Original Article

Common Hepatic Nuclear Factor-4α Variants Are Associated With High Serum Lipid Levels and the Metabolic Syndrome

Daphna Weissglas-Volkov,[1] Adriana Huertas-Vazquez,[1] Elina Suviolahti,[1] Jenny Lee,[1] Christopher Plaisier,[1] Samuel Canizales-Quinteros,[2] Teresa Tusie-Luna,[2] Carlos Aguilar-Salinas,[3] Marja-Riitta Taskinen,[4] and Päivi Pajukanta[1]

Hepatic nuclear factor-4α (HNF-4α), a transcription factor involved in the regulation of serum lipid and glucose levels, has recently been associated with type 2 diabetes. The HNF-4α gene (HNF4A) resides on chromosome 20q12-q13.1, which, in addition to type 2 diabetes, has also previously been linked to high triglycerides in Finnish familial combined hyperlipidemia (FCHL) families. FCHL, characterized by elevated levels of serum total cholesterol, triglycerides, or both, is a common dyslipidemia observed in up to 20% of patients with premature coronary heart disease. Considering the clear phenotypic overlap between type 2 diabetes and FCHL, both predisposing to high serum triglycerides and glucose intolerance, we tested this gene for association in dyslipidemic families originating from two distinct populations, Finnish and Mexican, and comprising 1,447 subjects. Our data show that common HNF4A variants and haplotypes are associated with elevated serum lipid levels and the metabolic syndrome ($P = 0.008–0.04$), as well as with elevated glucose parameters ($P = 0.008–0.03$), using family-based association analysis. Importantly, both Finnish and Mexican families shared two common lipid-associated HNF4A haplotypes ($P = 0.005$ for total cholesterol and 0.006 for triglycerides). In conclusion, we show for the first time that common HNF4A variants are associated with high serum lipid levels and the metabolic syndrome. *Diabetes* 55:1970–1977, 2006

From the [1]Department of Human Genetics, David Geffen School of Medicine at UCLA, University of California Los Angeles, Los Angeles, California; the [2]Department of Molecular Biology and Genomic Medicine, Instituto de Investigaciones Biomédicas de la Universidad National Autónoma de Mexico, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán (INCMNSZ), Mexico City, Mexico; the [3]Department of Endocrinology and Metabolism, INCMNSZ, Mexico City, Mexico; and the [4]Department of Medicine, Helsinki University Central Hospital, Helsinki, Finland.

Address correspondence and reprint requests to Päivi Pajukanta, MD, PhD, Assistant Professor, UCLA, Department of Human Genetics, Gonda Neuroscience and Genetics Research Center, 695 Charles E. Young Dr. South, Los Angeles, CA 90095-7088. E-mail: ppajukanta@mednet.ucla.edu.

Received for publication 6 January 2006 and accepted in revised form 24 March 2006.

D.W.-V. and A.H.-V. contributed equally to this work.

M.-R.T. has received honoraria and consulting fees from Merck Sharp and Dome, Pfizer, Laboratoires Fournier, GlaxoSmithKline, Bristol-Myers Squibb, Takeda, and Sanofi Aventis and research support from Laboratoires Fournier, Elli Lilly, and Novartis.

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

apo, apolipoprotein; AUC, area under the curve; CHD, coronary heart disease; FBAT, family-based association test; HBAT, haplotype-based association test; FCHL, familial combined hyperlipidemia; HNF-4α, hepatic nuclear factor-4α; IGT, impaired glucose tolerance; OGTT, oral glucose tolerance test; SNP, single nucleotide polymorphism; USF1, upstream transcription factor 1.

DOI: 10.2337/db06-0035

Recently, a number of independent groups identified associations between single nucleotide polymorphisms (SNPs) in the hepatic nuclear factor-4α (HNF-4α) gene HNF4A on 20q12-q13.1 and type 2 diabetes in several populations (1–3), including Finns (4,5). HNF4A encodes a transcription factor, and mutations in HNF4A are known to cause MODY1 (maturity-onset diabetes of the young type 1), an early-onset autosomal dominant form of type 2 diabetes (6). HNF-4α is expressed in various tissues, predominantly in liver and pancreas (7). Transcription of the gene is driven from two different promoters (8), and there are numerous alternative splice forms (9).

