

M. C. Y. Ng · K. Miyake · W. Y. So · E. W. M. Poon ·
V. K. L. Lam · J. K. Y. Li · N. J. Cox · G. I. Bell ·
J. C. N. Chan

The linkage and association of the gene encoding upstream stimulatory factor 1 with type 2 diabetes and metabolic syndrome in the Chinese population

Received: 10 January 2005 / Accepted: 1 May 2005 / Published online: 26 August 2005
© Springer-Verlag 2005

Abstract *Aims/hypothesis:* The transcription factor upstream stimulatory factor 1 (USF1) regulates the expression of genes involved in glucose and lipid metabolism and has been associated with familial combined hyperlipidaemia. *USF1* is located on chromosome 1q22–23, a region with evidence for linkage to type 2 diabetes and various traits of the metabolic syndrome in Chinese and other populations. The aim of this study was to investigate the linkage and association of *USF1* with type 2 diabetes and the metabolic syndrome in Chinese individuals. *Materials and methods:* We genotyped three haplotype-tagging single nucleotide polymorphisms (SNPs) (rs3737787, rs2516841 and rs2516839) at *USF1* in three samples of the Hong Kong Chinese population, including members of 179 families from the Hong Kong Family Diabetes Study, 1,383 hospital cases with type 2 diabetes and/or the metabolic syndrome and 454 normal control subjects. *Results:* We found significant association of individual polymorphisms

and haplotypes with type 2 diabetes and/or metabolic syndrome-related traits in the family samples using either family-based or unrelated normal control subjects. However, these variants could not explain much of the evidence for linkage in this region. Moreover, they were not associated with type 2 diabetes and/or the metabolic syndrome in the hospital cases. *Conclusions/interpretation:* The results are consistent with the hypothesis that variation at *USF1* contributes to the risk of type 2 diabetes and the metabolic syndrome in families with strong evidence for linkage in the chromosome 1q region. However, they provide little support for *USF1* as the susceptibility locus that generates the observed evidence for linkage at 1q21–25 for type 2 diabetes and/or the metabolic syndrome, and *USF1* does not appear to have a major contribution to these phenotypes in the general Chinese population.

Keywords Chinese · Genetics of type 2 diabetes · Metabolic syndrome · Upstream stimulatory factor 1

Electronic Supplementary Material Supplementary material is available in the online version of this article at <http://dx.doi.org/10.1007/s00125-005-1914-0>

M. C. Y. Ng (✉) · K. Miyake · G. I. Bell
Department of Biochemistry and Molecular Biology,
University of Chicago,
Chicago, IL, USA
e-mail: maggieng@uchicago.edu
Tel.: +1-773-7029118
Fax: +1-773-7029237

W. Y. So · E. W. M. Poon · V. K. L. Lam · J. K. Y. Li ·
J. C. N. Chan
Department of Medicine and Therapeutics,
Prince of Wales Hospital,
The Chinese University of Hong Kong,
Shatin, Hong Kong SAR

N. J. Cox · G. I. Bell
Department of Human Genetics, University of Chicago,
Chicago, IL, USA

N. J. Cox · G. I. Bell
Department of Medicine, University of Chicago,
Chicago, IL, USA

Abbreviations $\lambda_{s(\text{locus})}$: locus-specific sibling relative risk · FCHL: familial combined hyperlipidaemia · FPG: fasting plasma glucose · HKFDS: Hong Kong Family Diabetes Study · HOMA%IR: homeostasis model assessment for insulin resistance · HOMA% β : homeostasis model assessment for beta cell function · LD: linkage disequilibrium · LOD: log of the odds · OR: odds ratio · $p_{\text{empirical}}$: empirical p value · SNP: single-nucleotide polymorphism · STEPC: Statistical Explanation for Positional Cloning · USF1: upstream stimulatory factor 1

Introduction

The clustering of multiple metabolic abnormalities, such as obesity, dyslipidaemia and hypertension, is common in patients with type 2 diabetes. We have previously reported a locus on chromosome 1q21–25 with significant evidence for linkage to type 2 diabetes, the metabolic syndrome and waist circumference in families from the Hong Kong Family Diabetes Study (HKFDS) [1, 2]. This chromosomal

region has also been linked to type 2 diabetes or glucose intolerance, the metabolic syndrome, obesity, and dyslipidaemia or familial combined hyperlipidaemia (FCHL) in other populations [3–10].

