

## Genetic variants in the USF1 gene are associated with low-density lipoprotein cholesterol levels and incident type 2 diabetes mellitus in women: results from the MONICA/KORA Augsburg case-cohort study, 1984–2002

C. Holzapfel, J. Baumert, H. Grallert, A. M. Müller, B. Thorand, N. Khuseyinova, C. Herder, Christa Meisinger, H. Hauner, Heinz-Erich Wichmann, W. Koenig, T. Illig, N. Klopp

### Angaben zur Veröffentlichung / Publication details:

Holzapfel, C., J. Baumert, H. Grallert, A. M. Müller, B. Thorand, N. Khuseyinova, C. Herder, et al. 2008. "Genetic variants in the USF1 gene are associated with low-density lipoprotein cholesterol levels and incident type 2 diabetes mellitus in women: results from the MONICA/KORA Augsburg case-cohort study, 1984–2002." *European Journal of Endocrinology* 159 (4): 407–16. <https://doi.org/10.1530/eje-08-0356>.



# Genetic variants in the *USF1* gene are associated with low-density lipoprotein cholesterol levels and incident type 2 diabetes mellitus in women: results from the MONICA/KORA Augsburg case-cohort study, 1984–2002

C Holzapfel<sup>1,2</sup>, J Baumert<sup>1</sup>, H Grallert<sup>1</sup>, A M Müller<sup>1,3</sup>, B Thorand<sup>1</sup>, N Khuseyinova<sup>4</sup>, C Herder<sup>5</sup>, C Meisinger<sup>1</sup>, H Hauner<sup>2</sup>, H E Wichmann<sup>1,3</sup>, W Koenig<sup>4</sup>, T Illig<sup>1</sup> and N Klopp<sup>1</sup>

<sup>1</sup>Institute of Epidemiology, Helmholtz Center Munich, German Research Center for Environmental Health, Ingolstädter Landstraße 1, D-85764 Neuherberg, Germany, <sup>2</sup>Else Kröner-Fresenius-Center for Nutritional Medicine, Technical University of Munich, Munich, Germany, <sup>3</sup>Institute of Medical Information Processing, Biometry and Epidemiology, Chair of Epidemiology, Ludwig-Maximilians-University, Munich, Germany, <sup>4</sup>Department of Internal Medicine II-Cardiology, University of Ulm Medical Center, Ulm, Germany and <sup>5</sup>Institute for Clinical Diabetes Research, German Diabetes Center, Leibniz Institute at Heinrich-Heine-University, Düsseldorf, Germany

(Correspondence should be addressed to N Klopp; Email: klopp@helmholtz-muenchen.de)

## Abstract

**Objective:** Upstream transcription factor 1 (*USF1*) regulates genes of glucose and lipid metabolism. Polymorphisms in the *USF1* gene showed association with familial combined hyperlipidemia and lipid parameters. The aim of our study was to examine the associations between *USF1* polymorphisms and lipid parameters as well as incident type 2 diabetes mellitus (T2DM) in German Caucasians.

**Design:** We genotyped eight polymorphisms in the *USF1* gene in 2067 middle-aged (35–74 years) individuals including 498 incident T2DM cases and 1569 non-cases of the population-based case-cohort study from the MONICA/KORA Augsburg project.

**Methods:** Six polymorphisms and their haplotypes were analyzed using multivariable linear regression and Cox proportional hazards models.

**Results:** Polymorphism rs3737787 was inversely associated with incident T2DM in women with decreased risk for female heterozygotes compared with women homozygous for the major allele (Hazard ratio = 0.57; 95% confidence intervals: 0.38–0.87;  $P=0.008$ ). After correction for multiple testing, significance remained. Polymorphisms rs3813609 and rs1556259 were significantly associated with reduction in low-density lipoprotein (LDL) cholesterol ( $p_{\text{NOM}}=0.001$ ;  $p_{\text{NOM}}=0.00002$ ) in women. Analyses also indicated associations of haplotypes with LDL cholesterol in women, but the association lost statistical significance after correction for multiple testing. Total serum cholesterol (TC) and high-density lipoprotein (HDL) cholesterol were weakly associated ( $P<0.05$ ) with *USF1* polymorphisms in women. No significant associations were found in men.

**Conclusions:** In this large population-based study, statistically significant associations of *USF1* polymorphisms with incident T2DM and LDL cholesterol were found in women, but not in men. Genetic variants in the *USF1* gene showed weak or no associations with TC and HDL cholesterol.

*European Journal of Endocrinology* **159** 407–416

## Introduction

The upstream transcription factor 1 (*USF1*) gene encodes for the transcription factor USF1, which is strongly involved in the expression of genes in glucose and lipid metabolism. The *USF1* gene is localized on chromosome 1q23, consists of 11 exons and extends to 6.73 kb (National Centre for Biotechnology (NCBI), gene ID: 7391). USF1 is a member of the basic helix–loop–helix leucine zipper transcription factors, which includes a basic DNA-binding domain and a helix–loop–helix dimerization domain. The leucine zipper acts as a second dimerization domain. USF1 (43 kDa) interacts as a homodimer or together with

USF2 (44 kDa) as a heterodimer with the E-box motif ‘CANNTG’ (1–3). Variance in the promoter of *GNAS complex locus* (*GNAS*) gene affects the binding of the transcription factor USF1 (4).

Being involved in the control of the glucose-regulated expression of hormone-sensitive lipase, USF1 plays an important role in lipolysis (5) and might influence the regulation of the transcription of fatty acid synthase (6) and acetyl-CoA carboxylase (7). USF1 is also involved in the transcription of apolipoprotein AII (1), CIII (8) and A5 (9), hepatic lipase (10), ATP-binding cassette transporter A1 (11), glucokinase (12), and angiotensinogen (13). Furthermore, USF1 might modulate the

expression of insulin by interacting with the enhancer of the insulin gene (14) and has been reported to affect the expression of the islet-specific glucose-6-phosphatase catalytic subunit-related protein (15).

Taking into account the central role of *USF1* in glucose and lipid metabolism as well as the results of recent association studies, it could be assumed that the *USF1* gene represents an attractive candidate gene, which might contribute to lipid disturbances and type 2 diabetes mellitus (T2DM). *USF1* is associated with familial combined hyperlipidemia (FCHL) (16–18), which is the most common genetic dyslipidemia, and because of its characteristic features such as increased levels of apolipoprotein B, serum triglycerides, and total serum cholesterol (TC), a phenotypic overlap between FCHL, T2DM, and the metabolic syndrome exists (19, 20).

In some studies (18, 21–24), genetic variants within the *USF1* gene showed associations with serum triglycerides, whereas in two US cohort studies of cardiovascular disease risk (25), in French Caucasians (26), and in Korean adults (27), no associations with serum triglycerides were observed.

Also for low-density lipoprotein (LDL) cholesterol and polymorphisms within the *USF1* gene, significant associations were described. In Finnish FCHL families, haplotypes of the single nucleotide polymorphisms (SNPs) rs3737787 and rs2073658 were associated with small LDL peak particle size (21). In the Coronary Artery Risk Development in Young Adults cohort, rs3737787 was associated with lower and rs2516837 with higher LDL cholesterol levels, whereas no associations were found in the Cardiovascular Health Study (25). In Dutch families, rs3737787 was significantly associated with TC in men, in Utah families not rs3737787, but three other SNPs were associated with TC (18, 24). In US Whites suffering from coronary artery disease (CAD), no association with TC was found (18).

