**INSIG1** influences obesity-related hypertriglyceridemia in humans

E. M. Smith,\(^{1,2,\dagger,\ddagger}\) Y. Zhang,\(^{1,\ddagger,\dagger}\) T. M. Baye,\(^{6,\dagger}\) S. Gawrieh,\(^8\) R. Cole,\(^{\ddagger,\dagger}\) J. Blangero,\(^{\ddagger,\ddagger}\) M. A. Carless,\(^{1,\dagger}\) J. E. Curran,\(^{1,\dagger}\) T. D. Dyer,\(^{1,\dagger}\) L. J. Abraham,\(^{8}\) E. K. Moses,\(^{\ddagger,\ddagger}\) A. H. Kissebah,\(^8\) L. J. Martin,\(^{6,\ddagger,\dagger}\) and M. Olivier\(^3\)

Abstract In our analysis of a quantitative trait locus (QTL) for plasma triglyceride (TG) levels [logarithm of odds (LOD) = 3.7] on human chromosome 7q36, we examined 29 single nucleotide polymorphisms (SNPs) across **INSIG1**, a biological candidate gene in the region. Insulin-induced genes (**INSIGs**) are feedback mediators of cholesterol and fatty acid synthesis in animals, but their role in human lipid regulation is unclear. In our cohort, the **INSIG1** promoter SNP rs2721 was associated with TG levels \(P = 2 \times 10^{-5}\) in 1,560 individuals of the original linkage cohort, \(P = 8 \times 10^{-5}\) in 920 unrelated individuals of the replication cohort, combined \(P = 9.9 \times 10^{-6}\). Individuals homozygous for the T allele had 9% higher TG levels and 2-fold lower expression of **INSIG1** in surgical liver biopsy samples when compared with individuals homozygous for the G allele. Also, the T allele showed additional binding of nuclear proteins from HepG2 liver cells in gel shift assays. Finally, the variant rs7566605 in **INSIG2**, the only homolog of **INSIG1**, enhances the effect of rs2721 \(P = 0.00117\). The variant rs2721 alone explains 5.4% of the observed linkage in our cohort, suggesting that additional, yet-undiscovered genes and sequence variants in the QTL interval also contribute to alterations in TG levels in humans. – Smith, E. M., Y. Zhang, T. M. Baye, S. Gawrieh, R. Cole, J. Blangero, M. A. Carless, J. E. Curran, T. D. Dyer, L. J. Abraham, E. K. Moses, A. H. Kissebah, L. J. Martin, and M. Olivier. **INSIG1** influences obesity-related hypertriglyceridemia in humans. *J. Lipid Res.* 2010. 51: 701–708.

Increased plasma triglyceride (TG) levels are an important cardiovascular risk factor and are strongly associated with atherosclerotic heart disease (1, 2). Plasma TG levels vary widely between individuals, and both genetic and environmental factors have been shown to contribute to elevated plasma TG concentrations (3–7). Elevated plasma TG levels are often observed in obese and diabetic individuals and in individuals affected by the metabolic syndrome, a common chronic disorder associated with obesity, insulin resistance, hypertension, and alterations in plasma lipid profile such as elevated serum TG and low HDL levels. This characteristic pattern is similar to lipid abnormalities reported in familial combined hyperlipidemia (8).

The Metabolic Risk Complications of Obesity Genes (MRC-OB) project was established in 1994 to identify the genetic determinants of the metabolic syndrome and its metabolic abnormalities (9). As part of the project, basic anthropomorphid phenotypes, plasma lipid measures, and fasting glucose and insulin levels were ascertainment for 2,209 individuals from 507 families. A genome-wide linkage scan of these families identified a quantitative trait locus (QTL) on human chromosome 7q36 linked to plasma TG levels (LOD = 3.7) (10). This region has also been implicated in numerous other studies (11–15) and represents a common genetic variant associated with metabolic syndrome.

**Abbreviations**: BMI, body mass index; EMSA, electrophoretic mobility shift assay; GWAS, genome-wide association studies; **INSIGs**, insulin-induced genes; LD, linkage disequilibrium; LDL-C, LDL-cholesterol; ln, natural log; LOD, logarithm of odds; MAF, minor allele frequency; MRC-OB, Metabolic Risk Complications of Obesity Genes project; PBX1, Pre-B-cell leukemia homeobox 1; QTL, quantitative trait locus; SCAPs, SREBP cleavage-activating proteins; SNP, single nucleotide polymorphism; SREBP, sterol regulatory element binding protein; TG, triglyceride.