Recently, Pajukanta et al. found association and linkage of variation at the gene for the transcription factor upstream stimulatory factor 1 (*USF1*) on chromosome 1q22–23 with FCHL in a Finnish population [11]. Since the *USF1* transcription factor regulates both glucose and lipid metabolism [12], we hypothesised that *USF1* might be a candidate gene for susceptibility to type 2 diabetes, the metabolic syndrome or obesity. Our primary goal was to test whether variation at *USF1* could explain the observed evidence for linkage of type 2 diabetes and/or the metabolic syndrome in this region in the HKFDS families included in our genome scans. We also wanted to assess the evidence for association of variation at *USF1* with disease and metabolic phenotypes in the families, and association of this variation with disease phenotypes in independent samples of hospital cases and normal control subjects.

Subjects and methods

Subjects

All subjects were of southern Han Chinese ancestry, residing in Hong Kong. The study included individuals from three independent samples. The first sample (family cases) included 179 families (897 subjects, average family size 5.0 ± 2.3) recruited from the HKFDS. The details of ascertainment, exclusion criteria and phenotyping are described elsewhere [1, 2]. All families were recruited through a proband selected from the Prince of Wales Hospital Registry. Eighty-three percent of the probands were diagnosed at ≤ 40 years of age. The second sample (hospital cases) consisted of 1383 unrelated type 2 diabetic patients randomly selected from the same registry and 34% were diagnosed ≤ 40 years [13]. The third sample (control subjects) consisted of 454 normal control subjects (fasting plasma glucose [FPG] < 6.1 mmol/l) recruited from the general population participating in a community-based cardiovascular risk screening programme, as well as hospital staff. Control subjects with glucose intolerance (IFG, IGT or diabetes) [14], the metabolic syndrome (National Cholesterol Education Program in Adult Treatment Panel III criteria) [15] or a family history of diabetes were excluded. The clinical characteristics of subjects in the three samples are summarised in Table 1 of the Electronic supplementary material. Informed consent was obtained for each participating subject. This study was approved by the Clinical Research Ethics Committee of the Chinese University of Hong Kong.

Clinical studies

All subjects underwent detailed clinical investigation as described previously [2, 13]. A fasting blood sample was collected for measurement of plasma glucose, insulin, lip-

Table 1 Association of *USF1* SNPs and haplotypes with type 2 diabetes and metabolic syndrome using FBAT

Marker	Haplotype frequency	Type 2 diabetes ($n=343$) ^a <i>P</i> value	Metabolic syndrome ($n=273$) ^b <i>P</i> value
rs3737787	C: 0.896	0.004* (+)	0.057
rs2516841	C: 0.832	0.014* (–)	0.840
rs2516839	G: 0.707	0.070	0.165
Haplotype ^c	CCG: 0.548	0.022*	0.260
	CCA: 0.172		
	CTG: 0.148		
	TCA: 0.131		

Data were analysed by FBAT using the additive genetic model for the common allele in the Chinese population. Nucleotides refer to the complementary contig sequence NT_079484 (NCBI build 34). The order of markers in haplotypes is in the sequence rs3737787, rs2516841 and rs2516839

* Significant association ($p < 0.05$). The direction of allele transmission is indicated: + indicates over-transmission and – indicates under-transmission of the common allele

^aThe type 2 diabetes study included 343 affected individuals in 174 families

^bThe metabolic syndrome study included 273 affected individuals in 138 families

^cTest with four degrees of freedom

ids, liver and renal function, and DNA. All family members and 335 of the control subjects who had no history of diabetes underwent a 75 g OGTT. Three fasting blood samples collected at intervals of 5 min were assessed for mean FPG, insulin and C-peptide. Blood samples were also collected at 15, 30, 60 and 120 min during the OGTT for measurement of plasma glucose and insulin. Using the homeostasis model assessment (HOMA), the insulin resistance index (HOMA%IR) was assessed as fasting insulin (mU/l) \times FPG (mmol/l) / 22.5, and beta cell function (HOMA% β) was assessed as fasting insulin \times 20 / (FPG – 3.5) [16].

