Brief Genetics Report

Common Polymorphisms in the USF1 Gene Are Not Associated With Type 2 Diabetes in French Caucasians

Fernando Gibson,[1] Serge Hercberg,[2] and Philippe Froguel[1,3]

Upstream transcription factor 1 (USF1) is a ubiquitously expressed transcription factor of the basic helix-loop-helix leucine zipper family that has been shown to regulate the expression of a raft of key genes involved in glucose and lipid metabolism. The USF1 gene is located at chromosome 1q22-q23, within the most consistently replicated type 2 diabetes susceptibility locus in the human genome. In this study, we have examined the contribution of eight common USF1 single nucleotide polymorphisms (SNPs) to type 2 diabetes susceptibility in the French Caucasian population. None of the USF1 SNPs genotyped, including two SNPs previously associated with familial combined hyperlipidemia (rs2073658 and rs3737787), showed evidence of association with type 2 diabetes. In addition, USF1 SNPs were not associated with plasma levels of glucose, triglycerides, total cholesterol, or apolipoproteins A1 or B in normoglycemic subjects. A total of four common USF1 haplotypes were identified, accounting for >99% of chromosomes. There was no significant difference in the USF1 haplotype distribution of the case and control subjects. In conclusion, we report here that we were unable to find any evidence to support the hypothesis that genetic variation in the USF1 gene makes a significant contribution to type 2 diabetes susceptibility in the French Caucasian population. ***Diabetes*** **54:3040–3042, 2005**

Upstream transcription factor 1 (USF1) is a ubiquitously expressed transcription factor of the basic helix-loop-helix leucine zipper family. USF1 has been shown to regulate the expression of a raft of key genes involved in glucose and lipid metabolism (including insulin, glucokinase, fatty acid synthase, acetyl-CoA carboxylase, and pancreatic duodenal homeobox-1) by binding to palindromic E-box motifs in their promoter regions (1). The USF1 gene is located at chromosome 1q22-q23, within a susceptibility locus for type 2 diabetes that is notable for being the most consistently replicated locus in genome-wide scans for linkage to type 2 diabetes (2). Recently, Pajukanta et al. (3) reported that USF1 single nucleotide polymorphisms (SNPs) are strongly associated with familial combined hyperlipidemia (FCHL) in the Finnish population and proposed that USF1 is a good candidate for modulating susceptibility to type 2 diabetes and the metabolic syndrome. In this study, we have examined this hypothesis by evaluating the contribution of common USF1 SNPs to type 2 diabetes susceptibility in the French Caucasian population.

RESEARCH DESIGN AND METHODS

American Diabetes Association 2003 criteria (4) for the classification of subjects as diabetic or normoglycemic were applied. The type 2 diabetes case-control study was carried out with a cohort of 744 unrelated type 2 diabetic subjects (age at diagnosis 47 ± 12 years, BMI 27.0 ± 3.5 kg/m^2, men/women 55/45%) and 731 unrelated normoglycemic subjects (age at examination 54 ± 11 years, BMI 24.6 ± 3.9 kg/m^2, men/women 41/59%). The type 2 diabetic case subjects were composed of 372 probands from type 2 diabetic families, recruited by P.F.'s Centre National de la Recherche Scientifique, Institute Pasteur Unit in Lille, and 372 singleton patients with a family history of type 2 diabetes recruited at the Endocrinology-Diabetology Department of the Corbeil-Essonne Hospital. The control subjects included 361 normoglycemic husbands or wives from type 2 diabetic families and 370 normoglycemic subjects from the SUVIMAX (Supplementation en Vitamines et Mineraux Antioxydants) prospective population-based cohort study (5). Subjects were all of French Caucasian ancestry. A number of quantitative clinical phenotypes (including plasma levels of glucose, triglycerides, total cholesterol, and apolipoproteins A1 and B) were assayed in >95% of our normoglycemic cohort. Informed consent was obtained from all subjects, and the study was approved by the local ethics committees.

USF1 SNPs. Pajukanta et al. (3) identified 17 common (minor allele frequency >5%) SNPs in the USF1 gene by sequencing the entire 6.7-kb USF1 genomic sequence together with 2 kb upstream of the 5′ end of the gene in 31 FCHL probands. SNP location, dbSNP rs numbers, and linkage disequilibrium (LD) data are presented in Table 2 and Supplementary Table 3 of that report. SNP pairs stated as being in complete LD ($\Delta^2 = 1$) were deemed redundant.

SNP genotyping. Genotyping was performed with the Sequenom MassARRAY system (7). Every SNP genotyped in our case-control cohort was also genotyped in the subset of 55 French "strict-lean" type 2 diabetic families that produced a linkage signal at 1q (8). The case cohort contains 53 samples that are also present in the strict-lean family set. The genotyping concordance rate of these duplicate samples allowed us to estimate the accuracy of our genotyping.