The latest result was confirmed by Reiner *et al.* in a US cohort study and by Gibson *et al.* in a French study (25, 26). Also, no associations with high-density lipoprotein (HDL) cholesterol were found (24, 25, 27).

Several studies have reported a lack of association between *USF1* polymorphisms and glucose parameters or diabetes (18, 24, 25), whereas in US Whites and in Swedish men, associations with glucose levels were observed (18, 28). Also in haplotype analyses, associations with peak of glucose were found (29).

Considering the existing controversial results on the associations between *USF1* gene and lipids and T2DM, the aim of the present study was to determine whether genetic variants in the *USF1* gene are associated with the lipid parameters such as TC, LDL, and HDL cholesterol as well as with incident T2DM in a large prospective population-based case-cohort study from southern Germany. Therefore, eight SNPs within the *USF1* gene and their haplotypes were investigated in the case-cohort study of the Monitoring of Trends and

Determinants in Cardiovascular Disease/Cooperative Health Research in the Region of Augsburg (MONICA/KORA) project. Because of sex-specific dissimilarities in the development of T2DM (30–32) and sex-specific results concerning *USF1* polymorphisms and lipid parameters, there was a separate analysis for men and women (18, 21, 23).

## Subjects and methods

### Study design

The MONICA/KORA Augsburg study (1984–2002) is a population-based study of adults with German nationality (33). As part of the international World Health Organization MONICA project, three independent cross-sectional population-based surveys (S1–S3) covering the city of Augsburg and two adjacent counties were conducted in 1984/1985 (S1), 1989/1990 (S2), and 1994/1995 (S3) to estimate the prevalence and distribution of cardiovascular risk factors among men and women aged 25–64 (S1) or 25–74 years (S2 and S3). The surveys included a standardized interview and clinical investigations and were approved by the local ethic authorities including genetic analyses for cardiovascular and metabolic diseases. The total number of participants was 13 427 (6725 men and 6702 women). All subjects were prospectively followed within the framework of the KORA Augsburg project (34).

The incidence of T2DM was assessed using a written follow-up questionnaire sent to all participants of the three baseline surveys in 1997/1998 and 2002/2003. Furthermore, all S1 participants were invited to a follow-up examination in 1987/1988. Cases with self-reported incident diabetes were validated by a questionnaire mailed to the treating physician or medical chart review. Only subjects for whom the treating physician clearly reported a diagnosis of T2DM or for whom a diagnosis of T2DM was mentioned in the medical records or who were taking anti-diabetic medication were classified as T2DM cases for the present analysis. Owing to the low incidence of T2DM in younger subjects, the present study was limited to 10 718 persons (5382 men and 5336 women) aged 35–74 years at baseline. After exclusion of 1187 subjects with missing blood samples and 1595 participants with prevalent T2DM, with incident diabetes other than T2DM (e.g., type 1 or secondary diabetes), with self-reported, but non-validated incident diabetes, without follow-up information or with a follow-up time of less than 1 year, the source population of the present study comprised 7936 subjects (3894 men and 4042 women).

On this source population, we applied a case-cohort study consisting of a random sample (called sub-cohort;  $n=1885$ ) of the source population stratified by sex and survey, and all incident T2DM cases ( $n=555$ ) in the source population diagnosed up to December

31, 2002 as it was described more detailed previously (31, 35–37).

After the exclusion of subjects with missing DNA samples and missing values for lifestyle or metabolic risk factors, the present study including the subcohort and incident T2DM cases was based on 2067 participants (307 men and 191 women with incident T2DM; 835 men and 734 women without incident T2DM). Mean follow-up time ( $\pm$  s.d.) was 10.1 ( $\pm$  4.9) years. The final stratum-specific sample sizes of the subcohort were used together with the stratum-specific sizes of the source population to compute sampling fractions, and the inverse of the sampling fractions yielded survey- and sex-specific sampling weights.

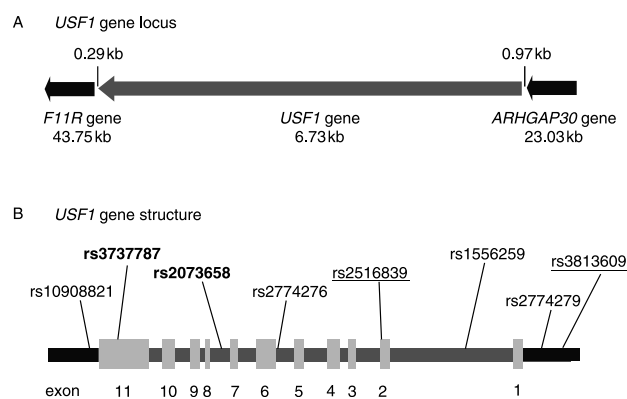
All cross-sectional analyses concerning SNP frequencies, tests for deviation from Hardy–Weinberg equilibrium (HWE) and associations of SNPs and lipid parameters were performed in random sample (subcohort) of the whole study population with available DNA (i.e., without prior to exclusion of subjects without follow-up information, prevalent diabetes). This sample included 1968 subjects.

### Assessment of demographic, lifestyle, and clinical characteristics

Standardized interviews were conducted by trained medical staff (mainly nurses) to assess information concerning sociodemographic variables, smoking habits, leisure time physical activity level, and alcohol consumption. In addition, participants underwent a standardized medical examination and a non-fasting venous blood sample was obtained. Body weight was measured in light clothing by trained investigators to the nearest 0.1 kg, and height to the nearest 0.5 cm. The body mass index (BMI;  $\text{kg}/\text{m}^2$ ) was calculated by body weight in kg divided by squared body height in m. TC and HDL cholesterol were measured by the Boehringer CHOD-PAP (Roche Diagnostics) and by the phosphotungstic acid method (Boehringer) respectively. After precipitation with dextran sulfate (Quantolip LDL, Immuno AG, Vienna, Austria), LDL cholesterol was determined. All survey methods have been described elsewhere in detail (30).

### Selection of polymorphisms

For SNP selection, the NCBI SNP database ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)) and the Haplotype Mapping (HapMap) database ([www.hapmap.org](http://www.hapmap.org)) were used. In addition, a literature search was performed. Eight polymorphisms (Fig. 1; Table 1) were selected for genotyping and, according to high correlation (linkage disequilibrium (LD),  $r^2 \geq 0.8$ ), six tagging SNPs were identified. Polymorphisms with a minor allele frequency (MAF) below 10% were not included in the genotyping process due to power calculation. The selected SNPs completely cover the common genetic variation (SNPs



**Figure 1** Gene locus (A) and gene structure (B) of the *USF1* gene. (A) *USF1* gene and its nearby genes *F11R* and *ARHGAP30* with sizes and distances in kilobases (kb). (B) Eight genotyped SNPs and their location; rs3813609 and rs2516839 (underlined) as well as rs2073658 and rs3737787 (bold) are highly correlated ( $r^2 \geq 0.8$ ).

with MAF greater 10%) in the *USF1* gene and flanking regions. Polymorphisms rs1556259, rs2516839, rs2774276, rs2073658, and rs3737787 are located in introns and exons of the *USF1* gene, rs3813609 and rs2774279 in the 5' region (*ARHGAP30* gene) and rs10908821 in the 3' region (*F11R* gene) of the *USF1* gene.