The metabolic syndrome was defined according to the NCEP III guidelines [15]. Subjects who had at least three of the following five risk factors were classified as having the metabolic syndrome: (1) hyperglycaemia with known diabetes or FPG ≥ 6.1 mmol/l; (2) known hypertension or BP $\geq 130/85$ mmHg; (3) hypertriglyceridaemia with triglyceride ≥ 1.7 mmol/l; (4) HDL < 1.0 mmol/l in men or < 1.3 mmol/l in women; and (5) central obesity with waist circumference > 90 cm in men or > 80 cm in women. The definition of central obesity was modified for Asian populations [17].

Genotyping

Among the common single nucleotide polymorphisms (SNPs) observed in *USF1* and the adjacent gene *F11R* by sequencing in a Finnish study [11], we genotyped nine SNPs (rs836, rs790056, rs4339888, rs3737787, rs2073658, rs2516841, rs2516839, rs2516838 and rs1556259), representing each of the linkage disequilibrium (LD) clusters in 38 unrelated family cases and 51 control subjects (Fig. 1).

We estimated pairwise LD between markers using the GOLD-ldmax program [18] and haplotype frequency using the PHASE program (v. 2.1) [19]. Four common haplotypes (frequency >0.01) constitute 99.7% of all haplotypes and three haplotype-tagging SNPs (rs3737787, rs2516841 and rs2516839) in *USF1* were genotyped in all subjects to capture these four common haplotypes ([20]; $r^2 \geq 0.5$, minor allele frequency ≥ 0.1). All SNPs were genotyped with the TaqMan assay (Applied Biosystems, Foster City, CA, USA). Blind duplicates of 68 samples showed no discordant genotypes. For family data, we removed Mendelian errors and potential genotyping errors identified in 0.4% of the data using PedCheck (v. 1.1) [21] and Merlin (v. 0.10.1) [22], respectively. A total of 0.4% of genotype data were missing. We assessed Hardy–Weinberg equilibrium with the χ^2 test in each study group separately. The marker rs2516839 showed moderate departure from Hardy–Weinberg equilibrium ($p=0.05$) in the hospital cases, but this is consistent with the possibility that variation in the region affects phenotype using a goodness-of-fit test [23]. Moreover, we could assign the four common haplotypes to most cases with high probability (>0.85) using PHASE. The departure from Hardy–Weinberg equilibrium is therefore less likely to be due to genotyping error but rather to reflect departure of this local region from Hardy–Weinberg equilibrium. Thus, all data were included in the subsequent analyses.

Statistical analyses

Continuous data were transformed by natural logarithm if necessary and expressed as mean \pm SD or geometric mean (95% CI). Categorical data were compared using the χ^2 test or Fisher's exact test as appropriate. All statistical tests were performed with SPSS (SPSS for Windows, v. 11.5; SPSS, Chicago, IL, USA) unless specified otherwise. Given the a priori hypotheses being tested and the correlations in phenotype data, no adjustment for multiple comparisons was made. A p value less than 0.05 was considered significant (two-tailed).

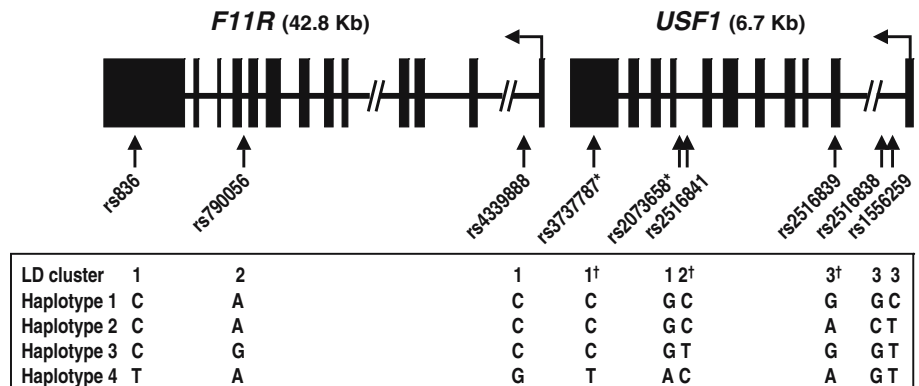
Family-based linkage and association analyses

We used Statistical Explanation for Positional Cloning (STEP-C) [24] to determine if the SNPs could fully explain the observed linkage for type 2 diabetes and the metabolic syndrome. The hypothesis was that, if a particular marker is the only site that influences the evidence for linkage at the trait, no residual linkage will remain after conditioning on the genotypes of the marker segregating in the families. We performed covariate analyses to determine whether the SNPs explained the evidence for linkage for waist circumference. We performed variance component linkage analyses for waist circumference using Merlin with adjustment for age and sex, and compared the result with results obtained using an additional individual SNP (in an additive model) as covariates.

