancestry.

The *USF1* analyses reported here are restricted to the 259 cases and 288 controls that were entirely unrelated, from the total of 600 described above and genotyped by the 1q consortium.

Utah: The cases in the case-control study were unrelated subjects from multiplex families with diabetes diagnosed before age 65. All families from the previous 1q linkage study [2] were represented in this collection. Generally, the individual chosen was the proband, although where DNA quantity was limiting, we chose the available individual with the next lowest age at diagnosis. Controls were Utah subjects with normal glucose tolerance (or fasting glucose below 5.6mmol/l) with no known first or second degree family history of type 2 diabetes. The intermediate trait studies employed 124 non-diabetic members of the same diabetes pedigrees, but there was no overlap with the case-control set.

Amish: All Amish subjects were participants in the Amish Family Diabetes Study (AFDS). Details of the AFDS design, recruitment, phenotyping, and pedigree structure have been described previously (Hsueh *et al.*, 2000; [7]). Briefly, probands with previously diagnosed type 2 diabetes (onset age 35–65 years) and, where possible, all first- and second-degree relatives of

probands and spouses >18 years of age were recruited and characterized. This study includes additional subjects recruited subsequent to the reports referenced above.

In all, 714 Amish samples have been genotyped for the 1q consortium, including the 150 cases with T2D (with age of diagnosis below 35y) and 361 controls (all with normal glucose tolerance beyond the age of 38) that form the basis of the case-control comparisons reported here. Of the other 203 samples, 159 have IGT and/or IFG: the remaining 44 were included to provide the "links" necessary to recreate the pedigrees used in the original affected-only linkage analysis. The intermediate trait studies featured the 631subjects from the 714 with relevant trait measures (including 103 of the cases, 342 of the NGT subjects, 150 with impaired glucose homeostasis, and 36 of the linking subjects (24 individuals with normal and 12 with unknown glucose tolerance).

Shanghai:

The 80 Shanghai cases were unrelated subjects with early onset T2D (diagnosis between 22 and 44) selected from the families previously described [9]. The 80 controls were selected from a community-based survey: all had normal glucose tolerance state confirmed after age 65, without dyslipidemia (TG<1.7mmol/l, HDL-c>0.91mmol/l) or hypertension, and with no family history of diabetes.

Hong Kong:

The Hong Kong cases were selected from the 64 families involved in the linkage study for T2D that were recruited through the Hong Kong Family Diabetes Study [8]. Cases were unrelated and selected based on the presence of T2D as well as metabolic syndrome (NCEPIII criteria) since the 1q region is also linked to metabolic syndrome in this population [14]. Sixty-four age- and sex-matched normal control subjects were selected from the general population participating in a community-based cardiovascular risk screening program. All control subjects did not have metabolic syndrome (NCEPIII criteria) or a family history of diabetes.

Pima:

The Pima case-control sample included 200 cases with age of diagnosis below 25y (the phenotype with the strongest linkage signal [10]), and 199 controls (non-diabetic after age 45y). One quarter of these samples came from the pedigree samples typed in the relevant linkage study [10]. Some degree of background relatedness between cases and controls is inevitable in the Pimas, but samples were chosen such that none was a first-degree relative of another.

A further 733 Pima samples from the original linkage pedigrees were typed as part of the consortium effort and, together with 99 of the case-control sample, analysed using family-based association methods. These 832 Pima samples comprised 570 T2D subjects (age of diagnosis below 45y); 104 nondiabetic siblings (aged over 45y) and 158 parents (to reconstruct family relationships).

Additional references for appendix

Hsueh WC, Mitchell BD, Aburomia R, Pollin T, Sakul H, Gelder EM, Michelsen BK, Wagner MJ, St Jean PL, Knowler WC, Burns DK, Bell CJ, Shuldiner AR: Diabetes in the Old Order Amish: characterization and heritability analysis of the Amish Family Diabetes Study. *Diabetes Care* 23:595-601,2000

Additional Acknowledgements (Online)

The principal funding for the present study was provided as a supplement to NIDDK award U01- DK58026. We also recognize the following additional sources of support and assistance.

Oxford: we acknowledge Diabetes UK for sample ascertainment and CAMR (Salisbury, UK) for assistance with DNA extraction.
Baltimore: Work in Baltimore was supported by research grants T32-AG00219, R01-DK54261, K24-DK02673, U01-DK58026, K07-CA67960, the University of Maryland General Clinical Research Center, grant M01 RR 16500, the General Clinical Research Centers Program, the National Center for Research Resources, the National Institutes of Health, and the Baltimore Veterans Administration Geriatric Research and Education Clinical Center. We gratefully acknowledge our Amish liaisons and field workers and the extraordinary cooperation and support of the Amish community.
Lille: The work in Lille was supported by the association "200 Familles pour vaincre le Diabète et l'Obésité" and the "Association Française des Diabétiques".
Little Rock: Studies in the Utah and Arkansas populations were supported by grant NIH/NIDDK DK39311 and the Research Service of the Department of Veteran Affairs. Subject ascertainment was supported in part by the American Diabetes Association. Experimental work was performed at the Arkansas General Clinical Research Center and supported by grant M01RR14288 from the National Center for Research Resources (NIH). We thank Horace Spencer for statistical consultation and Subraman Ranganathan for insulin assays.
Hong Kong: In Hong Kong, work was supported by grants from the Hong Kong Research Grants Committee (CUHK 4292/99M; 1/04C), the Chinese University of Hong Kong Strategic Grant Program (SRP9902) and the Hong Kong Innovation and Technology Support Fund (ITS/33/00). We thank Cherry Chiu and Delander Wong for recruitment of study subjects. We thank all nursing and medical staff at the Prince of Wales Hospital Diabetes and Endocrine Centre for their dedication and professionalism.
Shanghai: The research in Shanghai was supported by grants from the Major Project of National Nature Science Foundation of China (39630150), the Shanghai Medical Pioneer Development Project (96-3-004 and 996024), and Shanghai Science Technology Development Foundation (01ZB14047).
Phoenix: We thank the Gila River Indian Community in Arizona, U.S.A. for participation. This research was supported in part by the Intramural Research Program of the NIDDK.