### Genotyping

DNA samples were genotyped with the MassARRAY system using the iPLEX chemistry (Sequenom, San Diego, CA, USA). The allele-dependent primer extension products were loaded onto one 384 element chip using a nanoliter pipetting system (SpectroCHIP, SpectroPOINT Spotter, Sequenom). The samples were analyzed in a matrix-assisted laser desorption ionization time of flight mass spectrometer (MALDI TOF MS: Bruker Daltonik, Leipzig, Germany). The resulting mass spectra were analyzed automatically for peak identification via the SpectroTYPER RT 3.4 software (Sequenom). For quality reasons, 10% of the spectra were checked independently by two investigators. Negative controls were included in all assays. To control for reproducibility of genotyping data, 12.5% of randomly selected samples were genotyped in duplicate. Each SNP was tested for deviation from HWE by means of a  $\chi^2$ -test or Fisher's exact test depending on allele frequency.

### Statistical analysis

Means or proportions for baseline demographic and clinical characteristics of the study population were computed using the statistical analysis system (SAS) procedures, SURVEYREG or SURVEYFREQ, which estimate s.e.m. appropriate to the sampling scheme. Tests of differences between subjects with and without incident T2DM were based on these procedures. Associations between *USF1* polymorphisms or haplotypes and lipid

**Table 1** Characteristics of the *upstream transcription factor 1 (USF1)* polymorphisms genotyped in MONICA/KORA Augsburg case-cohort study population.

SNP no.	Gene position (dbSNP build 126)	MAF (in our subcohort)	HWE	Success rate	Genotype frequencies		
					0	1	2
rs3813609*	5' flanking region	G=0.38	0.93	98.5	38.5	47.2	14.4
rs2774279	5' flanking region	T=0.33	0.48	96.4	45.3	43.6	11.1
rs1556259	Intron 1	C=0.13	0.48	96.7	76.7	21.6	1.7
rs2516839*	Exon 2	G=0.38	0.86	96.7	38.5	47.3	14.2
rs2774276	Intron 5	C=0.25	0.60	98.2	56.0	37.3	6.6
rs2073658**	Intron 7	A=0.28	0.04	96.5	53.0	38.5	8.5
rs3737787**	Exon 11 (3'-UTR)	T=0.29	0.87	98.4	50.5	41.0	8.5
rs10908821	3' flanking region	G=0.13	0.47	97.2	75.4	22.6	2.0

For every SNP gene position, minor allele frequency (MAF), *P* value of Hardy-Weinberg equilibrium (HWE), genotyping success rate (%), and the genotype frequencies (0, homozygous in major allele; 1, heterozygous; 2, homozygous in minor allele; % weighted for the sampling scheme are given; \*SNP rs3813609 and rs2516839 as well as \*\*SNP rs2073658 and rs3737787 are highly correlated ( $r^2 \geq 0.8$ ).

parameters were assessed separately for men and women by linear regression analysis using SURVEYREG with adjustment for age and survey.

Cox proportional hazards regression was used to assess the association between polymorphisms or haplotypes and incident T2DM. To account for the case-cohort design, s.e.m.s were corrected using the 'sampling weight' approach developed by Barlow (38). Since sex-related differences seem to play a role in the development of diabetes, all analyses were done separately for men and women and carried out for each *USF1* SNP with a multivariable-adjusted model, including survey, age, BMI, systolic blood pressure, ratio of TC:HDL cholesterol, smoking status (never smoker, former smoker, and current smoker), alcohol consumption (men 0, 0.1–39.9,  $\geq 40$  g/day; women 0, 0.1–19.9,  $\geq 20$  g/day), and physical activity (inactive versus active, i.e., regular physical activity of  $\geq 1$  h/week in both summer and winter). *P* values are based on robust variance estimates using the Barlow approach. For each *USF1* SNP, we assumed that a nominal model and tests were carried out based on this model using ANOVA (with two degrees of freedom) with *P* values denoted as  $p_{\text{NOM}}$ . Additionally, a test based on an additive model was performed to assess the association between each *USF1* SNP and incident T2DM (*P* values denoted as  $p_{\text{ADD}}$ ).

As measures for pairwise LD between SNPs, Lewontin's disequilibrium coefficient  $D'$  and the squared correlation coefficient  $r^2$  were calculated. A correlation coefficient of  $r^2 \geq 0.8$  was classified as high correlation.

Haplotype reconstruction was performed within blocks of high  $D'$  between consecutive SNPs ( $D' \geq 0.8$ ) using the expectation-maximization algorithm within the SAS procedure HAPLOTYPE. To avoid large reconstruction errors, haplotype estimation is based only on subjects with complete genotype information. Haplotypes with frequencies below 5% were collected into a separate group of rare haplotypes. All haplotypes with frequency higher than 5% and the rare group were analyzed in a multivariable model with the

most frequent haplotype as reference assuming an additive model.

The global significance level of 5% was corrected for the number of independent tests following the Bonferroni procedure. The number of independent tests was calculated as the number of effective loci ( $n=5$ ) obtained through spectral decomposition of the correlation matrix of all SNPs analyzed in approach to Li & Ji (39).

Corresponding to an overall significance level of  $\alpha=0.05$ , the significance level accounting for associations of lipid markers and SNPs was reduced to 0.003. For analyses of T2DM in the case-cohort,  $\alpha=0.01$  was considered as significance level corrected for multiple testing. For haplotype analyses,  $\alpha=0.003$  (lipids) or 0.01 (T2DM) was statistically significant. Power estimation for the case-cohort study was performed by the method of Cai & Zeng (40). All statistical analyses were performed using the statistical package SAS Version 9.1 (SAS Institute, Cary, NC, USA).

## Results

### Description

Baseline characteristics of the MONICA/KORA Augsburg case-cohort study participants are given in Table 2. Compared with subjects without incident T2DM, cases were older with a mean age of 56.0 years (non-cases: 51.6 years), showed higher BMI values with a mean BMI of 30.0 kg/m<sup>2</sup> (non-cases: 26.7 kg/m<sup>2</sup>), and had a higher waist-to-hip ratio (0.927; non-cases: 0.864). Also concerning blood pressure, hypertension, and myocardial infarction, cases and non-cases differed significantly. Furthermore, significant differences were observed between the two study groups with respect to lipid parameters such as TC, LDL, and HDL cholesterol. Indeed, T2DM cases demonstrated significantly increased concentrations of TC (247.4 mg/dl; non-cases: 235.9 mg/dl) and LDL cholesterol (158.2 mg/dl;

**Table 2** Baseline characteristics of the study population with and without incident type 2 diabetes mellitus (T2DM) during follow-up ( $n=2067$ ).