Familial combined hyperlipidemia (FCHL) is a common dyslipidemia affecting 1–2% of Western populations (10) and 10–20% of subjects with premature coronary heart disease (CHD) (11). FCHL is characterized by elevated levels of serum total cholesterol, triglycerides, or both. FCHL also expresses low HDL cholesterol, elevated levels of serum apolipoprotein (apo)B, and glucose intolerance (12) as component traits (11,13).

We previously found suggestive evidence for linkage in both Finnish FCHL and low–HDL cholesterol families for a locus on chromosome 20q13.11, using the high-triglyceride and low–HDL cholesterol traits (14–16). Subsequently, evidence for linkage was found with high triglycerides in other populations for this locus (17,18). This region has also been linked to type 2 diabetes and obesity in numerous studies (19–24). There is a clear phenotypic overlap between FCHL and type 2 diabetes because patients with FCHL are often glucose intolerant and/or insulin resistant, and patients with type 2 diabetes are often hypertriglyceridemic (12,25). In addition, both diseases predispose to CHD.

It has been suggested that HNF-4α is a central regulator of glucose and lipid metabolism (26). However, no association between plasma lipid levels and HNF4A has been identified in families with FCHL or low HDL cholesterol. Considering the clear phenotypic overlap between type 2 diabetes and FCHL, it is possible that this gene contributes to both linkage results. We tested this gene region for association in Finnish and Mexican dyslipidemic families. This is the first study showing that common HNF4A

variants are associated with high plasma lipid levels and the metabolic syndrome.

RESEARCH DESIGN AND METHODS

A total of 99 Finnish dyslipidemic families (60 FCHL and 39 low–HDL cholesterol families) were recruited in the Helsinki and Turku University Central Hospitals (14,15,27). All study subjects gave their informed consent. The study design was approved by the ethics committees of the participating centers. The inclusion criteria for FCHL and low–HDL cholesterol families were premature CHD and abnormal lipid profile with total cholesterol and/or triglyceride levels greater than or equal to the age/sex-specific 90th Finnish population percentile for FCHL probands and HDL cholesterol levels less than or equal to the age/sex-specific 10th percentile for HDL cholesterol probands (14,15). Additional lipid criteria for the low–HDL cholesterol probands were total cholesterol ≤6.3 mmol/l in men and ≤6.0 mmol/l in women and triglycerides ≤2.3 mmol/l in both sexes. Exclusion criteria for the FCHL probands were type 1 diabetes, hypothyreosis, or renal disease, and for the low–HDL cholesterol probands they were type 1 diabetes, type 2 diabetes, hepatic or renal disease, and BMI >30 kg/m^2.

Mexican FCHL families. A total of 24 extended Mexican FCHL families with a history of premature CHD were included in this study, comprising 314 family members. The race of all of these subjects is Mestizos. These families were recruited in the Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán (INCMNSZ) in Mexico City. The inclusion and exclusion criteria for the Mexican FCHL families (28) are fully comparable with the Finnish criteria for FCHL (14,27). The age/sex-specific 90th percentiles for total cholesterol and triglycerides, based on a previous survey of the Mexican population (29), were used to classify the FCHL affection status. Each subject provided written informed consent. The study design was approved by the committee of biomedical research in humans of the INCMNSZ.

Biochemical analyses of Finnish and Mexican families. In the Finnish families, serum lipids, glucose parameters, and oral glucose tolerance tests (OGTTs) were performed as previously described (12,14,16). Patients who used lipid-lowering drugs were studied after their lipid-lowering treatment was withdrawn for 4 weeks. In the Mexican families, all of the measurements were performed by the INCMNSZ with commercially available standardized methods (28).

Phenotypic traits related to type 2 diabetes, impaired glucose tolerance, and the metabolic syndrome. In the Finnish families, classification of impaired glucose tolerance (IGT) was performed, using World Health Organization criteria (30). The subjects were classified as IGT affected if their serum glucose values at 2 h post–75-g glucose challenge were between 6.7 and 10.0 mmol/l (12). Because type 2 and type 1 diabetes were exclusion criteria for the low–HDL cholesterol probands, resulting in only 46 IGT-affected individuals in these families, we did not analyze this phenotypic trait in the low–HDL cholesterol families. The classification of the metabolic syndrome was performed in the Finnish FCHL and low–HDL cholesterol families, using National Cholesterol Education Program Adult Treatment Panel III criteria (31). In the Mexicans families, we were not able to classify the IGT or metabolic syndrome status because of a lack of some of the phenotypic data required to classify these traits.

