Upstream transcription factor 1 (USF1) in risk of type 2 diabetes

Meex, Steven J. R.; Ostaptchouk, Jana; van der Kallen, Carla J. H.; van Greevenbroek, Marleen M. J.; Schalkwijk, Casper G.; Feskens, Edith J. M.; Blaak, Ellen E.; Wijmenga, Cisca; Hofker, Marten H.; Stehouwer, Coen D. A.

Published in:
Molecular Genetics and Metabolism

DOI:
10.1016/j.ymgme.2008.03.011

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2008

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Upstream transcription factor 1 (USF1) in risk of type 2 diabetes: Association study in 2000 Dutch Caucasians

Steven J.R. Meex, Jana V. van Vliet-Ostaptchouk, Carla J.H. van der Kallen, Marleen M.J. van Greevenbroek, Casper G. Schalkwijk, Edith J.M. Feskens, Ellen E. Blaak, Marten H. Hofker, Coen D.A. Stehouwer, Tjerk W.A. de Bruin

Abstract

Type 2 diabetes shares substantial genetic and phenotypic overlap with familial combined hyperlipidemia. Upstream stimulatory factor 1 (USF1), a well-established susceptibility gene for familial combined hyperlipidemia, is postulated to be such a shared genetic determinant. We evaluated two established variants in familial combined hyperlipidemia (rs2073658 and rs3737787) for association with type 2 diabetes in two Dutch case-control samples (N=2011). The first case-control sample comprised 501 subjects with type 2 diabetes from the Breda cohort and 920 healthy blood bank donors of Dutch Caucasian origin. The second case-control sample included 211 subjects with type 2 diabetes, and 379 normoglycemic controls. SNP rs2073658 and SNP rs3737787 were in perfect linkage disequilibrium. In the first case-control sample, prevalence of the major allele was higher in patients than in controls (75% versus 71%, OR=1.22, p=0.018). A similar effect-size and -direction was observed in the second case-control sample (76% versus 72%, OR=1.25, p=0.16). A combined analysis strengthened the evidence for association (OR=1.23, p=0.006). Notably, the increased risk for type 2 diabetes could be ascribed to the major allele, and its high frequency translated to a substantial population attributable risk of 14.5%. In conclusion, the major allele of rs2073658 in the USF1 gene is associated with a modestly increased risk to develop type 2 diabetes in Dutch Caucasians, with considerable impact at the population level.

Keywords:
- Genetic association study
- Single nucleotide polymorphism
- Familial combined hyperlipidemia
- Complex genetic diseases
- Type 2 diabetes
- Insulin resistance

The first chromosomal locus for familial combined hyperlipidemia (FCHL) on 1q21–23 in Finnish pedigrees led to the identification of upstream transcription factor 1 (USF1), a transcriptional regulator critically involved in lipid and glucose homeostasis [1,2]. Since this original report, association of USF1 with FCHL was replicated in Mexican Americans and Utah Caucasians [3,4] and subsequent studies in distinct ethnic population samples also linked variants in the USF1 gene to inherited susceptibility for hyperlipidemia, the metabolic syndrome and its component traits [5–7]. Importantly, in a recent prospective study, specific alleles of the USF1 gene proved to modify cardiovascular risk and contribute both to cardiovascular disease and all-cause mortality at the population level [8].

Considerable genetic and phenotypic overlap exists between FCHL and type 2 diabetes and dyslipidemia is a very commonly observed phenomenon in patients with type 2 diabetes. FCHL and type 2 diabetes both develop against a background of insulin resistance and predispose to early cardiovascular disease. The USF1 chromosomal region 1q21–23 has not only been repeatedly attributed to FCHL [1,9–11], but also represents the most consistently replicated locus in genome wide scans for linkage to type 2 diabetes [12–21]. USF1 is therefore an attractive biological and positional candidate gene for type 2 diabetes. Yet, the direct contribution of the USF1 gene on type 2 diabetes susceptibility has been addressed in a more limited fashion and results are less unequivocal than for FCHL and lipid traits. Two studies thus far reported a significant genetic association with type 2 diabetes or its component traits with the single nucleotide polymorphisms (SNPs) (rs2073658 and rs3737787)—or SNPs in tight LD—as identified for...
FCHL and lipid traits [5,6]. Data from the two most recent studies did not provide additional support for a statistically significant effect [22,23]. In order to adequately evaluate the contribution of USF1 risk alleles to type 2 diabetes in distinct populations and provide robust assessment of effect-size, replication studies in large, well-defined cohorts are required.