Variable	T2DM cases	T2DM non-cases	<i>P</i> value <sup>a</sup>
<b>Demographic</b>			
Number (women/men)	191/307	734/835	
Age (years)	56.0 (0.4)	51.6 (0.3)	<0.001
Education < 12 years	81.7	74.8	0.001
<b>Clinical</b>			
Body mass index (kg/m <sup>2</sup> )	30.0 (0.2)	26.7 (0.1)	<0.001
Waist to hip ratio <sup>b</sup>	0.927 (0.005)	0.864 (0.003)	<0.001
Systolic blood pressure (mmHg)	142.2 (0.8)	132.7 (0.5)	<0.001
Diastolic blood pressure (mmHg)	85.3 (0.5)	81.5 (0.3)	<0.001
History of actual hypertension	66.5	39.8	<0.001
History of myocardial infarction	5.0	1.9	0.003
Total cholesterol (mg/dl)	247.4 (2.0)	235.9 (1.1)	<0.001
LDL cholesterol (mg/dl) <sup>c</sup>	158.2 (2.4)	145.8 (1.3)	<0.001
HDL cholesterol (mg/dl)	47.9 (0.6)	57.6 (0.4)	<0.001
<b>Lifestyle</b>			
Smoking status			0.002
Current smoker	26.9	24.2	
Former smoker	33.3	27.2	
Never smoker	39.8	48.6	
Frequency of exercise			<0.001
Inactive	69.1	58.3	
Active <sup>d</sup>	30.9	41.7	
Alcohol consumption <sup>e</sup>			0.112
0 g/day	32.1	28.5	
> 0–39.9/19.9 g/day	39.4	44.6	
≥ 40/20 g/day	28.5	26.9	
Current HRT <sup>f</sup>	8.8	12.5	0.219
Current use of OC <sup>g</sup>	9.1	14.8	0.238
<b>Survey</b>			
S1	36.5	28.9	<0.001
S2	39.4	36.1	
S3	24.1	35.0	

Data are weighted percentages for categorical variables and weighted means (S.E.M.) for normally distributed continuous variables; mean follow-up, 10.1 years ( $\pm 4.9$ ); weights: cases = 1; non-cases = 1/sampling fraction with sampling fraction = subcohort/full cohort without cases for each sex and survey; hormone replacement therapy (HRT); oral contraceptive (OC).

<sup>a</sup>*t*-test for continuous variables and  $\chi^2$ -test for categorical variables.

<sup>b</sup>Only measured in participants of S2 and S3 ( $n=1281$ ).

<sup>c</sup>Only measured in participants of S2 and S3 ( $n=1287$ ).

<sup>d</sup>Regular physical activity of  $\geq 1$  h/week in both summer and winter.

<sup>e</sup>Men: 0 g/day, > 0–39.9 g/day,  $\geq 40$  g/day; women: 0 g/day, > 0–19.9 g/day,  $\geq 20$  g/day.

<sup>f</sup>Only for women aged  $\geq 50$  years with no current use of OC ( $n=538$ ).

<sup>g</sup>Only for women aged < 50 years with no current HRT ( $n=375$ ).

non-cases: 145.8 mg/dl) as well as significantly decreased HDL cholesterol (47.9 mg/dl; non-cases: 57.6 mg/dl) values than non-cases (Table 2). The differences for lifestyle factors such as smoking status, frequency of exercise, alcohol consumption, hormone replacement therapy (HRT), and current use of oral contraceptives are also shown in Table 2.

Eight polymorphisms (rs3813609, rs2774279, rs1556259, rs2516839, rs2774276, rs2073658, rs3737787, and rs10908821) within the *USF1* gene and its flanking regions were successfully genotyped with a mean genotyping success rate of about 97%. Seven SNPs fulfilled the assumption of HWE ( $P > 0.05$ ), whereas rs2073658 revealed a slight deviation from HWE with a *P* value of 0.04. This polymorphism, however, was highly correlated with rs3737787 ( $r^2 = 0.99$ ;  $D' = 1.0$ ). The MAF and genotype frequencies as well as the characteristics of all SNPs are shown in Table 1. For statistical analyses, one representative polymorphism per LD block ( $r^2 \geq 0.8$ ) was selected. The chosen SNPs were those with the best genotyping success rates per LD block. Haplotype reconstruction following the reading direction of the gene revealed five haplotypes with a frequency of more than 5%: GTTGCC (33.1%), GCTGTC (28.6%), CCTCCG (13.0%), CCCGCC (12.9%), and CCTCCC (11.8%; data not shown).

### Association of *USF1* SNPs and haplotypes with incident T2DM

In the present study, a significant association was detected only between one analyzed polymorphism within the *USF1* gene and incident T2DM. SNP rs3737787 was inversely associated with incident T2DM in heterozygous women compared with women who were homozygous carriers of the major allele (Hazard ratio (HR) = 0.57; 95% confidence intervals (CI): 0.38–0.87;  $P = 0.008$ ) after correction for multiple testing (Table 3). In men, no significant associations were found between *USF1* SNPs and incident T2DM.

We performed two sensitivity analyses to confirm our results. Additional adjustment for lipid lowering or anti-hypertensive medication and intake of statins as well as intake of oral contraceptive and HRT did not confound substantially the associations between SNPs and incident T2DM. Following this adjustment, the effect of the heterozygote genotype of rs3737787 on the risk for incident T2DM was 0.61 (95% CI: 0.40–0.92;  $P = 0.020$ ) compared with the homozygote genotype with the major allele. However, this association remained not significant after correction for multiple testing. Moreover, the adjustment for glucose, HbA1c, and triglycerides (only available in S3, 262 men with 67 cases, and 249 women with 52 cases) did not change the results in a significant way. The effect of the heterozygote genotype of rs3737787 on the risk for incident T2DM was 0.30 (95% CI: 0.10–0.92) with *P* value 0.035 compared with the genotype homozygote for the major allele. This association lost statistical significance after correction for multiple testing.

Haplotype analysis did not show any significant association with incident T2DM in men and women (data not shown). Assuming a HR of 1.5, the power to detect a significant difference was over 90% in men and almost 80% in women when the frequency of the genotype with at least one minor allele is assumed to be

**Table 3** Sex-specific associations of polymorphisms within the *upstream transcription factor 1 (USF1)* gene and incident type 2 diabetes mellitus (T2DM).

SNP no.	N	Heterozygous		Homozygous	
		HR (95% CI)	P value	HR (95% CI)	P value
Incident T2DM (men)					
rs3813609	1123	1.04 (0.75–1.44)	0.812	0.86 (0.53–1.40)	0.551
rs2774279	1106	1.09 (0.79–1.50)	0.609	1.13 (0.70–1.83)	0.625
rs1556259	1105	1.06 (0.75–1.50)	0.725	0.34 (0.07–1.65)	0.182
rs2774276	1121	1.09 (0.79–1.50)	0.611	0.91 (0.47–1.76)	0.773
rs3737787	1125	1.01 (0.74–1.37)	0.954	0.84 (0.47–1.50)	0.549
rs10908821	1115	1.25 (0.87–1.78)	0.225	1.35 (0.48–3.78)	0.573
Incident T2DM (women)					
rs3813609	910	1.31 (0.88–1.94)	0.181	0.93 (0.47–1.86)	0.843
rs2774279	890	1.46 (0.97–2.19)	0.071	1.14 (0.63–2.05)	0.670
rs1556259	891	1.04 (0.65–1.66)	0.871	1.31 (0.29–5.86)	0.728
rs2774276	908	0.95 (0.64–1.43)	0.821	1.15 (0.48–2.76)	0.753
rs3737787	905	0.57 (0.38–0.87)	0.008	0.85 (0.43–1.70)	0.653
rs10908821	895	1.13 (0.71–1.81)	0.610	0.60 (0.07–5.27)	0.645

Associations are calculated in reference to homozygous carriers of the major allele. Hazard ratios (HRs) with 95% confidence intervals (95% CI) and *P* values are given; heterozygous and homozygous minor allele carriers were separately compared with the reference group (homozygous for the major allele); all models were adjusted for survey, age, BMI, systolic blood pressure, ratio of TC:HDL cholesterol, smoking status, alcohol consumption, and physical activity; after multiple testing polymorphism rs3737787 showed significant association ( $P=0.008$ ) with incident T2DM in women.

over 40%. In the case of a frequency of 20%, the power was estimated to be 82% in men and 64% in women.