We used subset analyses to assess whether the SNPs were associated with linkage results for type 2 diabetes and the metabolic syndrome. We used Merlin to select the affected individual from each family who shared the most alleles that were identical by descent at the peak linkage region (175 centimorgans) with the other affected family members [22]. These subjects tend to harbour the disease allele in linked families [25]. We divided the families into three subsets based on the genotype of the selected affected individuals. We performed non-parametric multipoint linkage analyses for type 2 diabetes and the metabolic syndrome in each subset and the combined samples using Merlin with the score-pairs option, as described previously [1]. Significance was assessed for the subset with the highest linkage signal in each SNP by performing 1,000 simulations with random selection of the same number of families as in the subset analyses. The empirical p value ($P_{\text{empirical}}$) was calculated as the proportion of simulations with a LOD (log of the odds) score greater than or equal to the observed LOD score.

We used the FBAT program (v. 1.5.5) [26] to test for the association between SNPs or haplotypes and the traits under an additive genetic model. We tested the null hypothesis of no association in the presence of linkage with the $-e$ option, which empirically estimated the variance. In haplotype analyses, the association was assessed with the

Fig. 1 Structure of *F11R* and *USF1* and the location of the SNPs studied. The LD clusters refer to groups of SNPs with $r^2 \geq 0.5$. Common haplotypes (frequency >0.01) are shown. †Haplotype tagging SNPs; *SNPs showing the most significant association with FCHL in the Finnish population [11]



overall haplotype frequency distribution for the four common haplotypes with at least ten informative families. In quantitative trait analyses, data on BP and glucose traits were not used if the subjects were taking BP- or glucose-lowering medications. Data were regressed on age and sex and residuals with outliers greater than four standard deviations from the mean were removed. The residuals were then standardised to mean zero and unit variance for FBAT analyses.

Case-control association analyses

We hypothesised that, if a SNP is associated with disease, the risk allele should have the highest frequency in family cases contributing to the linkage signal (linked-family cases), compared with all family cases and hospital cases, whereas control subjects will have the lowest frequency. We tested this hypothesis using three samples. The first sample included family cases in which one subject with either type 2 diabetes or the metabolic syndrome was selected per family. This corresponded to the subject showing the most evidence of allele sharing identity by descent with other affected family members (all family cases). Subsets of cases were also selected based on having a positive family LOD score at the linkage peak for the respective trait (linked-family cases). The second sample included 1,383 unrelated type 2 diabetic patients (hospital cases) for testing association with type 2 diabetes. A subset of 812 type 2 diabetic patients who also had the metabolic syndrome was selected for testing association with the metabolic syndrome. The third sample included 454 subjects without diabetes and the metabolic syndrome (control sample). The allele frequency of each case group was compared with that of the control subjects using a 2×2 χ^2 table. Odds ratios (OR) with 95% CI are presented. We also compared the overall haplotype frequency distribution in each case group vs. the control subjects using PHASE under the condition of no recombination, and assessed the significance through 100 permutations [19].

Power of association studies

We estimated the locus-specific sibling relative risk ($\lambda_{s(\text{locus})}$) by calculating the ratio of expected vs observed probability of identity by descent=0 at the peaks of our linkage signals using ASPEX (v. 2.3) [27]. We then assessed the posterior power of our study to detect associations for genetic models that were consistent with the observed $\lambda_{s(\text{locus})}$ [28, 29].

Results

Linkage disequilibrium in the *USF1* region in Chinese individuals

We initially typed nine SNPs in the *USF1-F1IR* region (Fig. 1). All SNPs were in strong LD ($D' \geq 0.6$) and formed

three clusters with $r^2 \geq 0.5$ (cluster 1: rs836, rs4339888, rs3737787, rs2073658; cluster 2: rs790056, rs2516841; cluster 3: rs2516839, rs2516838, rs1556259). Subsequently, three haplotype-tagging SNPs, including rs3737787, rs2516841 and rs2516839, were selected for typing in all samples. The pattern of LD in the *USF1* region in Chinese was similar to that reported in the Finnish population [11] and the Han Chinese and CEPH populations of the Hap-Map project [30]. There were, however, differences in the minor allele frequencies in Chinese family samples compared with the Finnish family samples (allele T, 0.10 vs 0.34; allele T, 0.17 vs 0.28; allele A, 0.29 vs 0.61 for rs3737787, rs2516841 and rs2516839, respectively) (Table 1).