Genotyping and sequencing. SNPs were genotyped, using the pyrosequencing technique (Pyrosequencing, Foxboro, MA). Primers were designed for PCR, using the Primer3 program, and for detection, using SNP Primer Design software (Pyrosequencing). Sequencing was performed, using an ABI 377XL automated DNA sequencer (Applied Biosystems, Foster City, CA).

All selected SNPs were first tested for association in 42 Finnish FCHL families with 230 genotyped individuals. The five implicated SNPs were further tested for association in an extended sample of 60 Finnish FCHL families with 706 genotyped individuals. This sample of 60 FCHL families consisted of the 42 FCHL families where genotyping was extended to all available family members, as well as 18 additional FCHL families. The same inclusion and exclusion criteria were used to collect all 60 FCHL families. The five implicated SNPs were also genotyped in an independent sample of 39 Finnish low–HDL cholesterol families with 427 genotyped individuals and 24 Mexican FCHL families with 314 genotyped individuals, resulting in a total of 1,447 genotyped subjects. This strategy aimed to minimize multiple testing and verify the initial signals in several sample sets.

For quality control, we replicated 3.5% of genotyped samples. The percentage agreement between samples was >99%. The genotype call rate was 93–99% and genotype error rate 0.0–0.6%. The genotypes of the individuals with Mendelian errors were called as 0 in the analyses. All SNPs were in Hardy-Weinberg equilibrium.

Statistical analysis. Association analyses of qualitative traits for each individual SNP were performed, using the HHRR (haplotype-based haplotype relative risk) program (32). For quantitative traits, association analysis was performed, using the family-based association test (FBAT) program (33). The haplotypes were tested, using the haplotype-based association test (HBAT) program (33). The option "–e" of HBAT was used in the Finnish families because it is a test of association given linkage. In the relatively small Mexican FCHL sample, we also used the option "–o" of HBAT to assess not only preferential transmission of susceptibility haplotypes to affected individuals, but also less preferential transmissions to unaffected individuals. We also used the "–p" option of HBAT, which performs the Monte-Carlo permutation procedure to correct for multiple testing. Allele frequencies were estimated, using the DOWNFREQ program (34). The PedCheck program was used to assess pedigree inconsistencies (35). Pairwise linkage disequilibrium was tested, using the JLIN (Java LINkage disequilibrium plotter) program (available at www.genepi.com.au/project/jlin). For the subset analyses, the MLINK program (36) was used to perform the parametric two-point analysis, and the location score analysis of the SimWalk2 program (37) was used to perform the parametric multipoint analysis.

For the analysis of FCHL and low–HDL cholesterol Finnish families, triglycerides, total cholesterol, apoB, IGT, and the metabolic syndrome were tested as qualitative traits. Measurements of glucose, insulin, and C-peptide and the area under the curve were tested as quantitative traits. In Mexican FCHL families, three qualitative lipid traits were tested: triglycerides, total cholesterol, and apoB. However, no refined phenotypes for glucose metabolism were available for these families. The affection status for triglycerides, total cholesterol, and apoB traits in the family members was defined, using the 90th Finnish and Mexican age/sex-specific percentiles (14,26,28). The AUC was calculated, using the trapezoid rule (12). Log transformation was performed for the traits that were not normally distributed. The significance of age and/or sex as covariates was tested for all analyzed quantitative traits. The residuals were calculated, using only the significant one(s) of these two, and were used in the association analyses of quantitative traits.

Functional investigation of HNF4A-associated variants. The DIG gel shift kit, second generation, was used for sequence-specific DNA binding assays (Roche Applied Science, Indianapolis, IN). Double-stranded oligomers (25 bp) were designed to contain the variant of interest flanked by 12-bp sequences on both sides. All gel mobility shift experiments were repeated twice. Evolutionary conservation across species was evaluated, using the PhastCons program (available at www.genome.ucsc.edu/). The ElDorado tool (available at www.genomatix.de/products/ElDorado/) was used to predict a loss and/or gain of transcription factor binding sites caused by the investigated SNPs.

RESULTS

Association analyses of lipid phenotypes in Finns. To study whether variants within the HNF4A gene on 20q12-q13.1 are associated with high serum lipids, we first genotyped the SNP rs2144908, previously associated with type 2 diabetes (Fig. 1) (1,4), in 42 Finnish families with FCHL. We found evidence of association for high triglycerides (Table 1). Consequently, we examined this SNP in an extended sample of 60 Finnish families with FCHL and obtained a P value of 0.008 for high triglyceride levels (Table 1).