In the present study we investigated whether variants in the USF1 gene contribute to the inherited susceptibility for developing type 2 diabetes in 2011 subjects of the Dutch population, using two independent case-control samples comprising 712 patients with type 2 diabetes and 1299 healthy controls.

Materials and methods

Subjects

The first case-control sample comprised 501 subjects with type 2 diabetes from the Breda cohort [24] and 920 healthy blood bank donors of Dutch Caucasian origin [25]. Patients were diagnosed according to the WHO criteria (random plasma glucose level >11.1 mmol/l or a fasting plasma glucose level >7.0 mmol/l). The clinical characteristics of the patients (HbA1c, total cholesterol, HDL-cholesterol and triglycerides) were available, as well as the level of obesity (body mass index) in each individual. Clinical characteristics of the first case-control sample are provided in Table 1. The second case-control sample was collected as described previously [26,27]. Briefly, more than 2700 subjects with one or more cardiovascular risk factors, including hypertension, BMI >25 kg/m², a positive family history for type 2 diabetes mellitus, or a history of gestational diabetes, were screened for type 2 diabetes. Exclusion criteria were the use of medication that affects glucose metabolism and non-Caucasian ethnicity. The case-control sample comprised all newly diagnosed subjects with type 2 diabetes (N=211), and a random selection of 379 normoglycemic control subjects. Clinical characteristics of the second case-control sample are provided in Table 2. The Human Investigation Review Committee of the Academic Hospital Maastricht and the Medical Ethics Committee of the University Medical Center Utrecht approved the study protocol and all subjects gave written informed consent.

Genotyping

Previous studies that investigated the genetic contribution of USF1 to FCHL and type 2 diabetes identified two specific SNPs—or variants in tight LD with these SNPs—that were consistently most strongly associated with disease. These variants, rs2073658 and rs3777787 are 1239 bp apart and located in intron 7 and the 3’ untranslated region of the USF1 gene respectively. In previous reports [2–4,22] both variants were in (almost) complete linkage disequilibrium.

SNPs rs2073658 and rs3777787 were genotyped using Taqman assays (Applied Biosystems). Assays were performed according to the manufacturer’s instructions. The DNA samples were processed in 384-well plates. Each plate contained eight negative controls and 16 genotyping controls. Genotypes were analyzed using a TaqMan 7900 HT (Applied Biosystems, Nieuwkerk a/d IJssel, The Netherlands).

Statistical analyses

Power calculations were performed using the genetic power calculator of Purcell et al. [28] (available at http://pngu.mgh.harvard.edu/~purcell/gpc/). Assuming a disease prevalence of 0.1, a genotype relative risk of 1.25 and 1.50, and an allele frequency of 0.75, the combined case-control sample afforded an estimated power of 0.77 at p<0.05. The genotype frequencies were tested for Hardy–Weinberg equilibrium by \( \chi^2 \) analysis. Odds ratios and p-values for case-control analyses were calculated using \( \chi^2 \) tests. Adjustments for age sex and body mass index were done using logistic regression analyses. \( D^2 \) and \( r^2 \) between the USF1 SNPs were calculated using Haploview.

Results

We genotyped rs2073658 in all subjects from both cohorts. Genotype success rate was >98%, and the SNP was in Hardy–Weinberg equilibrium. In addition, we genotyped SNP rs3777787 in a subset of 590 subjects. In accordance with previous studies both SNPs were in perfect linkage disequilibrium (\( D^2 =1, r^2 =1 \)). The results of SNP rs2073658, which has the best credentials as a potential causal variant in USF1 [29], are presented in Table 3.