Family-based linkage and association

STEP analysis showed significant residual evidence for linkage conditioning on any of the individual SNPs for type 2 diabetes ($0.0002 < p < 0.05$) and the metabolic syndrome ($0.00001 < p < 0.0009$). Comparison of linkage results for waist circumference with or without using *USF1* SNPs as a covariate showed no effect on linkage peak (original vs SNP-adjusted LOD scores, 4.79 vs 4.16–4.40 at 178 centimorgans and 3.80 vs 3.45–4.09 at 200 centimorgans). Thus, none of the polymorphisms can explain the observed linkage signals for type 2 diabetes, the metabolic syndrome or waist circumference. We also examined the association of *USF1* SNPs with the evidence of linkage for type 2 diabetes and the metabolic syndrome by subset linkage analyses based on the genotype of an affected individual in each family. We did not observe significant associations, but the rs3737787 CC genotype and rs2516841 CT genotype showed a trend of association with the evidence for linkage for type 2 diabetes ($p_{\text{empirical}} = 0.10$ and 0.09, respectively) and similarly for the rs2516839 GG genotype

Table 2 Association of *USF1* SNPs with metabolic traits in all family members in HKFDS ($n=897$); p values are shown

Trait	rs3737787	rs2516841	rs2516839	Haplotype ^a
Waist circumference	0.057	0.130	0.692	0.010*
Systolic BP	0.512	0.744	0.959	0.765
Diastolic BP	0.841	0.596	0.452	0.523
Total cholesterol	0.858	0.578	0.239	0.560
Triglyceride	0.494	0.423	0.443	0.552
HDL	0.531	0.457	0.490	0.557
LDL	0.832	0.417	0.269	0.287
HOMA%IR	0.174	0.544	0.064	0.140
HOMA%β	0.181	0.286	0.161	0.034*

Quantitative traits were analysed with the FBAT program using the additive genetic model for the common allele in the Chinese population. Data are logarithmically transformed for triglyceride, HOMA%IR and HOMA%β

* Significant association ($p < 0.05$). Direction of effect is indicated as + for increased and – for decreased trait values associated with the common allele

^aHOMA%IR insulin resistance index, HOMA%β beta cell function
^aTest with 4 degrees of freedom

with the evidence for linkage for the metabolic syndrome ($p_{\text{empirical}}=0.08$).

We further assessed the evidence of association of *USF1* SNPs and haplotypes with type 2 diabetes and the metabolic syndrome in the presence of linkage in our families, using FBAT. We found that the common C allele of rs3737787 and the rare T allele of rs2516841 were significantly associated with higher risk of type 2 diabetes ($p=0.004$ and 0.014 , respectively) with an overall significant haplotype effect ($p=0.022$). The C allele of rs3737787 also showed a trend to increasing risk of the metabolic syndrome ($p=0.057$) (Table 1). Moreover, we assessed the association of these SNPs and haplotypes with age- and sex-adjusted metabolic traits in all family members. Significant haplotype associations were observed for waist circumference ($p=0.010$) and HOMA% β ($p=0.034$), with a trend to association of rs3737787 with waist circumference ($p=0.057$) (Table 2).

Case-control association

We also examined the association of *USF1* with type 2 diabetes and the metabolic syndrome in both unrelated

family and hospital cases compared with normal control subjects. There was an overall significant haplotype effect in a comparison of family cases and normal control subjects for type 2 diabetes ($p=0.01$) but not for the metabolic syndrome. Moreover, the common C allele at rs3737787 was associated with significantly higher risk of type 2 diabetes and the metabolic syndrome in all family cases (OR=1.55–1.56, $p<0.05$) (Table 3). This association was even stronger in the subset of cases from linked (positive LOD score at 175 centimorgans) families (OR 2.33 for type 2 diabetes and 2.60 for the metabolic syndrome, $p<0.05$) (Table 3). In contrast to results for the family cases, there was no association of *USF1* SNPs or haplotypes with type 2 diabetes and the metabolic syndrome when comparing hospital cases and normal control subjects (Table 3). Using genetic models consistent with the $\lambda_{s(\text{locus})}$ of 1.7 for type 2 diabetes and 2.8 for the metabolic syndrome in our linkage studies, we had more than 99% power to detect association in the case-control studies at $r^2 \geq 0.5$ between the marker and disease locus under our primary hypothesis that the disease locus accounts for the evidence for linkage in this region.