\(^{1}\) E. M. Smith and Y. Zhang contributed equally to this work.

\(^{2}\) Present address: Department of Biological Sciences, University of Warwick, Coventry, CV4 7AL, UK.

\(^{3}\) To whom correspondence should be addressed. e-mail: molivier@mcw.edu

This work was supported by grant HL74168 from the Heart, Lung, and Blood Institute of the National Institutes of Health (M.O.). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health or other granting agencies.

Manuscript received 12 October 2009 and in revised form 14 October 2009.

Published, JLR Papers in Press, October 14, 2009

DOI 10.1194/jlr.M001404

Copyright © 2010 by the American Society for Biochemistry and Molecular Biology, Inc.

This article is available online at http://www.jlr.org
one of the most replicated linkages for dyslipidemia and elevated TG levels. The genomic interval on chromosome 7q36 includes 22 known genes, including insulin-induced gene 1 (INSIG1), the only biological candidate.

INSIG1 was initially described as an insulin-induced transcript, designated CL-6 (16). It was later renamed INSIG1 following cloning and assignment to chromosome 7q36 (17). INSIG1 has been shown to affect the feedback regulation of cellular cholesterol (18, 19). Sterol regulatory element binding proteins (SREBPs), which are transcription factors located in the endoplasmic reticulum, control the synthesis of cholesterol and fatty acids. When cellular cholesterol levels fall, SREBPs bind to SREBP cleavage-activating proteins (SCAPs) to be transported to the Golgi, where they undergo proteolytic cleavage (20). The NH2-terminal fragment translocates to the nucleus to activate the transcription of target genes, including INSIG1 (21), that are involved in the synthesis of cholesterol and fatty acids. As cellular sterol levels rise, SCAP binds to INSIG1, thus preventing the transport of the SCAP/SREBP complex to the Golgi (18) and halting the synthesis of cholesterol and fatty acids. Overexpression of INSIG1 has been shown to reduce high levels of TG in both liver and plasma of Zucker diabetic fatty rats (25). It has been proposed that this effect is due to the mechanism outlined above (18). In addition, single knockout of either INSIG1 or INSIG2 in mice leads to increased total content of both cholesterol and TG in mouse livers, an effect exacerbated significantly when both INSIG genes are knocked out (26). These studies indicate that INSIG1 and INSIG2 interactively contribute to the control and regulation of lipid homeostasis in animals.

Based on this evidence, we hypothesized that sequence variation in INSIG1, which is located in the QTL region on human chromosome 7q36 linked to plasma TG levels, contributes to the observed genetic effect on plasma lipid levels. Here we present data on our comprehensive association analysis of sequence variants in INSIG1 in the initial MRC-OB cohort and a second replication cohort of unrelated individuals, and evidence for a synergistic effect of an associated sequence variant in INSIG1 with a variant in INSIG2 on plasma TG levels. Furthermore, our data suggest that a putative causal promoter variant in INSIG1 affects gene expression of INSIG1 in liver tissue, potentially by altering the binding of nuclear factors to the promoter sequence.

METHODS

Study design
We performed a family-based study on genetic contributions to plasma TG levels.

Study cohort. Cohort recruitment and individual phenotyping have been described in detail previously (9, 10). In brief, families with at least two obese siblings [body mass index (BMI) >30], the availability of at least one parent, and one or more never obese siblings (BMI <27) were recruited from the Take Off Pounds Sensibly, Inc. membership in 10 Midwestern U.S. states. Health information of all participants was obtained by a questionnaire. Individuals with the following conditions were excluded from recruitment: pregnancy; type 1 diabetes; history of cancer, renal, or hepatic disease; severe coronary artery disease; substance abuse; corticosteroids or thyroid dosages above replacement dose; history of weight loss of >10% in the preceding 12 months; and individuals receiving lipid-lowering medications. Phenotypic measurements included height, weight, waist and hip circumferences, and fasting plasma levels of glucose, insulin, total cholesterol, LDL-cholesterol (LDL-C), HDL-cholesterol, and TG. A total of 2,209 individuals distributed over 507 families of Northern European descent qualified for the above-mentioned criteria and thus formed the initial study population, of which 1,560 individuals from 261 families that contributed to the linkage on chromosome 7q36 (family LOD >0) were selected for further studies reported here.

Confirmatory studies were performed from a cohort of unrelated individuals of Northern European descent recruited by the same criteria from the same geographic regions. Samples from 920 unrelated individuals were available for this study. All protocols were approved by the Institutional Review Board of the Medical College of Wisconsin, and all participants signed an informed consent. Details on the two cohorts used in this study are included in Table 1.