Statistical analyses. Power calculations were performed with the program of Purcell et al. (9) (available at http://statgen.iop.kcl.ac.uk/gpc/). Assuming a disease prevalence of 0.1, a genotype relative risk of 1.25, and an allele frequency of 0.73 (the frequency of the major allele of rs3737787, previously associated with FCHL [3]), the case-control sample afforded an estimated 83% power at $P < 0.05$. Comparisons of SNP allele and haplotype frequencies in case and control groups were performed using the χ^2 test. Quantitative traits were log transformed and corrected for age, sex, and BMI, as appropriate. Testing SNPs for association with quantitative traits was carried out with the ANOVA test, using a codominant model. Pairwise SNP LD was calculated with the GOLD software package (10) from the haplotype counts output by PHASE (11). The Haplotype Trend Regression program (12) was used to test inferred haplotypes for association with quantitative traits.

From the [1]Department of Genomic Medicine, Imperial College, Hammersmith Campus, London, U.K.; the [2]Institut National de la Santé et de la Recherche Médicale, Unite de Surveillance et d'Epidemiologie Nutritionnelle, InVS/CNAM, Institut Scientifique et Technique de la Nutrition et de l'Alimentation/CNAM, Paris, France; and the [3]Centre National de la Recherche Scientifique, Institut de Biologie de Lille, Institut Pasteur, Lille, France.

Address correspondence and reprint requests to Fernando Gibson, U557 INSERM and Unite de Surveillance et d'Epidemiologie Nutritionnelle, InVS/CNAM, Institut Scientifique et Technique de la Nutrition et de l'Alimentation/CNAM, Paris, France. E-mail: fernando.gibson@imperial.ac.uk.

Received for publication 29 April 2005 and accepted in revised form 28 June 2005.

Additional information for this article can be found in an online appendix at http://diabetes.diabetesjournals.org.

FCHL, familial combined hyperlipidemia; LD, linkage disequilibrium; SNP, single nucleotide polymorphism; USF1, upstream transcription factor 1.

TABLE 1
Common USF1 SNPs genotyped in our type 2 diabetic case-control cohort

dbSNP ID	Gene region	Position*	Alleles	MAF
rs2516837	intron 1	1030	C/T	0.44
rs1556259	intron 1	1108	T/C	0.19
rs2516838	intron 1	1387	G/C	0.29
rs2073653	intron 2	2997	A/G	0.11
rs2774276	intron 5	4041	G/C	0.27
rs2516841	intron 7	4983	G/A	0.31
rs2073658	intron 7	4995	G/A	0.23
rs3737787	3′ UTR	6234	C/T	0.24

*Position relative to the first nucleotide of the National Center for Biotechnology Information RefSeq USF1 mRNA sequence (NM_007122). Note that USF1 is transcribed in the reverse orientation relative to the reference human genome sequence (chromosome 1 nucleotides: 157,828,830–157,822,115). MAFs were taken from Pajukanta et al. (3). The USF1 gene spans 11 exons and 6.7 kb, with a genomic structure as follows: exon 1, nucleotides 1–110; exon 2, 2608–2700 (ATG codon 2693–2695); exon 3, 3086–3135; exon 4, 3298–3413; exon 5, 3751–3852; exon 6, 4122–4317; exon 7, 4557–4644; exon 8, 5096–5154; exon 9, 5298–5392; exon 10, 5638–5766; and exon 11, 5959–6716 (TAA codon 6046–6048). MAF, minor allele frequency.

TABLE 2
USF1 SNP allele frequencies in type 2 diabetic case and normoglycemic control subjects

SNP	*n*	Allele 1	Allele 2	*P*
rs2516837				
T2D	721	C:872 (0.60)	T:570 (0.40)	0.44
NG	681	C:843 (0.62)	T:519 (0.38)	
rs1556259				
T2D	713	T:1229 (0.86)	C:197 (0.14)	0.76
NG	679	T:1165 (0.86)	C:193 (0.14)	
rs2516838				
T2D	710	G:935 (0.66)	C:485 (0.34)	0.93
NG	678	G:895 (0.66)	C:461 (0.34)	
rs2073653				
T2D	723	A:1246 (0.86)	G:200 (0.14)	0.86
NG	683	A:1174 (0.86)	G:192 (0.14)	
rs2774276				
T2D	704	G:1042 (0.74)	C:366 (0.26)	0.29
NG	672	G:1018 (0.76)	C:326 (0.24)	
rs2516841				
T2D	718	G:1073 (0.75)	A:363 (0.25)	0.28
NG	682	G:1043 (0.76)	A:321 (0.24)	
rs2073658				
T2D	705	G:1039 (0.74)	A:371 (0.26)	0.18
NG	654	G:934 (0.71)	A:374 (0.29)	
rs3737787				
T2D	682	C:1009 (0.74)	T:355 (0.26)	0.44
NG	597	C:867 (0.73)	T:327 (0.27)	

Genotype data was generated from a total cohort of 744 type 2 diabetes case and 731 normoglycemic control subjects of French Caucasian ancestry. NG, normoglycemic; T2D, type 2 diabetic.