In the first case-control sample, prevalence of the major allele was higher in patients than in controls (75% versus 71%, odds ratio 1.25, \( p =0.018 \)). In the second case-control sample, a similar frequency difference and concomitant odds ratio were observed: 76% in patients versus 72% in control subjects, odds ratio 1.22. Although insufficient power of the second case-control sample prevented this frequency difference to meet the threshold for statistical significance (\( p =0.16 \)), the effect-size and -direction were identical to the first case-control sample and a combined analysis of both samples further strengthened the evidence for association (odds ratio 1.23, \( p =0.006 \)). Adjustment for age and sex did not substantially

<table>
<thead>
<tr>
<th>Trait</th>
<th>Case-control sample 1 (N=1421)</th>
<th>Control (N=2073)</th>
<th>Odds ratio</th>
<th>95% CI</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N (female/male)</td>
<td>501 (276/230)</td>
<td>920 (354/557)</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at diagnosis (years)</td>
<td>70.7 ± 9.9</td>
<td>47.8 ± 12.7</td>
<td>1.25</td>
<td>0.93–1.61</td>
<td>0.155</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>27.6 ± 4.9</td>
<td>NA</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HbA1c (%)</td>
<td>6.6 ± 2.4</td>
<td>NA</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HDL-cholesterol (mmol/L)</td>
<td>1.1 ± 0.5</td>
<td>NA</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total cholesterol (mmol/L)</td>
<td>4.9 ± 1.8</td>
<td>NA</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Triglyceride (mmol/L)</td>
<td>1.7 ± 1.1</td>
<td>NA</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The data are presented as means ± SD. HbA1c, haemoglobin A1c (glucose bound to haemoglobin); NA, not available.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Case-control sample 2 (N=590)</th>
<th>Control (N=742)</th>
<th>Odds ratio</th>
<th>95% CI</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N (female/male)</td>
<td>141 (63/78)</td>
<td>139 (48/91)</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>74 (35.6)</td>
<td>152 (40.8)</td>
<td>1.05</td>
<td>0.51–2.14</td>
<td>0.896</td>
</tr>
<tr>
<td>BMI</td>
<td>121 (58.2)</td>
<td>193 (51.7)</td>
<td>1.35</td>
<td>0.67–2.71</td>
<td>0.396</td>
</tr>
<tr>
<td>HbA1c (%)</td>
<td>100 (24.0)</td>
<td>206 (27.9)</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HDL-cholesterol (mmol/L)</td>
<td>316 (76.0)</td>
<td>538 (72.1)</td>
<td>1.22</td>
<td>0.93–1.61</td>
<td>0.155</td>
</tr>
<tr>
<td>Total cholesterol (mmol/L)</td>
<td>487 (6.3)</td>
<td>115 (9.0)</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Triglyceride (mmol/L)</td>
<td>248 (35.6)</td>
<td>507 (39.6)</td>
<td>1.17</td>
<td>0.81–1.70</td>
<td>0.400</td>
</tr>
<tr>
<td>HDL-cholesterol (mmol/L)</td>
<td>401 (57.5)</td>
<td>658 (51.4)</td>
<td>1.44</td>
<td>1.02–2.09</td>
<td>0.038</td>
</tr>
<tr>
<td>Total cholesterol (mmol/L)</td>
<td>344 (24.7)</td>
<td>737 (28.8)</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Triglyceride (mmol/L)</td>
<td>1050 (75.3)</td>
<td>1823 (71.2)</td>
<td>1.23</td>
<td>1.06–1.43</td>
<td>0.006</td>
</tr>
</tbody>
</table>

The data are presented as means ± SD. HbA1c, haemoglobin A1c (glucose bound to haemoglobin); NA, not available.
influence the result (adjusted odds ratio 1.20, p = 0.02). Information on BMI was not available for control subjects in sample 1. In sample 2, BMI was significantly higher in patients than control subjects. To explore the possibility that differences in allele frequency between patients and controls are primarily due to differences in body weight, we performed age-, sex- and BMI-adjusted analyses in case-control sample 2. Adjustment for these traits only slightly reduced the effect-size in case-control sample 2 (odds ratio unadjusted: 1.22; odds ratio age-, sex- and BMI-adjusted: 1.15). This suggests that the observed differences in allele frequency are not primarily due to differences in body weight, age or sex, but rather reflect a direct effect on type 2 diabetes.

Importantly, the increased risk for type 2 diabetes susceptibility can be ascribed to the major allele, and its high frequency translates to a considerable population attributable risk of 14.5%. Together, these data indicate that the major allele of rs2073658 is associated with a modest but consistent increased risk for developing type 2 diabetes in Dutch Caucasians, with substantial impact at the population level.