Measurements

Genotyping. Single nucleotide polymorphisms (SNPs) were genotyped using Invader technology as described for previous studies (27–30). Genotyping was carried out in a total volume of 6 µl, containing 0.5 µl of PCR product, 0.02 µl of each primary probe, 0.002 µl of Invader probe, 2.75 µl TE buffer, 0.35 µl Cleavase (TWT), and 1.24 µl FRET mix (TWT). Reactions were denatured at 95°C for 5 min and then heated at 65°C for 15–120 min. Additional SNPs were genotyped on an Affymetrix MegAllele custom-designed 5K array using molecular inversion probe technology, as described previously (28). The location of all SNPs included in the analysis is shown in Fig. 1.

Sequencing and SNP identification. A 4,139 bp region was sequenced in 47 unrelated individuals selected from the MRC-OB cohort. The region was amplified by PCR using 21 overlapping amplicons 261–444 bp in size. PCR products were purified using MultiScreen-FB plates (Millipore) and eluted in sterile water for sequencing. Sequencing was carried out using BigDye v3.1 chemistry (Applied Biosystems, Foster City, CA) with standard

| TABLE 1. Summary statistics for the cohorts used in this study, the original 1,560 individuals of the MRC-OB family cohort and 920 unrelated individuals of the replication cohort |
|----------------|----------------|----------------|----------------|
|                | MRC-OB Family Cohort | Replication Cohort |
|                 | Average | Range     | Average | Range     |
| n                | 1,560   | 929       |
| % female         | 70%     | 80%       |
| Age (years)      | 46, 18–90 | 54, 18–89 |
| BMI (kg/m2)      | 31.9, 17.1–75.3 | 32.6, 15.9–86.1 |
| TGs (mg/dl)      | 128.0, 27–457 | 135.8, 36–424 |
| LDL-C (mg/dl)    | 106.8, 15–245 | 131.0, 33–251 |
techniques and analyzed on an Applied Biosystems 3730xl DNA Analyzer. SNP loci were identified by aligning all sequences using POLYPHRED (31), and all putative sites were checked manually for confirmation. Results were independently confirmed by aligning all traces with the chromosome 7 reference sequence using the anchored alignment algorithm in POLYBAVES (32) and using POLYPHRED to screen for SNPs.

Expression profiling. RNA isolation, anti-sense RNA generation, and Illumina BeadChip hybridization of the liver samples as well as transcript detection and normalization were performed as previously described (33). Briefly, a liver biopsy was obtained from each of 73 morbidly obese subjects of Northern European descent undergoing bariatric surgery. RNA was extracted and analyzed using Illumina Human-6 Expression BeadChips, which allow interrogation of almost 48,000 transcripts. Gene expression data for all liver samples were analyzed according to Goring et al. (33). To identify transcripts with detectable quantitative expression in the samples, distribution of expression values for a given transcript were compared with those of the controls imbedded in each chip. All significantly expressed transcripts were identified using an FDR of 0.05. To minimize the influence of overall signal levels, which may reflect RNA quantity and quality rather than a true biological difference between individuals, abundance values of all transcripts were standardized by z-scoring within individuals (using decile percent-age bins of transcripts, grouped by average log-transformed raw signals across individuals), followed by linear regression against the individual-specific average log-transformed raw signal and its squared value, as discussed by Goring et al. (33).

For the analysis of allele-specific expression differences in INSIG1, individuals were grouped according to genotype (for rs2721), and their gene-specific normalized expression z-scores were compared for genotype-specific differences using ANOVA. Differences between individual genotypes were compared using t-tests.