We subsequently genotyped nine additional SNPs (Fig. 1) selected because of their previous association with type 2 diabetes or based on the linkage disequilibrium structure in the region, which we constructed using both the HapMap data and recently published Finnish linkage disequilibrium data (4). These SNPs spanned a total of 77 kb (Fig. 1). The allele frequencies of the genotyped SNPs in probands and spouses are shown in Table 1 of the online appendix (available at http://diabetes.diabetesjournals.org). Although rs2144908 was the only single SNP showing evidence for association with triglycerides, we examined whether HNF4A haplotypes are associated with high serum lipids. Haplotype analysis revealed that a common haplotype of SNPs rs6031558-rs745975-rs3212198, haplotype H1A (1-1-2; 1 indicates the common allele), was also associated with high triglyceride levels (P = 0.008) (Table 2), whereas haplotype H1B (2-1-1) was implicated as a protective haplotype for high total cholesterol levels (Table 2). These SNPs represent three nonredundant SNPs

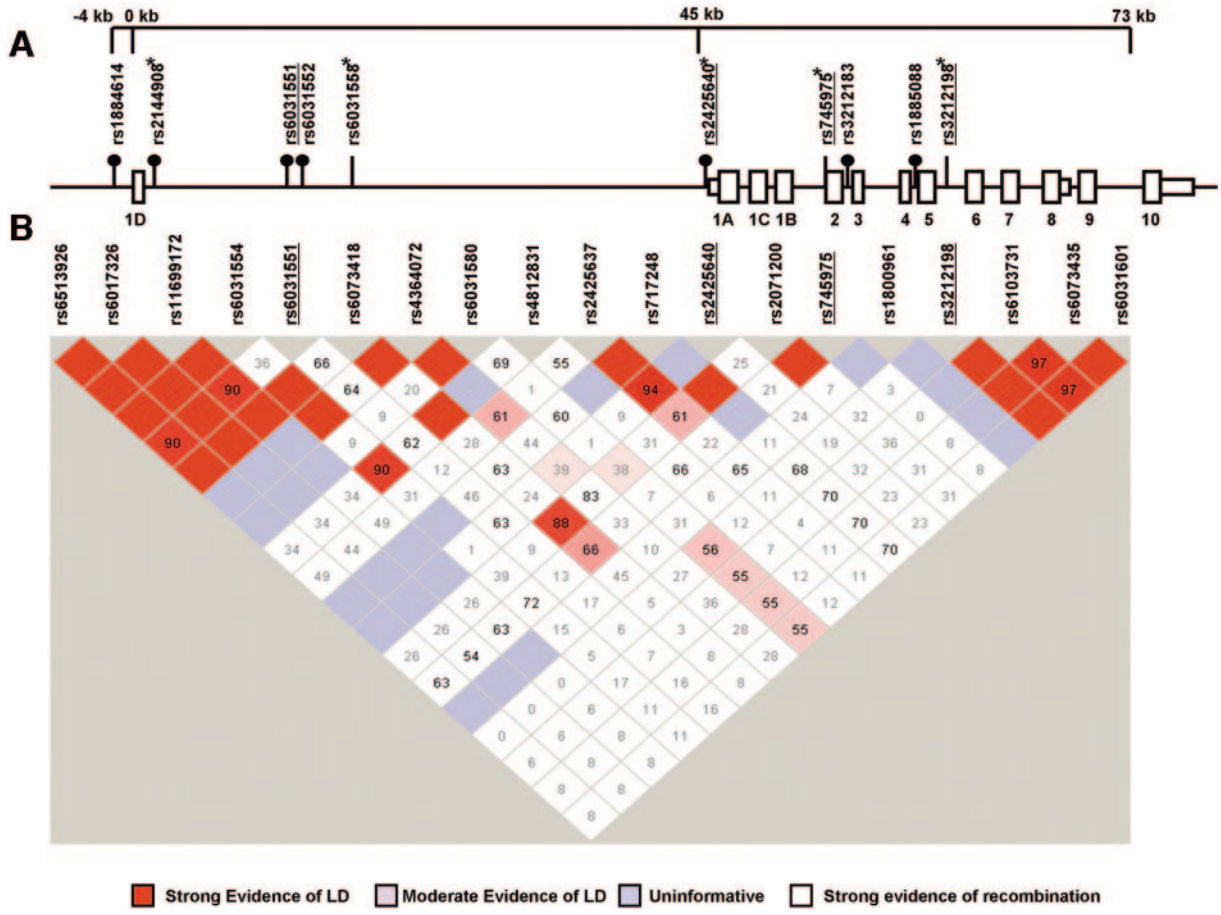

FIG. 1. HNF4A gene structure and linkage disequilibrium map. *A*: Positions of 10 SNPs within the HNF4A locus selected for analysis in the Finnish FCHL families. HapMap SNPs are underlined. ●, SNPs previously associated with type 2 diabetes disease status; *SNPs genotyped in both Finnish and Mexican FCHL families. *B*: Linkage disequilibrium (LD) plot across the HNF4A region. Linkage disequilibrium calculated using 30 CEPH trios with northern and western European ancestry genotyped by the HapMap project (release no. 16c.1/phaseI, June 2005). Linkage disequilibrium is displayed, using the Haploview program v3.11 (available at www.broad.mit.edu/mpg/haploview/).

because the linkage disequilibrium between them was <0.8 (D′) (Figs. 1 and 2). We also examined these haplotypes in an extended study sample of 60 Finnish families with FCHL. The inclusion of the additional families did not strengthen the result for high triglycerides; however, the *P* value for high total cholesterol levels improved slightly (Table 2). The apoB trait did not show evidence for association with single SNPs or haplotypes.