**Discussion**

USF1 is a positional and functional candidate gene for type 2 diabetes. The main finding of this study is that a polymorphism within the USF1-gene with strongest prior odds for association and best credentails for functional involvement (rs2073658) is associated with type 2 diabetes in an analysis with 2011 Dutch Cardiovascular subjects. The observed effect-size of the USF1 risk allele is modest (OR = 1.25), but consistent in the two independent Dutch case-control samples.

USF1 is a ubiquitously expressed transcription factor of the basic helix-loop-helix leucine zipper family and regulates the expression of some 40 genes, several of them involved in lipid and glucose metabolism. USF1 mediates its transcriptional regulation through binding to E-box motifs in the promoter region of target genes with the consensus sequence CACGTG, either as homodimer but in most cases as a heterodimer with the related transcription factor USF2 [30].

USF1 was originally identified as a major susceptibility gene for FCHL, underlying the linkage signal on 1q21–23 in Finnish families with the disease [2]. Two polymorphisms in USF1, rs2073658 and rs3737778, in strong LD with each other, were the most strongly associated variants. Subsequent replication studies in various FCHL samples of distinct ethnicities implicated the same polymorphisms (or variants in tight LD) in the genetic background of FCHL [3,4,31]. Furthermore, differential expression of USF1-regulated genes in adipose tissue from subjects with different allelic variants has been reported [29]. Complementary to these findings, significant associations with lipid traits, parameters of adipose tissue metabolism, and the metabolic syndrome in populations not specifically ascertainment for FCHL, further extended the potential significance of USF1 as a cardiovascular risk determinant [6,7,32–34]. Recent important data from a prospective follow up study demonstrate that USF1 risk alleles confer an approximately two fold increased risk to cardiovascular disease and all-cause mortality in women from two independent Finnish cohorts representative of the general population [8]. These latter data provide important decisive evidence and illustrate how a gene originally identified in high risk families proves to be important also at the population level.

No amino acid changes have been detected in the USF1 gene that could functionally account for the observed associations. However, rs2073658 resides in a 20-bp DNA sequence that binds nuclear proteins and possibly represents a transcriptional regulatory element [29]. In all studies to date rs2073658 and rs3737778 were among the strongest associated variants, suggesting that—irrespective of whether one of these SNP is a true etiological variant—they represent reliable markers for FCHL across populations of different ethnicity.

Triggered by the convincing results of USF1 variants in relation to FCHL and lipid traits, some research groups scrutinized the USF1 gene for its potential contribution to inherited DM2 susceptibility. Four studies thus far investigated the potential association of USF1 SNPs with type 2 diabetes or related traits [5,6,22,23]. Putt et al. were the first to genotype three USF1 SNPs in 800 male subjects and report haplotypic associations (comprising SNPs in tight LD with rs3737787 and rs2073658) with glucose levels during an oral glucose tolerance test [5]. In Hong Kong Chinese subjects three haplotype tagging SNPs, including rs3737787, were genotyped. Rs3737787 was associated with type 2 diabetes in 897 family cases from 179 families (40 of which showed 1q-linkage), but no association was found in a second case-control cohort including 1383 unrelated patients with type 2 diabetes and 454 control subjects [6]. Analyses in 744 French Caucasian patients with type 2 diabetes and 731 normoglycemic controls revealed no significant association with any of 8 genotyped USF1 polymorphisms (strongest effect for rs2073658, p = 0.018) [22]. A large recent study in several populations with evidence of chromosome 1q linkage (N = 3726, 22 genotyped variants) found no significant associations [23].

Although the USF1 risk allele in the present study confers only a modest risk to individuals carrying the risk variant, it may have a substantial impact at the population level. As much as ~75% of subjects carry the risk allele, which translates to a population attributable risk (PAR) of 14%. This means, that if the population were monomorphic for the protective allele, the prevalence of type 2 diabetes would be 14% lower. This resembles the established type 2 diabetes variant Pro12Ala in PPARγ, which confers modest risk for the individual (OR = 1.25), but high risk at the population level (PAR = 25%) [35].

Our present study conclude that the major allele of rs2073658 in the USF1 gene is associated with a modestly increased risk to develop DM2 in Dutch Caucasians, with considerable impact at the population level.

**Acknowledgment**

We thank all patients who participated in this study.

**References**