EMSA. The human HepG2 cell line was maintained in DMEM, supplemented with 2 mM L-glutamine, 100 µg/ml each of streptomycin and penicillin, and 15% fetal bovine serum, at 37°C with 5% CO2. Complementary oligonucleotides representing both allelic forms of the INSIG1 promoter SNP (rs2721, G>T at −2253 to the start codon) were obtained commercially (Integrated DNA Technologies) and purified by HPLC. Electrophoretic mobility shift assay (EMSA) was performed as described previously (34). Briefly, nuclear extracts were prepared from approximately 8 × 10^6 cells. Extracts were frozen in liquid N2 and stored at −80°C. For EMSA, nuclear proteins (3 µg) were preincubated for 10 min on ice with 1 µg of poly (dI-dC) (Pharmacia) in a binding buffer [4% Ficoll, 20 mM HEPES (pH 7.9), 1 mM EDTA, 1 mM DTT, 50 mM KCl]. Nuclear proteins were then incubated with biotin-labeled double-stranded oligonucleotide probes (20 fmol) for 30 min on ice and then analyzed on 6% Novex® DNA retardation gels (Invitrogen, CA) and electroblotted onto a positively charged nylon membrane (Ambion, Novex® DNA retardation gels (Invitrogen, CA) and electrophoresis. Probes (20 fmol) were incubated for 30 min on ice and then analyzed on 6% Novex® DNA retardation gels (Invitrogen, CA) and electroblotted onto a positively charged nylon membrane (Ambion, Novex® DNA retardation gels (Invitrogen, CA) and electroblotted onto a positively charged nylon membrane (Ambion, Novex® DNA retardation gels (Invitrogen, CA) and electroblotted onto a positively charged nylon membrane (Ambion, Novex® DNA retardation gels (Invitrogen, CA) and electroblotted onto a positively charged nylon membrane (Ambion, Novex® DNA retardation gels (Invitrogen, CA) and electroblotted onto a positively charged nylon membrane (Ambion, NO

RESULTS

Association of INSIG1 sequence variants with TG in the MRC-OB cohort

The QTL on chromosome 7q35-q36 linked to plasma TG levels has been replicated in multiple cohorts (10–15). Among the 22 identified genes in this QTL, INSIG1 is the only candidate with known biological roles that are related to lipid metabolism and therefore has been the focus of this study.

Initially, we tested for INSIG1 association using 11 SNPs selected from the Hapmap database using the tagger function of Haploview (35, 36) (Fig. 1 and Table 2). These SNPs spanned the entire INSIG1 gene region (39.7 kb).
One variant in the putative promoter region of *INSIG1*, rs2721, was significantly associated with plasma TG levels in 1,560 individuals from our cohort (\(P = 0.002\)). Two additional SNPs (rs9770068 and rs9769506) located in intron 6 of *INSIG1* showed marginal association (\(P = 0.047\) and \(P = 0.037\), respectively), as well as one additional SNP in the promoter region (rs9690040, \(P = 0.045\)). As these three marginally associated SNPs are in almost complete linkage disequilibrium with rs2721 (Fig. 1), we speculate that their high degree of correlation may have led to the observed associations.

The promoter variant (rs2721) is located 2,253 bp upstream of the transcription start site of *INSIG1*. To include all other variants surrounding rs2721 in the association analysis, we resequenced the promoter region and exon 1 in 47 unrelated individuals from our cohort. SNPs in the remaining exons of the gene were in high linkage disequilibrium (LD) with each other and not in LD with rs2721; thus, we focused solely on the region surrounding rs2721. All selected individuals were parents from families contributing significantly to the initially observed linkage on chromosome 7 in the MRC-OB cohort. The resequencing effort uncovered 17 additional SNPs, of which 8 were novel and not reported in dbSNP. The location of all SNPs is shown in Fig. 1.

All 28 SNPs were genotyped in the 1,560 individuals from the MRC-OB cohort. The results of the association analysis are summarized in Table 2. No additional SNPs were associated with plasma TG levels in this cohort, suggesting that the effect is primarily mediated by rs2721. The TG levels of individuals homozygous for the G allele (123.1 ± 2.3 mg/dl, \(n = 1073\)) were on average 9% lower than those of individuals homozygous for the T allele (135.9 ± 16.2 mg/dl, \(n = 31\)).

Next, we examined the effect of the SNP on the initial linkage. Overall, the QTL interval on chromosome 7q36 explains 31% of the variance in TG levels in our cohort. When including SNP rs2721 in *INSIG1* as a covariate, the linkage drops 5.4% for our family-based cohort, suggesting a modest but significant effect on the initial linkage on

Fig. 1. Schematic illustration of the *INSIG1* gene region. All SNPs investigated across the gene region (Chr7:154,696,933 bp–154,736,613 bp) are shown relative to the *INSIG1* gene, depicted by the black line. Translated exons are shown by white boxes, untranslated exonic regions are indicated by gray boxes. The associated promoter SNP rs2721 is highlighted by the red box. The colored diagram depicts the pairwise linkage disequilibrium between SNPs (\(D'\), as calculated for the family-based cohort of 1,560 individuals using the program Haploview.
INSIG1 affects human triglyceride levels

All SNPs and their positions are listed, together with p-values for association with plasma TG levels in the MRC-OB family cohort and the replication cohort. SNPs shown in italics were used in the initial association analysis. The results for SNP rs2721 are highlighted by the box.