### Association of USF1 SNPs and haplotypes with TC, HDL, and LDL cholesterol

For two lipid markers, TC and HDL cholesterol, significant associations with *USF1* SNPs were observed in women. Three polymorphisms were associated with decreased TC (rs3813609:  $p_{\text{NOM}}=0.020$ ; rs2774276:  $p_{\text{NOM}}=0.010$ ; rs10908821:  $p_{\text{NOM}}=0.011$ ) and one polymorphism with elevated HDL cholesterol (rs1556259:  $p_{\text{NOM}}=0.016$ ). None of the analyzed genetic variants was significantly associated with these lipid parameters in men. After correction for multiple testing, none of the associations with TC and HDL cholesterol in women remained significant (Table 4). Also in haplotype analysis, no significant association was found with TC and HDL cholesterol (data not shown).

Several *USF1* polymorphisms showed statistically significant associations with decreased LDL cholesterol concentrations in women, but not in men (Table 4). For instance, the association between the intronic polymorphism rs1556259 and reduction in LDL cholesterol concentrations in the nominal model remained significant even after accounting for multiple testing ( $p_{\text{NOM}}=0.00002$ ;  $p(\text{TC})=0.085$ ;  $p(\text{CC})<0.0001$ ). Monotone changes in LDL cholesterol levels with increasing number of copies of the minor allele indicate a possible additive effect. Assuming an additive model, the association of rs1556259 remained significant ( $p_{\text{ADD}}=0.021$ ).

Furthermore, rs3813609 located in the 5' flanking region, and rs2774276 within intron 5 of the *USF1* gene were significantly associated with decreased LDL cholesterol ( $p_{\text{NOM}}=0.001$  and  $p_{\text{NOM}}=0.006$  respectively; Table 4). The corresponding *P* values were 0.224

and 0.0002 for GC and CC genotypes of polymorphism rs3813609 respectively and 0.691 and 0.002 for GC and CC genotypes of polymorphism rs2774276 respectively. Similar to the polymorphism rs1556259, the association between rs3813609 or rs2774276 and decreased LDL cholesterol were also significant in the additive model ( $p_{\text{ADD}}=0.001$ ;  $p_{\text{ADD}}=0.045$ ). All significant associations with LDL cholesterol were detected only in women. After correction for multiple testing, the only associations between rs1556259 and rs3813609 and LDL cholesterol concentrations remained statistically significant ( $P\leq 0.003$ ) in the nominal model and rs3813609 also in the additive model (Table 4).

Additional adjustment for BMI, smoking status, alcohol consumption, physical activity, systolic blood pressure, lipid lowering medication, anti-hypertensive treatment, and intake of statins as well as intake of oral contraceptive and HRT in women revealed similar associations between SNPs and mean lipid markers in comparison with models adjusted only for age and survey.

Haplotype analyses revealed that haplotypes CCCGCC and CCTCCC were borderline significantly associated with reduction in LDL cholesterol concentrations ( $P=0.017$  and  $P=0.040$  respectively), compared with the most frequent haplotype GTTGCC in women, but not in men (data not shown). Both of these haplotypes are observed with frequency of more than 10%. The association, however, was strongly attenuated and was no longer statistically significant after correction for multiple testing.

## Discussion

In this case-cohort study, none of the analyzed *USF1* polymorphisms showed consistent association with

**Table 4** Sex-specific association of polymorphisms within the *upstream transcription factor 1 (USF1)* gene and the lipid parameters total cholesterol (TC), high-density lipoprotein (HDL), and low-density lipoprotein (LDL) cholesterol.

SNP no.	Genotype	N (men)	Mean $\pm$ s.e.m.	P value nominal	P value additive	N (women)	Mean $\pm$ s.e.m.	P value nominal	P value additive
Total cholesterol									
rs3813609	GG	393	235.4 $\pm$ 2.3	0.507	0.907	359	243.6 $\pm$ 2.3	0.020	0.029
	GC	490	238.3 $\pm$ 2.1			419	242.0 $\pm$ 2.1		
	CC	170	234.5 $\pm$ 3.2			107	232.1 $\pm$ 3.5		
rs2774279	CC	471	235.0 $\pm$ 2.1	0.394	0.567	384	240.4 $\pm$ 2.2	0.414	0.263
	CT	448	238.8 $\pm$ 2.1			379	241.0 $\pm$ 2.1		
	TT	111	234.7 $\pm$ 4.5			104	247.2 $\pm$ 4.7		
rs1556259	TT	777	236.7 $\pm$ 1.6	0.952	0.977	671	242.3 $\pm$ 1.7	0.345	0.146
	TC	231	236.3 $\pm$ 3.0			189	237.5 $\pm$ 3.1		
	CC	25	238.8 $\pm$ 7.6			10	235.8 $\pm$ 9.7		
rs2774276	GG	586	235.7 $\pm$ 1.9	0.182	0.882	515	242.0 $\pm$ 2.0	0.010	0.155
	GC	391	239.6 $\pm$ 2.2			319	242.7 $\pm$ 2.4		
	CC	73	230.5 $\pm$ 5.1			49	227.2 $\pm$ 4.7		
rs3737787	CC	551	237.4 $\pm$ 1.9	0.376	0.297	427	240.2 $\pm$ 2.0	0.827	0.578
	CT	418	237.1 $\pm$ 2.3			376	242.0 $\pm$ 2.2		
	TT	86	229.9 $\pm$ 5.0			78	242.2 $\pm$ 5.4		
rs10908821	CC	780	235.7 $\pm$ 1.6	0.353	0.412	670	241.5 $\pm$ 1.7	0.011	0.414
	CG	237	240.5 $\pm$ 3.1			189	242.0 $\pm$ 3.1		
	GG	23	231.8 $\pm$ 11.0			13	218.6 $\pm$ 7.6		
HDL cholesterol									
rs3813609	GG	393	48.4 $\pm$ 0.7	0.122	0.097	359	63.0 $\pm$ 0.9	0.427	0.397
	GC	490	50.4 $\pm$ 0.7			419	62.9 $\pm$ 0.8		
	CC	170	50.1 $\pm$ 1.1			107	65.0 $\pm$ 1.5		
rs2774279	CC	471	49.6 $\pm$ 0.6	0.956	0.887	384	64.2 $\pm$ 0.9	0.068	0.599
	CT	448	49.7 $\pm$ 0.7			379	61.8 $\pm$ 0.8		
	TT	111	49.2 $\pm$ 1.5			104	65.0 $\pm$ 1.7		
rs1556259	TT	777	49.3 $\pm$ 0.5	0.391	0.355	671	63.2 $\pm$ 0.6	0.016	0.631
	TC	231	50.9 $\pm$ 1.1			189	62.7 $\pm$ 1.2		
	CC	25	48.7 $\pm$ 3.4			10	73.9 $\pm$ 3.8		
rs2774276	GG	586	48.9 $\pm$ 0.6	0.090	0.395	515	63.0 $\pm$ 0.7	0.836	0.594
	GC	391	50.8 $\pm$ 0.8			319	63.4 $\pm$ 0.9		
	CC	73	48.0 $\pm$ 1.7			49	64.3 $\pm$ 2.0		
rs3737787	CC	551	50.2 $\pm$ 0.7	0.306	0.128	427	63.9 $\pm$ 0.8	0.168	0.503
	CT	418	49.2 $\pm$ 0.7			376	61.9 $\pm$ 0.9		
	TT	86	48.1 $\pm$ 1.3			78	64.7 $\pm$ 2.0		
rs10908821	CC	780	49.3 $\pm$ 0.5	0.074	0.963	670	63.3 $\pm$ 0.7	0.892	0.836
	CG	237	50.8 $\pm$ 0.9			189	63.3 $\pm$ 1.1		
	GG	23	43.8 $\pm$ 3.3			13	61.2 $\pm$ 4.3		
LDL cholesterol									
rs3813609	GG	239	150.0 $\pm$ 2.8	0.588	0.325	229	154.1 $\pm$ 2.9	0.001	0.001
	GC	333	148.4 $\pm$ 2.3			280	149.5 $\pm$ 2.4		
	CC	110	145.3 $\pm$ 3.6			81	136.6 $\pm$ 3.7		
rs2774279	CC	310	147.0 $\pm$ 2.3	0.621	0.384	273	147.3 $\pm$ 2.5	0.263	0.115
	CT	297	150.0 $\pm$ 2.3			247	149.7 $\pm$ 2.5		
	TT	70	150.6 $\pm$ 5.7			68	157.5 $\pm$ 5.8		
rs1556259	TT	507	149.0 $\pm$ 1.8	0.430	0.437	463	151.0 $\pm$ 2.0	0.00002	0.021
	TC	156	148.2 $\pm$ 3.5			121	144.1 $\pm$ 3.5		
	CC	17	137.5 $\pm$ 8.6			5	121.3 $\pm$ 6.2		
rs2774276	GG	365	149.5 $\pm$ 2.3	0.864	0.671	318	151.3 $\pm$ 2.4	0.006	0.045
	GC	265	147.7 $\pm$ 2.4			230	149.8 $\pm$ 2.7		
	CC	48	148.5 $\pm$ 5.6			41	134.7 $\pm$ 4.7		
rs3737787	CC	367	148.1 $\pm$ 2.1	0.919	0.905	285	147.5 $\pm$ 2.5	0.530	0.266
	CT	262	149.3 $\pm$ 2.5			255	150.7 $\pm$ 2.5		
	TT	53	147.4 $\pm$ 6.2			49	153.5 $\pm$ 6.7		
rs10908821	CC	503	149.0 $\pm$ 1.9	0.951	0.824	436	149.9 $\pm$ 2.0	0.084	0.322
	CG	162	147.8 $\pm$ 3.2			143	148.9 $\pm$ 3.6		
	GG	16	149.3 $\pm$ 9.9			12	132.5 $\pm$ 7.6		