Association analyses of glucose phenotypes in Finns. Next, we examined whether the SNPs rs6031558-rs745975-rs3212198, forming the lipid-associated H1A and H1B haplotypes, were associated with glucose traits in the Finnish FCHL families. Subjects were classified for IGT, resulting in 132 IGT-affected individuals. We found evidence of association with a *P* value of 0.03 for a different risk haplotype (1-2-2). We also tested the SNPs and haplotypes for association with glucose values 2 h after an OGTT. As with IGT, we found evidence of association for increased glucose values in OGTT with a *P* value of 0.02 for the same risk haplotype (1-2-2). Furthermore, we found that SNP rs2425640, previously associated with type 2 diabetes (2,4), was individually associated with an increase of several glucose parameters (glucose AUC, insulin AUC, and C-peptide AUC; $P = 0.01$, 0.008, and 0.03,

TABLE 1
Association analysis of high triglycerides and apoB with the SNP rs2144908 in the Finnish and Mexican dyslipidemic study samples

Study sample	*n**	Major/minor allele†	Heterozygosity/minor allele frequency‡	Triglyceride *P* value	apoB *P* value
42 Finnish FCHL families	230, 117, 127	**G**/A	0.30/0.21	0.02	NS
60 Finnish FCHL families	706, 202, 257	**G**/A	0.31/0.20	0.008	NS
39 Finnish low–HDL cholesterol families	427, 49, 113	**G**/A	0.31/0.21	NS	NS
99 Finnish FCHL and low–HDL cholesterol families	1,133, 251, 370	**G**/A	0.31/0.20	0.02	0.03
24 Mexican FCHL families	314, 59, 142	**A**/G	0.50/0.40	NS (0.07)	0.04

*The first value is the number of genotyped individuals, the second value indicates the number of triglyceride affected subjects, and the third value the number of apoB-affected subjects; †risk alleles are indicated in bold; ‡allele frequency in all family members.

TABLE 2
Association analysis of lipid phenotypes and the metabolic syndrome with rs6031558-rs745975-rs3212198 haplotypes in Finnish dyslipidemic families

Study sample and trait	*n**	Haplotype	Frequency†	*P* value
42 Finnish FCHL families	230			
Triglycerides	117	H1A	0.28	0.008
Total cholesterol	131	H1B	0.15	0.04
60 Finnish FCHL families	706			
Triglyceride	202	H1A	0.35	0.03
Total cholesterol	205	H1B	0.15	0.02
Metabolic syndrome	158	H1A	0.35	NS (0.07)
39 Finnish low–HDL cholesterol families	427			
Metabolic syndrome	60	H1B	0.16	0.01
99 Finnish FCHL and low–HDL cholesterol families	1,133			
Total cholesterol	246	H1B	0.16	0.02
Metabolic syndrome	218	H1A	0.35	0.02
		H1B	0.16	0.02

*The number of genotyped and trait-affected subjects for study sample and trait, respectively; †haplotype frequency in all family members. H1A, risk haplotype of 1-1-2 alleles, where 1 indicates the common allele; H1B, protective haplotype of 2-1-1 alleles.

respectively) (Table 3). Haplotype analysis also revealed that the combined analysis of rs2425640 and rs745975 is significantly associated with glucose parameters for common risk (H2A) and protective (H2B) haplotypes (Table 3).