RESULTS AND DISCUSSION

The USF1 gene consists of 11 exons that span only 6.7 kb of genomic DNA. Pajukanta et al. (3) previously identified 17 common (minor allele frequency >5%) noncoding SNPs in the USF1 gene by sequencing the entire USF1 genomic sequence gene in 31 FCHL probands. From the LD information provided in that report, we identified a set of nine nonredundant SNPs. After removing rs2073657, which failed to successfully genotype using either the Sequenom (7) or Taqman methods (13), the remaining eight SNPs (Table 1) were genotyped in our type 2 diabetic case-control cohort. Since rs2073657 is in complete LD ($\Delta^2 = 1$) with the typed SNP rs2516837 in the Finnish population (3), it is highly probable that the loss of rs2073657 had very little or no effect on the proportion of common variation captured in the French Caucasian population. The genotyping success rate for the eight USF1 SNPs genotyped was 93%, and the concordance rate of duplicate genotypes was >99%. The genotype distribution was in accordance with Hardy-Weinberg equilibrium for all SNPs (data not shown). The pairwise LD values, quantified by the metrics D′ and Δ^2, indicate that there is a high level of LD across the USF1 gene in the French population (Fig. 1), corroborating the Finnish LD data (3). The average D′ and Δ^2 values were 1 and 0.26, respectively. Three SNP pairs, including rs2073658 and rs3737787, previously associated with FCHL (3) exhibited complete LD ($\Delta^2 = 1$).

The results of evaluating each of the USF1 SNPs for association with type 2 diabetes is presented in Table 2. None of the SNPs, including two SNPs previously associated with FCHL (rs2073658 and rs3737787), showed evidence of association with type 2 diabetes. There were no significant differences in SNP allele or haplotype frequencies between men and women in the case-control cohort (data not shown). In addition, none of the USF1 SNPs were associated with BMI, homeostasis model assessment of insulin resistance, or with plasma levels of glucose,

	D′							
	rs2516837	rs1556259	rs2516838	rs2073653	rs2774276	rs2516841	rs2073658	rs3737787
rs2516837		1.000	1.000	1.000	0.995	0.998	0.997	0.990
rs1556259	0.257		1.000	1.000	1.000	1.000	1.000	1.000
rs2516838	0.329	0.084		1.000	0.992	0.996	0.996	0.992
Δ^2 rs2073653	0.257	0.997	0.085		1.000	1.000	1.000	1.000
rs2774276	0.515	0.054	0.168	0.054		1.000	1.000	0.989
rs2516841	0.517	0.054	0.169	0.054	0.998		1.000	0.989
rs2073658	0.233	0.060	0.190	0.060	0.122	0.122		1.000
rs3737787	0.231	0.061	0.190	0.061	0.120	0.120	0.993	

FIG. 1. Pairwise LD measures for common SNPs spanning the USF1 gene. Pairwise LD was calculated with the GOLD software package (10) from the combined haplotype counts of the type 2 diabetic case-control cohort output by PHASE (11). *P* values for allelic associations were all <0.0001.

TABLE 3
USF1 haplotype distribution in type 2 diabetic case and normoglycemic control subjects

Haplotype	T2D (%)	NG (%)	*P*
n	1,428	1,356	
11211111	34.1	33.8	0.87
11111122	26.1	27.6	0.37
21112211	25.5	23.8	0.30
22121111	13.8	14.4	0.65
Rare	0.5	0.7	
Total	100	100	

Haplotype frequencies for the eight common SNPs in Table 2 were estimated with PHASE. Haplotypes with a frequency <5% were defined as rare. *n*, number of chromosomes. The proportion of missing SNP genotypes in the dataset used for haplotype estimation was 2.9%. NG, normoglycemic; T2D, type 2 diabetic.

triglycerides, cholesterol, or apolipoproteins A1 or B in the normoglycemic subjects (online appendix Table 1 [available at http://diabetes.diabetesjournals.org]). These findings are in agreement with those of Putt et al. (14), who also reported no associations between USF1 SNPs and anthropometric or plasma lipid phenotypes in a study of >800 European Caucasian men. No associations with quantitative phenotypes were uncovered by stratifying for sex (data not shown). Finally, and perhaps not surprisingly, none of the USF1 SNPs were associated with the evidence for linkage of type 2 diabetes to 1q21-q24 (data not shown).