Table 3 Case-control association of *USF1* SNPs with type 2 diabetes and metabolic syndrome

	Type 2 diabetes				Metabolic syndrome			
	<i>n</i>	Common allele frequency	Odds ratio (95% CI)	<i>p</i> _{allele}	<i>n</i>	Common allele frequency	Odds ratio (95% CI)	<i>p</i> _{allele}
rs3737787								
Control subjects	454	0.841	1		454	0.841	1	
Hospital cases	1,383	0.837	0.97 (0.79–1.19)	0.752	812	0.834	0.95 (0.76–1.18)	0.645
All family cases	175	0.891	1.55 (1.06–2.26)	0.024*	139	0.892	1.56 (1.03–2.36)	0.037*
Linked family cases	40	0.925	2.33 (1.02–5.32)	0.046*	37	0.932	2.60 (1.07–6.35)	0.036*
rs2516841								
Control subjects	454	0.852	1		454	0.852	1	
Hospital cases	1,383	0.844	0.94 (0.76–1.16)	0.550	811	0.854	1.01 (0.80–1.27)	0.921
All family cases	175	0.834	0.87 (0.62–1.22)	0.423	139	0.820	0.79 (0.55–1.13)	0.193
Linked family cases	40	0.788	0.64 (0.37–1.13)	0.122	37	0.797	0.68 (0.38–1.23)	0.204
rs2516839								
Control subjects	454	0.694	1		454	0.694	1	
Hospital cases	1,381	0.667	0.88 (0.75–1.04)	0.128	812	0.660	0.86 (0.72–1.02)	0.083
All family cases	175	0.700	1.03 (0.79–1.35)	0.831	139	0.712	1.09 (0.81–1.47)	0.559
Linked family cases	40	0.687	0.97 (0.59–1.59)	0.906	37	0.730	1.19 (0.70–2.03)	0.518
Haplotype								
Control subjects	454				454			
Hospital cases	1,383			0.13	812			0.24
All family cases	175			0.01*	139			0.10
Linked family cases	40			0.13	37			0.07

Allele frequencies were compared in each disease group with the control subjects using a 2×2 χ^2 table. Odds ratios and 95% CIs are shown for disease associations with respect to the common allele. An overall haplotype effect was compared in each disease group with the control subjects using PHASE with 100 permutations

*Significant associations ($p<0.05$)

Discussion

We have previously demonstrated significant linkage of chromosome 1q21–25 to type 2 diabetes, the metabolic syndrome and waist circumference in Chinese [1, 2]. Here we show that variation at *USF1* in this region was unable to explain much of the evidence for the linkage using STEPC, covariate and subset linkage analyses. However, *USF1* is associated with type 2 diabetes, the metabolic syndrome and related traits (waist circumference and HOMA%β) in the family data using both family-based and/or case-control association methods (Tables 1, 2, 3). In contrast to the associations observed in the family cases, we observed no association of *USF1* with either type 2 diabetes or the metabolic syndrome in the population-based (hospital) type 2 diabetic cases.

Although we did not resequence *USF1* in our population to search for Chinese-specific polymorphisms, private SNPs tend to have very low minor allele frequency and the strong signals observed in our linkage studies would require cumulative susceptibility allele frequencies in the intermediate range [31]. In addition, the LD pattern of D' and r^2 around *USF1* are strong and similar between Han Chinese and CEPH populations in the HapMap data [30], suggesting that the SNPs we studied are representative of *USF1*. The strongest associations were observed for SNP rs3737787 in most analyses. This SNP is located in a region of strong LD (D') among the genes *ITLN2*, *F11R*, *USF1* and *PVRL4*, but is in strong LD only with SNPs in *F11R* and *USF1*. Collectively, the results suggest that variation at *USF1*, or an as yet unidentified variant in strong LD located at or near *USF1*, contributes to the risk of developing type 2 diabetes and/or metabolic syndrome in families with evidence for linkage to this region. However, the results provide little support for *USF1* as a major contributor to risk of type 2 diabetes or the metabolic syndrome in the general Chinese population.