We next examined the four SNPs that formed the haplotypes (rs6031558, rs2425640, rs745975, and rs3212198) and the individually associated rs2144908 in an independent study sample of 427 individuals from 39 Finnish low–HDL cholesterol families. The lipid trait implicated in the low–

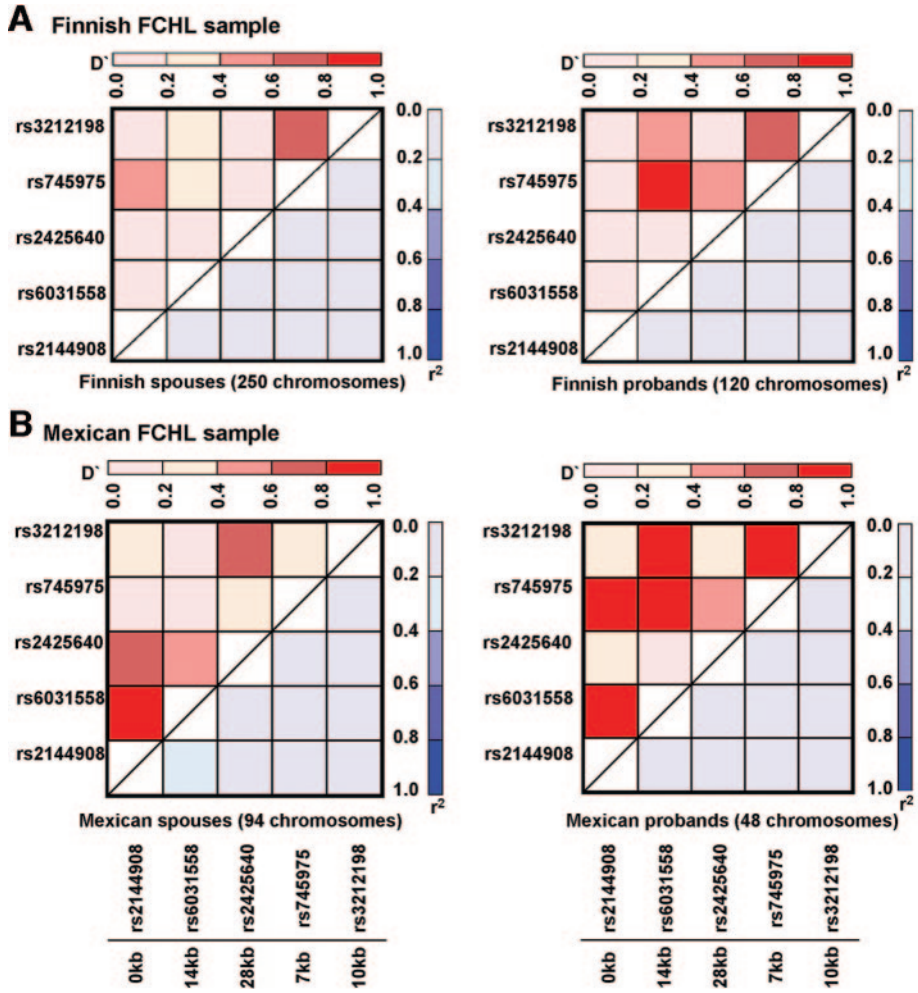

FIG. 2. Pairwise linkage disequilibrium and distances between SNPs (in bp). *A*: Linkage disequilibrium between markers tested in Finnish FCHL spouses (*left*) and probands (*right*). *B*: Linkage disequilibrium between markers tested in Mexican FCHL spouses (*left*) and probands (*right*). D′ is shown in the *upper triangle*, and the scale for D′ is presented at the *top*. r^2 is shown in the *lower triangle*, and the scale for r^2 is presented on the *right*.