Table 2. Summary of SNP association results

| SNP   | Position | MRC-OB Family Cohort | P  | Replication cohort | MAF | P  |
|-------|----------|----------------------|----|--------------------|-----|----|
| rs6597489 | 154180935  | 0.206                | 0.372 | 0.0008              | 0.15 |
| rs10901042 | 154097319  | 0.184                | 0.371 |                    |      |    |
| rs6915390 | 154170180  | 0.250                | 0.162 |                    |      |    |
| rs9801134 | 154170357  | 0.126                | 0.028 |                    |      |    |
| rs9801301 | 154099667  | 0.516                | 0.462 |                    |      |    |
| rs6909400 | 154170649  | 0.045                | 0.372 |                    |      |    |
| rs6901897 | 154178125  | 0.255                | 0.434 |                    |      |    |
| SNP1    | 154178324  | 0.127                | 0.018 |                    |      |    |
| rs2721  | 154178767  | 0.002                | 0.145 |                    |      |    |
| rs1128640 | 154178751  | 0.542                | 0.114 |                    |      |    |
| rs10271719 | 154179098  | 0.284                | 0.129 |                    |      |    |
| rs12154226 | 154179026  | 0.146                | 0.425 |                    |      |    |
| rs12154352 | 154179048  | 0.165                | 0.423 |                    |      |    |
| rs12154437 | 154179051  | 0.153                | 0.423 |                    |      |    |
| SNP2    | 154179102  | 0.146                | 0.427 |                    |      |    |
| SNP3    | 154179128  | 0.119                | 0.422 |                    |      |    |
| SNP4    | 154179186  | 0.111                | 0.423 |                    |      |    |
| rs1128636 | 154179304  | 0.098                | 0.42  |                    |      |    |
| rs6902271 | 154179881  | 0.185                | 0.423 |                    |      |    |
| SNP5    | 154179345  | 0.464                | 0.078 |                    |      |    |
| SNP6    | 154179202  | 0.464                | 0.001 |                    |      |    |
| SNP7    | 154179230  | 0.207                | 0.001 |                    |      |    |
| SNP8    | 154179254  | 0.202                | 0.133 |                    |      |    |
| rs9767875 | 154179399  | 0.535                | 0.132 |                    |      |    |
| rs9768687 | 154179437  | 0.148                | 0.117 |                    |      |    |
| rs9779063 | 154179695  | 0.047                | 0.363 |                    |      |    |
| rs9765906 | 154179788  | 0.037                | 0.421 |                    |      |    |
| rs12391375 | 154179915 | 0.151                | 0.255 |                    |      |    |

Interaction between INSIG1 and INSIG2

In humans, the INSIG1 protein is 78% similar to its only homolog, INSIG2. Interestingly, while INSIG1 is involved in the feedback regulation of lipid synthesis, INSIG2 has been shown to be associated with obesity and adipocyte metabolism in a number of recent studies (41–43). However, no study to date has addressed whether these two genes work cooperatively in humans to affect lipid levels, as has been shown in mice (26).

Correlation analysis based on family-specific LOD scores of the locus of INSIG1 and the locus of INSIG2 suggested a nonadditive interaction between the two genes (p = 0.239).

Fig. 2. EMSA of rs2721 using nuclear extract of liver HepG2 cells. Oligos of DNA sequences from INSIG1 promoter region centered with either allele of rs2721 (T or G) were incubated with total nuclear proteins from cultured liver cells. Arrows point at mobility shift of oligonucleotides with the T allele caused by differential protein binding. Results are representative of at least two independent experiments.
In INSIG2 (rs2422166 and rs7566605) that have been repeatedly shown to be associated with BMI and LDL-C in other studies (41–43). Gene-gene interaction analysis was performed on the two INSIG2 SNPs and rs2721 of INSIG1. In this analysis, SNP rs2422166 did not show any effect on TG levels directly or through interaction with rs2721. In contrast, SNP rs7566605 showed marginal interaction with rs2721 (P = 0.0473) although it was not directly associated with TG levels. To measure the impact of genotype at rs7566605 on the association of rs2721 with lipid traits, we grouped the data based on genotype at rs7566605. Measured genotype analysis was then performed in this subset (Table 3). In the GG subgroup, the associations with TG levels disappeared. In contrast, in the CC-CG subgroup, the association was significantly strengthened (β = 0.13 ± 0.04, P = 0.00117). Here, individuals homozygous for the T allele of rs2721 had 31% higher TG levels than individuals homozygous for the G allele (138.4 ± 1.9 mg/dl vs. 105.6 ± 1.7 mg/dl). Before we stratified the study cohort based on genotypes of INSIG2 rs7566605, the difference between the INSIG1 rs2721 homozygous groups was only 8%.