Genetic variation at *USF1* has also been associated with FCHL. Pajukanta et al. [11] showed association and linkage of a *USF1* haplotype (the two common alleles for rs3737787 and rs2073658) with FCHL in Finnish families. The strongest association was with high triglyceride levels in men. We observed a similar trend for association of the rs3737787 common C allele with higher triglyceride levels in men from the family study ($p=0.051$). We also selected a subset of 88 patients having combined hyperlipidaemia (>90 percentiles of the age- and sex-specific population references of both total cholesterol and triglyceride) [32] from the population-based (hospital) cases. We found no association of this phenotype with variation at *USF1* (data not shown).

The discrepancy in the results between family-based and population-based (hospital) studies described here was also observed in studies of *USF1* and FCHL. The Finnish family study showed association and linkage of variation at *USF1* with FCHL in families with evidence for linkage on chromosome 1q21–23 [11]. However, a case-control study

in Europeans was not able to replicate this result using population-based cases with FCHL [33]. Why do we and others [11, 33] observe association of *USF1* variation with metabolic diseases in families with linkage to chromosome 1q region but not in population-based samples? We cannot exclude the possibility that this is due to differences in ascertainment criteria, clinical heterogeneity and the ethnicity of the study samples. On the other hand, variation at *USF1* showed stronger association with the diseases in cases from families demonstrating linkage, but the segregation of that variation in families could not explain much of that evidence for linkage. Given our strong linkage signals, we expect a large effect size and good power to detect association, given the hypothesis that it is variation at *USF1* that explains the evidence for linkage. Our results implicate additional susceptibility gene(s) in this region for at least the type 2 diabetes and the metabolic syndrome phenotypes that we studied. These findings have implications for the positional cloning of the gene(s) contributing to the development of metabolic diseases in the chromosome 1q22–23 region. It is plausible that an unidentified gene(s) may play a major role in the development of type 2 diabetes and related diseases and that *USF1* could act as a modifier that interacts with this gene; this is consistent with the absence of associations in population-based studies. Thus, it will be interesting to test whether *USF1* accounts for the evidence for linkage in the Finnish family study [11]. Clearly, however, further studies in other populations, including both population-based and family samples, and in particular family samples showing evidence for linkage on chromosome 1q, may provide a better understanding of the contribution of this gene to the regulation of glucose and lipid metabolism.

Acknowledgements We thank Ms Cherry Chiu for recruitment of families in HKFDS and Dr Ying Wang for recruitment of control subjects. We thank Mr Kevin Yu for management of the PWH Diabetes Registry. We thank Mr William Wen for computational support. We thank all nursing and medical staff at the PWH Diabetes and Endocrine Centre for their dedication and professionalism. We thank all subjects participating in this study. This work was supported by grants from the Hong Kong Research Grants Committee, the Chinese University of Hong Kong Strategic Grant Program and the Innovation and Technology Support Fund, Hong Kong (ITS/33/00), and US Public Health Service Grants DK-20595, DK-47486, DK-55889 and DK-58026.

References

1. Ng MCY, So WY, Cox NJ et al (2004) Genome-wide scan for type 2 diabetes loci in Hong Kong Chinese and confirmation of a susceptibility locus on chromosome 1q21–q25. *Diabetes* 53: 1609–1613
2. Ng MCY, So WY, Lam VKL et al (2004) Genome-wide scan for metabolic syndrome and related quantitative traits in Hong Kong Chinese and confirmation of a susceptibility locus on chromosome 1q21–q25. *Diabetes* 53:2676–2683
3. Pajukanta P, Nuotio I, Terwilliger JD et al (1998) Linkage of familial combined hyperlipidaemia to chromosome 1q21–q23. *Nat Genet* 18:369–373