Means and s.e.m.s are given; P values are shown for the nominal and the additive model; all models were adjusted for age and survey; after correction for multiple testing the associations with decreased LDL cholesterol remain statistically significant ( $P < 0.003$ ) in women; LDL cholesterol was measured only in S2 and S3.

incident T2DM, with the only exception for female heterozygotes of rs3737787, who demonstrated a ~30% decreased risk for the development of T2DM during a 10-year follow-up period. Furthermore, haplotypes which were inferred from these particular SNPs, reflected the results of single SNP-outcome analyses, revealing no associations with incident T2DM. Moreover, the polymorphisms rs1556259 and rs3813609 showed significant associations in the nominal model with decreased LDL cholesterol in women after correction for multiple testing. One of the associated SNPs (rs3813609) causes an amino acid exchange (Leu>Val). In haplotype analyses, a weak association with LDL cholesterol was found for two haplotype variants with frequency of more than 10%. No relation was seen in men and none of the analyzed SNPs and haplotypes was significantly associated with TC and HDL cholesterol in our population-based case-cohort study after accounting for multiple testing. The sex-specific differences, observed in the present analyses, are in accordance with several reports from the literature demonstrating sex-specific results for genetic variants within the *USF1* gene and lipid markers (18, 21, 23), whereas in other studies, only men were analyzed (28, 41). Also for the association between *USF1* and cardiovascular disease and mortality, sex-specific results were found (42). Lee *et al.* found an association between BMI and *USF1* polymorphisms (27) only in women. Mechanisms for sex-specific differences are unknown, although it is known that hormonal differences could play a role.

In this study, a significantly inverse association between one *USF1* polymorphism (rs3737787) and incident T2DM was found only in female heterozygotes. This finding is consistent with a Chinese family study, in which the same polymorphism (rs3737787) was associated with T2DM (43). In female US Whites with CAD, SNP rs3737787 was associated with higher glucose levels (18), while in our study, this *USF1* SNP was associated with decreased risk of incident T2DM. However, in a multiethnic study, there was no evidence that genetic variants in the *USF1* gene may contribute to T2DM susceptibility (24). In contrast to our study, in this report, prevalent T2DM was analyzed. Also in French Caucasians (26) and in a Chinese population study (43), none of the *USF1* polymorphisms was associated with T2DM.

Our study supports the previously published results on the association between *USF1* polymorphisms and LDL cholesterol. In Dutch FCHL families, suggestive associations for the *USF1* polymorphisms rs3737787 and rs2073658 as well as haplotypes were found for small-dense LDL cholesterol (44). A study in Utah families also showed significant associations between the most common *USF1* polymorphism rs3737787 and LDL cholesterol, whereas the association was more evident when only males were analyzed (23). In European-Americans, the minor T allele of SNP

rs3737787 was associated with lower LDL cholesterol, but there was no evidence that there are sex-specific differences (25). In a Swedish male cohort, the minor allele of rs2073658 highly correlated with rs3737787 was associated with higher LDL cholesterol levels (28). By contrast, a population-based Korean study and the EARSII study found no association between LDL cholesterol and genetic variants within the *USF1* gene, but there was indication of interaction between genetic variants in the *USF1* gene and fasting LDL cholesterol (27, 29).

No significant associations between various *USF1* polymorphisms or haplotypes, inferred from these SNPs and TC and HDL cholesterol concentrations, were found in our study, which again is in agreement with most published studies so far (18, 24, 26, 27, 29). In Dutch FCHL families, genetic variants within the *USF1* gene were associated with TC (18, 44). In addition to TC, HDL, and LDL cholesterol, serum triglycerides were also found to be associated with genetic variants within the *USF1* gene in several studies (18, 21–24). However, because of the non-fasting status of our participants, we were not able to investigate serum triglycerides. *In vitro* tests showed that polymorphisms in the *USF1* gene are associated with an increase in lipolytic activity (45, 46) and have a functional role (20). Both association studies and *in vitro* tests strengthen the associations between the *USF1* gene and the lipid parameters.

In contrast to Coon *et al.* we found significant results for LDL cholesterol in women but not in men (23). Although it is well known that men more often suffer from dyslipidemia than women and associations between polymorphisms within the *USF1* gene and serum triglycerides in men were reported, our results in women give evidence that also in female genetic variants within the *USF1* gene might play a role for lipid metabolism.

Our study is the first population-based case-cohort study that analyzed the association between genetic variants in the *USF1* gene and TC, HDL, and LDL cholesterol in German Caucasians. The prospective study design with incident T2DM cases, the long follow-up period and the large sample size represent the strengths of this study. The systematic genetic dissection of the *USF1* gene and the complete coverage of the whole gene locus confirm our conclusions concerning the significant associations between the gene and LDL cholesterol. Nevertheless, the number of subjects homozygous for the minor allele was very small in the case of rs1556259 and the power to detect significant associations between SNPs and incident T2DM might be not sufficient in women due to low case numbers especially for low genotype frequencies. Furthermore, the incidence of T2DM might be underestimated because of the self-reported diabetes and the missing glucose tolerance test.