**TABLE 3. Interaction analysis of rs7566605 in INSIG2 and rs2721 in INSIG1.** The C-allele of rs7566605 significantly enhances the association of rs2721 with plasma TG levels

| rs7566605 | rs2721 | TG (mg/dl) |
|-----------|--------|------------|
| CC and CG | GG     | 105.6 ± 1.7|
| (n = 763) | GT     | 117.9 ± 1.8|
|          | TT     | 138.4 ± 1.9*|
| GG       | GG     | 107.8 ± 1.7|
| (n = 768) | GT     | 108.9 ± 1.8|
|          | TT     | 98.5 ± 1.6  |

* = significant difference when compared with GG of rs2721 (P = 0.00117).

**DISCUSSION**

Studies conducted in animals have revealed intriguing and critical roles of INSIG genes in lipid homeostasis (26). By anchoring a sterol-sensing retainer protein that is able to block the access to the nucleus of a number of transcription factors, INSIG proteins have been recognized as key regulators of both cholesterol and fatty acid synthesis (18, 19, 21). No study, however, has shown evidence of their contrast, SNP rs7566605 showed marginal interaction with rs2721. In the GG subgroup, the associations with TG levels disappeared. However, our analysis also suggests that this variant accounts for only a portion of the overall linkage (~5%), suggesting that other variants and genes in the QTL interval also affect plasma TG levels. Overall, variance at rs2721 explains 1.6% of the observed variance in TG levels in our family cohort. This contribution of this polymorphism, however, may be underestimated, because the interaction with other genes and beyond INSIG2 locus have not been considered in this study.

The SNP rs2721 resides 2,253 bp upstream of the start codon of INSIG1 in the putative promoter region. To test the genotype-specific effect on the transcription of INSIG1 in humans, expression profiling was conducted in liver tissue, the putative target tissue for INSIG1 action on lipid homeostasis in mice. Our analysis revealed that individuals homozygous for the T allele had significantly lower INSIG1 expression levels, suggesting that reduced expression of INSIG1 increases circulating plasma TG levels. While the expression analysis revealed only two individuals homozygous for the T allele in our sample set of 73 individuals, the data nonetheless provide suggestive evidence for an effect of this promoter variant on gene repression in vivo. Additional samples and tissues will need to be investigated to further substantiate this effect of rs2721.

To further validate this potential effect of SNP rs2721 on the promoter function of INSIG1, we examined the binding of nuclear proteins to the sequence surrounding rs2721 by EMSA. Separate custom oligonucleotides specific for the alleles of rs2721 and comprising the immediate sequence surrounding the SNP location were designed. Oligonucleotides specific for the T allele were able to bind nuclear proteins extracted from HepG2 cells, in contrast to the G allele (Fig. 2). These data demonstrate that the T allele of rs2721 shows differential binding of nuclear proteins in vitro, further supporting the hypothesis that this sequence variant affects promoter function of INSIG1. Analysis of the sequence using TFBIND (http://tfbind.imis.u-tokyo.ac.jp/) reveals that rs2721 is located in a putative binding site for the transcription factor PreB-cell leukemia homeobox 1 (PBX1), which is expressed in the liver and a wide variety of other tissues. The binding site is altered in the G allele, reducing the probability of PBX1 binding more than 6-fold over the T allele. Additional transcription factors with differential binding affinity to the two alleles of rs2721 include Forkhead transcription factor and Caudal type homeobox A. This differential binding affinity may explain the observed EMSA results and may be related to the observed allele-specific expression difference, although none of the transcription factors have been reported to modulate plasma TG levels. Nonetheless, PBX1 has been reported to affect glucose tolerance and steroidogenesis, and it has been suggested that mutations in PBX1 may promote susceptibility to diabetes (44). However, at this point, the specific proteins mediating this effect (binding to the T allele, reducing expression of INSIG1, and thereby elevating plasma TG levels)
remain to be identified. While the EMSA results, in conjunction with the expression analysis of the liver biopsy samples, suggest a functional mechanism for rs2721, it is clear that EMSA results do not prove functional interactions, and additional molecular studies using transfection assays may be required to confirm the initial findings.