The USF1 haplotype frequency distribution of the type 2 diabetic case and control subjects is shown in Table 3. A total of four common haplotypes were identified, accounting for >99% of chromosomes in both the case and control subjects. It is clear that there was no significant difference in the USF1 haplotype distribution of the case and control subjects. In addition, none of the four common USF1 haplotypes were associated with any of the quantitative phenotypes in our normoglycemic subjects (data not shown).

In conclusion, we report here that we were unable to find any evidence to support the hypothesis that genetic variation in the USF1 gene contributes to type 2 diabetes susceptibility in the French Caucasian population. The data presented do not rule out the possibility that distant USF1 regulatory variants lying outside the USF1 gene region surveyed by Pajukanta et al. (3) could make a contribution to type 2 diabetes susceptibility. The case-control sample size afforded >80% power to detect an association with an allele of a frequency equivalent to that of the two SNPs that were previously associated with FCHL (3), assuming a genotype relative risk of 1.25. Nevertheless, a more highly powered study involving thousands of subjects may yet uncover a role for USF1 variants in type 2 diabetes susceptibility.

ACKNOWLEDGMENTS

This work is supported by a Wellcome Trust University Award to F.G. (GR065414MF) and grants from Diabetes U.K. (BDA: RD01/0002308), the Association Française des Diabétologues de Langue Française (ALFEDIAM), the European Union–funded GIFT Grant, the Direction de la Recherche Clinique/Assistance Publique-Hopitaux de Paris, and the Program Hospitalier de Recherche Clinique (AOM 96088).

We thank Cliona Boyle and Amina Bibi for technical support and two anonymous referees for critically appraising draft versions of the manuscript.

REFERENCES

1. Shoulders CC, Naoumova RP: USF1 implicated in the aetiology of familial combined hyperlipidaemia and the metabolic syndrome. *Trends Mol Med* 10:362–365, 2004
2. McCarthy MI: Progress in defining the molecular basis of type 2 diabetes mellitus through susceptibility-gene identification. *Hum Mol Genet* 13 (Spec no. 1):R33–R41, 2004
3. Pajukanta P, Lilja HE, Sinsheimer JS, Cantor RM, Lusis AJ, Gentile M, Duan XJ, Soro-Paavonen A, Naukkarinen J, Saarela J, Laakso M, Ehnholm C, Taskinen MR, Peltonen L: Familial combined hyperlipidemia is associated with upstream transcription factor 1 (USF1). *Nat Genet* 36:371–376, 2004
4. Expert Committee on the Diagnosis and Classification of Diabetes Mellitus: Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care* 26 (Suppl. 1):S5–S20, 2003
5. Hercberg S, Preziosi P, Briancon S, Galan P, Triol I, Malvy D, Roussel AM, Favier A: A primary prevention trial using nutritional doses of antioxidant vitamins and minerals in cardiovascular diseases and cancers in a general population: the SU.VI.MAX study: design, methods, and participant characteristics. SUpplementation en VItamines et Mineraux AntioXydants. *Control Clin Trials* 19:336–351, 1998
6. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC: Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* 28:412–419, 1985
7. Jurinke C, van den Boom D, Cantor CR, Koster H: Automated genotyping using the DNA MassArray technology. *Methods Mol Biol* 187:179–192, 2002
8. Vionnet N, Hani El H, Dupont S, Gallina S, Francke S, Dotte S, De Matos F, Durand E, Lepretre F, Lecoeur C, Gallina P, Zekiri L, Dina C, Froguel P: Genomewide search for type 2 diabetes-susceptibility genes in French whites: evidence for a novel susceptibility locus for early-onset diabetes on chromosome 3q27-qter and independent replication of a type 2-diabetes locus on chromosome 1q21–q24. *Am J Hum Genet* 67:1470–1480, 2000
9. Purcell S, Cherny SS, Sham PC: Genetic power calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* 19:149–150, 2003
10. Abecasis GR, Cookson WO: GOLD: graphical overview of linkage disequilibrium. *Bioinformatics* 16:182–183, 2000
11. Stephens M, Smith NJ, Donnelly P: A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 68:978–989, 2001
12. Zaykin DV, Westfall PH, Young SS, Karnoub MA, Wagner MJ, Ehm MG: Testing association of statistically inferred haplotypes with discrete and continuous traits in samples of unrelated individuals. *Hum Hered* 53:79–91, 2002
13. Livak KJ: Allelic discrimination using fluorogenic probes and the 5′ nuclease assay. *Genet Anal* 14:143–149, 1999
14. Putt W, Palmen J, Nicaud V, Tregouet DA, Tahri-Daizadeh N, Flavell DM, Humphries SE, Talmud PJ: Variation in USF1 shows haplotype effects, gene : gene and gene : environment associations with glucose and lipid parameters in the European Atherosclerosis Research Study II. *Hum Mol Genet* 13:1587–1597, 2004