4. Coon H, Myers RH, Borecki IB et al (2000) 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. *Arterioscler Thromb Vasc Biol* 20:2275–2280
5. Pei W, Baron H, Muller-Myhsok B et al (2000) Support for linkage of familial combined hyperlipidemia to chromosome 1q21–q23 in Chinese and German families. *Clin Genet* 57:29–34
6. Arya R, Blangero J, Williams K et al (2002) Factors of insulin resistance syndrome-related phenotypes are linked to genetic locations on chromosomes 6 and 7 in nondiabetic Mexican-Americans. *Diabetes* 51:841–847
7. Broeckel U, Hengstenberg C, Mayer B et al (2002) A comprehensive linkage analysis for myocardial infarction and its related risk factors. *Nat Genet* 30:210–214
8. Feitosa MF, Borecki IB, Rich SS et al (2002) Quantitative-trait loci influencing body-mass index reside on chromosomes 7 and 13: the National Heart, Lung, and Blood Institute Family Heart Study. *Am J Hum Genet* 70:72–82
9. McCarthy MI (2003) Growing evidence for diabetes susceptibility genes from genome scan data. *Curr Diabetes Rep* 3: 159–167
10. Langefeld CD, Wagenknecht LE, Rotter JI et al (2004) Linkage of the metabolic syndrome to 1q23–q31 in Hispanic families: the Insulin Resistance Atherosclerosis Study family study. *Diabetes* 53:1170–1174
11. Pajukanta P, Lilja HE, Sinsheimer JS et al (2004) Familial combined hyperlipidemia is associated with upstream transcription factor 1 (*USF1*). *Nat Genet* 36:371–376
12. Shoulders CC, Naoumova RP (2004) *USF1* implicated in the aetiology of familial combined hyperlipidaemia and the metabolic syndrome. *Trends Mol Med* 10:362–365
13. Chan JCN, Lau M, Wong R et al (1997) Delivery of diabetes care—the experience at the Prince of Wales Hospital. *Hosp Authority Qual Bull* 2:3–21
14. Alberti KGMM, Zimmet PZ (1998) Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1. Diagnosis and classification of diabetes mellitus, provisional report of a WHO consultation. *Diabet Med* 15:539–553
15. Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (2001) Executive summary of the third report of the National Cholesterol Education Program (NCEP) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (Adult Treatment Panel III). *JAMA* 285:2486–2497
16. Matthews RD, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC (1985) Homeostasis model assessment: insulin resistance and β cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* 28:412–419
17. International Obesity Task Force (2000) The Asia-Pacific perspective: redefining obesity and its treatment. World Health Organization, Western Pacific Region
18. Abecasis GR, Cookson WO (2000) GOLD—graphical overview of linkage disequilibrium. *Bioinformatics* 16:182–183
19. Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 68:978–989
20. Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, Nickerson DA (2004) Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am J Hum Genet* 74:106–120
21. O'Connell JR, Weeks DE (1998) PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 63:259–266
22. Abecasis GR, Cherny SS, Cookson WO, Cardon LR (2002) Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 30:97–101
23. Wittke-Thompson JK, Pluzhnikov A, Cox NJ (2005) Rational inferences about departures from Hardy–Weinberg equilibrium. *Am J Hum Genet* 76:967–986
24. Sun L, Cox NJ, McPeck MS (2002) A statistical method for identification of polymorphisms that explain a linkage result. *Am J Hum Genet* 70:399–411
25. Fingerlin TE, Boehnke M, Abecasis GR (2004) Increasing the power and efficiency of disease-marker case-control association studies through use of allele-sharing information. *Am J Hum Genet* 74:432–443
26. Lake SL, Blacker D, Laird NM (2000) Family-based tests of association in the presence of linkage. *Am J Hum Genet* 67: 1515–1525
27. D.A. Hinds, N. Risch, (1996) The ASPEx package: affected sib-pair exclusion mapping. Available from <http://aspex.sourceforge.net>
28. Risch N (1990) Linkage strategies for genetically complex traits: I. Multilocus models. *Am J Hum Genet* 46:222–228
29. Gordon D, Finch SJ, Nothnagel M, Ott J (2002) Power and sample size calculations for case-control genetic association tests when errors are present: application to single nucleotide polymorphisms. *Hum Hered* 54:22–33
30. International HapMap Project. Available from <http://hapmap.org>. Accessed 23 March 2005
31. Pritchard JK, Cox NJ (2002) The allelic architecture of human disease genes: common disease–common variant ... or not? *Hum Mol Genet* 11:2417–2423
32. Janus ED (1997) The Hong Kong cardiovascular risk factor prevalence study 1995–1996. Department of Clinical Biochemistry, Queen Mary Hospital of Hong Kong, Hong Kong
33. Putt W, Palmen J, Nicaud V et al (2004) 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