It has been shown in animal models that double knock-out animals for *INSIG1* and its only homolog *INSIG2* have more significant lipid abnormalities compared with knock-out animals for either gene individually (26). We therefore explored the putative interactive effect of sequence variants in both human *INSIG1* and *INSIG2* on plasma lipids. By focusing on the functional SNP rs2721 of *INSIG1*, we tested the gene-gene interaction with two *INSIG2* variants that have shown repeatedly to be associated with BMI and lipid abnormalities in obesity studies. Based on our analysis, *INSIG1* does interact with *INSIG2* in relation to lipid homeostasis. The two *INSIG2* SNPs have no independent effect on plasma TG levels in our cohort. However, rs7566605 significantly enhances the effect seen with rs2721 of *INSIG1*, and their interaction strengthened the association with plasma TG levels (Table 3). Specifically, people possessing the GG genotype at rs7566605 of *INSIG2* no longer showed a detectable effect of the TT genotype of rs2721 in *INSIG1* on TG levels. In contrast, individuals with at least one C allele at rs7566605 of *INSIG2* showed an enhanced effect of the T allele of rs2721 on TG. These data suggest that *INSIG2* (or at least the G allele of rs7566605) may be able to compensate for the effect of rs2721 in *INSIG1*. While the exact mechanism remains to be elucidated, our analysis overall clearly suggests that rs2721 affects plasma TG levels by lowering *INSIG1* gene expression in the liver. This effect can be reduced or enhanced by the different alleles of rs7566605 in *INSIG2*.

Clinically, a reduction of 10 mg/dl in plasma TG levels, independently of the impact of treatment on LDL-C levels, resulted in a 1.4% reduction in risk of developing coronary heart disease in a report by Miller M et al. (45). In our study, the elevation in TG levels in individuals with the TT genotype of rs2721 was 12.8 mg/dl and 14.4 mg/dl for families and unrelated individuals, respectively, when the population was not stratified by genotypes of *INSIG2* at rs7566605. This effect is further exacerbated by the cooccurrence of the C allele of rs7566605 in *INSIG2* and the TT allele of rs2721, resulting in a 32.8 mg/dl elevation in TG levels.

Despite the replication of the linkage finding on chromosome 7q36, all genome-wide association studies (GWAS) to date exploring lipid traits have failed to detect any significant association in this genomic interval. Neither *INSIG1* nor any other gene in the QTL region has been implicated in any of the lipid trait-related GWAS efforts. Given the low linkage disequilibrium between rs2721 and any of the adjacent SNPs on common GWAS platforms, it is not surprising why the association of rs2721 with plasma TG levels detected in our analysis was not uncovered by GWAS. However, the complete lack of association evidence for the entire genomic interval, despite repeated linkage evidence, needs to be investigated further. This effect has been reported before and is not specific to this QTL region or lipid traits (46).

The lack of comparable findings in other studies may also be due to differences in the recruitment of the study cohorts. In our study design, we excluded individuals with hepatic disease, severe coronary artery disease, dramatic weight loss in the preceding 12 months, or individuals taking lipid-lowering medications. Given the suggested function of the *INSIG* genes, these criteria may themselves be associated with *INSIG* function and thus may have a potential influence on the results. Likewise, it remains to be seen whether the expression profiles obtained from liver samples collected from bariatric surgery patients reflect the levels of *INSIG1* expression in the general population.

While *INSIG1* is not the only gene and rs2721 not the only sequence variant responsible for the initially observed linkage on human chromosome 7q36, our analysis, for the first time, demonstrates a significant role for this gene in the control and regulation of plasma TG levels and suggests a functional mechanism by which the causal variant rs2721 may exert its effect. Further QTL-wide analyses will be required to elucidate the other remaining genetic factors affecting plasma TG levels in our cohort and validate the effect on this important cardiovascular disease risk factor in other patient cohorts.

**REFERENCES**

1. Forrester, J. S. 2001. Triglycerides: risk factor or fellow traveler? *Curr. Opin. Cardiol.* 16: 261–264.

2. Malloy, M. J., and J. P. Kane. 2001. A risk factor for atherosclerosis: triglyceride-rich lipoproteins. *Adv. Intern. Med.* 47: 111–136.

3. Connelly, P. W., A. Petrasovits, S. Stachenko, D. R. MacLean, J. A. Little, and A. Chockalingam. 1999. Prevalence of high plasma triglyceride combined with low HDL-C levels and its association with smoking, hypertension, obesity, diabetes, sedentariness and LDL-C levels in the Canadian population. Canadian Heart Health Surveys Research Group. *Can. J. Cardiol.* 15: 428–433.