In conclusion, there is some evidence that genetic variants within the *USF1* gene are associated with LDL

cholesterol and incident T2DM in German Caucasians. The significant results only in women suggest that there are sex-specific differences. Further population-based studies are needed in order to replicate our results and to produce more consistent data.

## Declaration of interest

The authors declare that there is no conflict of interest that would prejudice the impartiality of this work.

## Funding

The MONICA/KORA Augsburg cohort study was financed by the Helmholtz Center Munich, German Research Center for Environmental Health and supported by grants from the Federal Ministry of Education and Research. The present study was funded by the German Research Foundation (DFG, TH-784/2-1) and the German Federal Ministry of Education, Science, Research and Technology (National Genome Research Net-2) and was supported by the Munich Center of Health Sciences (McHealth) as part of LMUinnovativ.

## Acknowledgements

We thank all members of the Helmholtz Center Munich, Institute of Epidemiology, who were involved in the planning and conduct of the MONICA/KORA Augsburg studies, and the KORA Augsburg team. Furthermore, we are grateful to Prof. Lloyd Chambless (School of Public Health, University of North Carolina at Chapel Hill, NC, USA) for statistical advice concerning the analysis of the case-cohort dataset and to Melanie Kolz (Helmholtz Center Munich) for advice regarding the genotyping. Finally, we are indebted to all study participants.

## References

- Ribeiro A, Pastier D, Kardassis D, Chambaz J & Cardot P. Cooperative binding of upstream stimulatory factor and hepatic nuclear factor 4 drives the transcription of the human apolipoprotein A-II gene. *Journal of Biological Chemistry* 1999 **274** 1216–1225.
- Corre S & Galibert MD. USF as a key regulatory element of gene expression. *Medical Science* 2006 **22** 62–67.
- Littlewood TD & Evan GI. Transcription factors 2: helix–loop–helix. *Protein Profile* 1995 **2** 621–702.
- Frey UH, Hauner H, Jöckel KH, Manthey I, Brockmeyer N & Siffert W. A novel promoter polymorphism in the human gene GNAS affects binding of transcription factor upstream stimulatory factor 1, Galphas protein expression and body weight regulation. *Pharmacogenetics and Genomics* 2008 **18** 141–151.
- Smih F, Rouet P, Lucas S, Mairal A, Sengenès C, Lafontan M, Vaulont S, Casado M & Langin D. Transcriptional regulation of adipocyte hormone-sensitive lipase by glucose. *Diabetes* 2002 **51** 293–300.
- Wang D & Sul HS. Upstream stimulatory factor binding to the E-box at -65 is required for insulin regulation of the fatty acid synthase promoter. *Journal of Biological Chemistry* 1997 **272** 26367–26374.
- Travers MT, Vallance AJ, Gourlay HT, Gill CA, Klein I, Bottema CB & Barber MC. Promoter I of the ovine acetyl-CoA carboxylase- $\alpha$  gene: an E-box motif at -114 in the proximal promoter binds upstream stimulatory factor (USF)-1 and USF-2 and acts as an insulin-response sequence in differentiating adipocytes. *Biochemical Journal* 2001 **359** 273–284.
- Pastier D, Lacorte JM, Chambaz J, Cardot P & Ribeiro A. Two initiator-like elements are required for the combined activation of the human apolipoprotein C-III promoter by upstream stimulatory factor and hepatic nuclear factor-4. *Journal of Biological Chemistry* 2002 **277** 15199–15206.
- Nowak M, Helleboid-Chapman A, Jakel H, Martin G, Duran-Sandoval D, Staels B, Rubin EM, Pennacchio LA, Taskinen MR, Fruchart-Najib J & Fruchart JC. Insulin-mediated down-regulation of apolipoprotein A5 gene expression through the phosphatidylinositol 3-kinase pathway: role of upstream stimulatory factor. *Molecular and Cellular Biology* 2005 **25** 1537–1548.
- Botma G, Verhoeven AJ & Jansen H. Hepatic lipase promoter activity is reduced by the C-480T and G-216A substitutions present in the common LIPC gene variant, and is increased by upstream stimulatory factor. *Atherosclerosis* 2001 **154** 625–632.
- Yang XP, Freeman LA, Knapper CL, Amar MJ, Remaley A, Brewer HB Jr & Santamarina-Fojo S. The E-box motif in the proximal ABCA1 promoter mediates transcriptional repression of the ABCA1 gene. *Journal of Lipid Research* 2002 **43** 297–306.
- Iynedjian PB. Identification of upstream stimulatory factor as transcriptional activator of the liver promoter of the glucokinase gene. *Biochemical Journal* 1998 **333** 705–712.
- Yanai K, Saito T, Hirota K, Kobayashi H, Murakami K & Fukamizu A. Molecular variation of the human angiotensinogen core promoter element located between the TATA box and transcription initiation site affects its transcriptional activity. *Journal of Biological Chemistry* 1997 **272** 30558–30562.
- Read ML, Clark AR & Docherty K. The helix–loop–helix transcription factor USF (upstream stimulatory factor) binds to a regulatory sequence of the human insulin gene enhancer. *Biochemical Journal* 1993 **295** 233–237.
- Martin CC, Svitek CA, Oeser JK, Henderson E, Stein R & O'Brien RM. Upstream stimulatory factor (USF) and neurogenic differentiation/beta-cell E box transactivator 2 (NeuroD/BETA2) contribute to islet-specific glucose-6-phosphatase catalytic-subunit-related protein (IGRP) gene expression. *Biochemical Journal* 2003 **371** 675–686.
- Pajukanta P, Nuotio I, Terwilliger JD, Porkka KV, Ylitalo K, Pihlajamäki J, Suomalainen AJ, Syvänen AC, Lehtimäki T, Viikari JS, Laakso M, Taskinen MR, Ehnholm C & Peltonen L. Linkage of familial combined hyperlipidaemia to chromosome 1q21-q23. *Nature Genetics* 1998 **18** 369–373.
- Coon H, Myers RH, Borecki IB, Arnett DK, Hunt SC, Province MA, Djousse L & Leppert MF. Replication of linkage of familial combined hyperlipidemia to chromosome 1q with additional heterogeneous effect of apolipoprotein A-I/C-III/A-IV locus. The NHLBI Family Heart Study. *Arteriosclerosis, Thrombosis, and Vascular Biology* 2000 **20** 2275–2280.
- Lee JC, Weissglas-Volkov D, Kyttala M, Sinsheimer JS, Jokiaho A, de Bruin TW, Lüsüs AJ, Brennan ML, van Greevenbroek MM, van der Kallen CJ, Hazen SL & Pajukanta P. USF1 contributes to high serum lipid levels in Dutch FCHL families and US Whites with coronary artery disease. *Arteriosclerosis, Thrombosis, and Vascular Biology* 2007 **27** 2222–2227.
- De Graaf J, Van der Vleuten G & Stalenhoef AF. Diagnostic criteria in relation to the pathogenesis of familial combined hyperlipidemia. *Seminars in Vascular Medicine* 2004 **4** 229–240.
- Naukkarinen J, Gentile M, Soro-Paavonen A, Saarela J, Koistinen HA, Pajukanta P, Taskinen MR & Peltonen L. USF1 and dyslipidemias: converging evidence for a functional intronic variant. *Human Molecular Genetics* 2005 **14** 2595–2605.
- Pajukanta P, Lilja HE, Sinsheimer JS, Cantor RM, Lüsüs 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). *Nature Genetics* 2004 **36** 371–376.
- Huertas-Vazquez A, Aquilar-Salinas C, Lüsüs AJ, Cantor RM, Canizales-Quinteros S, Lee JC, Mariana-Nunez L, Riba-Ramirez RM, Jokiaho A, Tusie-Luna T & Pajukanta P. Familial combined hyperlipidemia in Mexicans: association with upstream transcription factor 1 and linkage on chromosome 16q24.1. *Arteriosclerosis, Thrombosis, and Vascular Biology* 2005 **25** 1985–1991.