4. Tai, E. S., S. C. Emmanuel, S. K. Chew, B. Y. Tan, and C. E. Tan. 1999. Isolated low HDL cholesterol: an insulin-resistant state only in the presence of fasting hypertriglyceridaemia. *Diabetes.* 48: 1088–1092.

5. Austin, M. A., M. C. King, R. D. Bawol, S. B. Hulley, and G. D. Friedman. 1987. Risk factors for coronary heart disease in adult female twins. Genetic heritability and shared environmental influences. *Am. J. Epidemiol.* 125: 308–318.

6. Heller, D. A., U. de Faire, N. L. Pedersen, G. E. Dahlen, and G. E. Sonnenberg. 1999. Genetic and environmental influences on serum lipid levels in twins. *N. Engl. J. Med.* 329: 1150–1156.

7. Perusse, L., T. Rice, J. P. Despres, J. Bergeron, M. A. Province, J. Gagnon, A. S. Leon, D. C. Rao, J. S. Skinner, J. H. Wilmore, et al. 1997. Familial resemblance of plasma lipids, lipoproteins and postheparin lipoprotein and hepatic lipases in the HERITAGE Family Study. *Arterioscler. Thromb. Vasc. Biol.* 17: 3263–3269.

8. Kissebah, A. H., S. Alfarsi, and P. W. Adams. 1981. Integrated regulation of very low density lipoprotein triglyceride and apolipoprotein-B kinetics in man: normalipemic subjects, familial hypertriglyceridaemia and familial combined hyperlipidaemia. *Metabolism.* 30: 856–868.

9. Kissebah, A. H., G. E. Sonnenberg, J. Myklebust, M. Goldstein, K. Broman, R. G. James, J. A. Marks, G. R. Krakower, H. J. Jacob, J. Weber, et al. 2000. Quantitative trait loci on chromosomes 3 and 17 influence phenotypes of the metabolic syndrome. *Proc. Natl. Acad. Sci. USA.* 97: 14478–14483.

10. Sonnenberg, G. E., G. R. Krakower, L. J. Martin, M. Olivier, A. E. Koitek, A. G. Comuzzie, J. Blangero, and A. H. Kissebah. 2004. Genetic determinants of obesity-related lipid traits. *J. Lipid Res.* 45: 610–615.
11. Li, W. D., C. Dong, D. Li, C. Garrigan, and R. A. Price. 2005. A genome scan for serum triglyceride in obese nuclear families. *J. Lipid Res.* 46: 432–438.

12. Duggirala, R., J. Blangero, L. Almasy, T. D. Dyer, K. L. Williams, R. J. Leach, P. O’Connell, and M. P. Stern. 2000. A major susceptibility locus influencing plasma triglyceride concentrations is located on chromosome 1q24 in Mexican Americans. *Am. J. Hum. Genet.* 66: 1237–1247.

13. Horne, B. D., A. Malhotra, and N. J. Camp. 2003. Comparison of linkage analysis methods for genome-wide scanning of extended pedigrees, with application to the TG/HDL-C ratio in the Framingham Heart Study. *BMC Genet.* 4 (Suppl. 1): S93.

14. Lin, J. P. 2003. Genome-wide scan on plasma triglyceride and high density lipoprotein cholesterol levels, accounting for the effects of correlated quantitative phenotypes. *BMC Genet.* 4 (Suppl. 1): S47.

15. Shearman, A. M., J. M. Ordovas, L. A. Cupples, E. J. Schaefer, M. D. Harmon, Y. Shao, J. D. Keen, A. L. DeStefano, O. Joost, P. W. Wilson, et al. 2000. Evidence for a gene influencing the TG/HDL-C ratio on chromosome 7q23.3-qter: a genome-wide scan in the Framingham study. *Hum. Mol. Genet.* 9: 1315–1320.

16. Diamond, R. H., K. Du, V. M. Lee, K. L. Mohn, B. A. Haber, D. S. Tewari, and R. Taub. 1993. Novel delayed-early and highly insulin-induced growth response genes. Identification of HRS, a potential regulator of alternative pre-mRNA splicing. *J. Biol. Chem.* 268: 15185–15192.

17. Li, W. D., C. Dong, D. Li, C. Garrigan, and R. A. Price. 2005. A genome scan for serum triglyceride in obese nuclear families. *J. Lipid Res.* 46: 432–438.