- 23 Coon H, Xin Y, Hopkins PN, Cawthon RM, Hasstedt SJ & Hunt SC. Upstream stimulatory factor 1 associated with familial combined hyperlipidemia, LDL cholesterol, and triglycerides. *Human Genetics* 2005 **117** 444–451.
- 24 Zeggini E, Damcott CM, Hanson RL, Karim MA, Rayner NW, Groves CJ, Baier LJ, Hale TC, Hattersley AT, Hitman GA, Hunt SE, Knowler WC, Mitchell BD, Ng MC, O'Connell JR, Pollin TI, Vaxillaire M, Walker M, Wang X, Whittaker P, Xiang K, Jia W, Chan JC, Froguel P, Deloukas P, Shuldiner AR, Elbein SC & McCarthy MI. International Type 2 Diabetes 1q Consortium. Variation within the gene encoding the upstream stimulatory factor 1 does not influence susceptibility to type 2 diabetes in samples from populations with replicated evidence of linkage to chromosome 1q. *Diabetes* 2006 **55** 2541–2548.
- 25 Reiner AP, Carlson CS, Jenny NS, Durda JP, Siscovick DS, Nickerson DA & Tracy RP. USF1 gene variants, cardiovascular risk, and mortality in European Americans: analysis of two US cohort studies. *Arteriosclerosis, Thrombosis, and Vascular Biology* 2007 **27** 2736–2742.
- 26 Gibson F, Herberg S & Froguel P. Common polymorphisms in the USF1 gene are not associated with type 2 diabetes in French Caucasians. *Diabetes* 2005 **54** 3040–3042.
- 27 Lee SK, Kim HJ, Kim BJ, Jo YS, Park KS, Baik HW, Hyun SH, Lee JC & Kim SA. Body mass index is associated with USF1 haplotype in Korean premenopausal women. *Journal of Korean Medical Science* 2008 **23** 83–88.
- 28 Auro K, Kristiansson K, Zethelius B, Berne C, Lannfelt L, Taskinen MR, Jauhiainen M, Perola M, Peltonen L & Syvänen AC. USF1 gene variants contribute to metabolic traits in men in a longitudinal 32-year follow-up study. *Diabetologia* 2008 **51** 464–472.
- 29 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. *Human Molecular Genetics* 2004 **13** 1587–1597.
- 30 Meisinger C, Thorand B, Schneider A, Stieber J, Döring A & Löwel H. Sex differences in risk factors for incident type 2 diabetes mellitus: the MONICA Augsburg cohort study. *Archives of Internal Medicine* 2002 **162** 82–89.
- 31 Thorand B, Baumert J, Kolb H, Meisinger C, Chambless L, Koenig W & Herder C. Sex differences in the prediction of type 2 diabetes by inflammatory markers: results from the MONICA/KORA Augsburg case-cohort study, 1984–2002. *Diabetes Care* 2007 **30** 854–860.
- 32 Kolz M, Baumert J, Müller M, Khuseynova N, Klopp N, Thorand B, Meisinger C, Herder C, Koenig W & Illig T. Association between variations in the TLR4 gene and incident type 2 diabetes is modified by the ratio of total cholesterol to HDL-cholesterol. *BMC Medical Genetics* 2008 **9** 9.
- 33 Wichmann HE, Gieger C, Illig T & for the MONICA/KORA Study Group. . KORA-gen – resource for population genetics, controls and broad spectrum of disease phenotypes. *Gesundheitswesen* 2005 **67** 26–30.
- 34 Holle R, Happich M, Löwel H, Wichmann HE & MONICA/KORA Study Group. KORA – a research platform for population based health research. *Gesundheitswesen* 2005 **67** S19–S25.
- 35 Thorand B, Kolb H, Baumert J, Koenig W, Chambless L, Meisinger C, Illig T, Martin S & Herder C. Elevated levels of interleukin-18 predict the development of type 2 diabetes: results from the MONICA/KORA Augsburg Study, 1984–2002. *Diabetes* 2005 **54** 2932–2938.
- 36 Herder C, Baumert J, Thorand B, Koenig W, de Jager W, Meisinger C, Illig T, Martin S & Kolb H. Chemokines as risk factors for type 2 diabetes: results from the MONICA/KORA Augsburg Study, 1984–2002. *Diabetologia* 2006 **49** 921–929.
- 37 Herder C, Klopp N, Baumert J, Müller M, Khuseynova N, Meisinger C, Martin S, Illig T, Koenig W & Thorand B. Effect of macrophage migration inhibitory factor (MIF) gene variants and MIF serum concentrations on the risk of type 2 diabetes: results from the MONICA/KORA Augsburg case-cohort study, 1984–2002. *Diabetologia* 2008 **51** 276–284.
- 38 Barlow WE. Robust variance estimation for the case-cohort design. *Biometrics* 1994 **50** 1064–1072.
- 39 Li J & Ji L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity* 2005 **95** 221–227.
- 40 Cai J & Zeng D. Sample size/power calculation for case-cohort studies. *Biometrics* 2004 **60** 1015–1024.
- 41 Kristiansson K, Ilveskoski E, Lehtimäki T, Peltonen L, Perola M & Karhunen PJ. Association analysis of allelic variants of USF1 in coronary atherosclerosis. *Arteriosclerosis, Thrombosis, and Vascular Biology* 2008 **28** 983–989.
- 42 Komulainen K, Alanne M, Auro K, Kilpikari R, Pajukanta P, Saarela J, Ellonen P, Salminen K, Kulathinal S, Kuulasmaa K, Silander K, Salomaa V, Perola M & Peltonen L. Risk alleles of USF1 gene predict cardiovascular disease of women in two prospective studies. *PLoS Genetics* 2006 **2** e69.
- 43 Ng MC, Miyake K, So WY, Poon EW, Lam VK, Li JK, Cox NJ, Bell GI & Chan JC. The linkage and association of the gene encoding upstream stimulatory factor 1 with type 2 diabetes and metabolic syndrome in the Chinese population. *Diabetologia* 2005 **48** 2018–2024.
- 44 Van der Vleuten GM, Isaacs A, Hijmans A, van Duijn CM, Stalenhoef AFH & de Graaf J. The involvement of upstream stimulatory factor 1 (USF1) in Dutch patients with familial combined hyperlipidemia. *Journal of Lipid Research* 2007 **48** 193–200.
- 45 Hoffstedt J, Rydén M, Wahrenberg H, van Harmelen V & Arner P. Upstream transcription factor-1 gene polymorphism is associated with increased adipocyte lipolysis. *Journal of Clinical Endocrinology and Metabolism* 2005 **90** 5356–5360.
- 46 Kantartzis K, Fritsche A, Machicao F, Stumvoll M, Machann J, Schick F, Häring HU & Stefan N. Upstream stimulatory factor 1 gene polymorphisms are associated with high antilipolytic insulin sensitivity and show gene-gene interactions. *Journal of Molecular Medicine* 2007 **85** 55–61.