TABLE 3
Association analysis of glucose parameters with the SNP rs2425640 and rs2425640-rs745975 haplotypes in 706 genotyped individuals from 60 Finnish FCHL families

SNP or haplotype	Frequency*	AUC glucose *P* value	AUC insulin *P* value	AUC C-peptide *P* value
rs2425640	0.38	0.01	0.008	0.03
rs2425640-rs745975 (H2A)†	0.32	0.01	0.009	0.02
rs2425640-rs745975 (H2B)‡	0.51	0.01	0.006	NS

*Allele/ haplotype frequency in all family members; †H2A, risk haplotype of 2-1 alleles, where 1 indicates the common allele; ‡H2B, protective haplotype of 1-1 alleles.

HDL cholesterol families was high total cholesterol for a different protective haplotype of SNPs rs6031558-rs745975-rs3212198, 1-1-1 ($P = 0.04$). Total cholesterol was also implicated in the combined analysis of 1,133 subjects from 60 FCHL and 39 low–HDL cholesterol families for the same protective haplotype (H1B) as for FCHL families ($P = 0.02$) (Table 2). In the low–HDL cholesterol sample, we did not obtain significant evidence for association with other lipid traits. It is worth noting, however, that a plasma triglyceride level >2.3 mmol/l was an exclusion criterion for the probands of the low–HDL cholesterol families, and there were only 49 subjects affected with high triglycerides in these families.

Association analyses of the metabolic syndrome in Finns. We examined whether the haplotypes are associated with the metabolic syndrome status in the Finnish dyslipidemic families. There were 218 affected individuals for the metabolic syndrome in the Finnish FCHL and low–HDL cholesterol families. Interestingly, the same risk (H1A) and protective (H1B) haplotypes associated with high triglycerides and total cholesterol levels, respectively, were also associated with the metabolic syndrome (Table 2).

Investigation of the previously known nonsynonymous HNF4A SNPs. We examined 16 unrelated affected Finnish FCHL individuals for frequencies of all validated nonsynonymous HNF4A SNPs (rs1800961, rs6093980, rs6031602, and rs10632391) by sequencing. We found that two affected individuals were heterozygotes for the rare nonsynonymous mutation T130I (rs1800961). None of the other variants were polymorphic in the 16 subjects. We genotyped the 42 FCHL Finnish families for the SNP rs1800961. The genotype results revealed only three additional heterozygotes from two separate families. We examined the serum lipid and glucose parameters of the mutation carriers, but the variant did not cosegregate with lipid or glucose traits.

Association analyses of lipid phenotypes in Mexicans. To validate the associations observed in the Finns, we examined the four SNPs that formed the haplotypes (rs6031558, rs2425640, rs745975, and rs3212198) and individually associated rs2144908 in 24 Mexican FCHL families with 314 genotyped individuals. The major allele (A) of rs2144908 was individually associated with elevated apoB levels ($P = 0.04$) (Table 1). The same risk (H1A) and protective (H1B) haplotypes (rs6031558-rs745975-rs3212198) as in the Finns were implicated with high triglycerides in the Mexicans (Table 4), although in the Mexicans, most of the association seems to be derived from the first SNP (rs6031558) of the H1A and H1B haplotypes and the two first SNPs (rs6031558-rs745975) of the H1A and H1B haplotypes (Table 4). We were not able to test the haplotypes for IGT, OGTT, glucose parameters, or the metabolic syndrome in the Mexican families because of a lack of this information. No other lipid traits showed evidence for association with single SNPs or haplotypes.

Association analyses of lipid phenotypes in the combined study sample of Finnish and Mexican FCHL families. We also performed a combined association analysis of 1,020 subjects from 60 Finnish and 24 Mexican FCHL families for the haplotypes implicated in both study samples (Table 5). In the combined analyses, the significance of the risk (H1A) and protective (H1B) haplotypes improved compared with the Finnish FCHL sample alone, resulting in *P* values of 0.006 and 0.005 in the combined analysis (Table 5).

The analyzed traits are highly correlated lipid and glucose parameters, making the Bonferroni correction a conservative approach to correct for multiple testing. Therefore, we also tested the haplotypes, using the –p option of HBAT to correct for multiple testing. Comparable results were obtained with this approach (data not shown).

TABLE 4
Association analysis of the 24 Mexican FCHL families for high triglycerides

SNP or haplotype	Implicated allele or haplotype	Frequency*	Triglycerides *P* value†	Triglycerides *P* value‡
rs6031558	1 Risk allele	0.81	0.004	0.02
H1A haplotype and its combinations (risk haplotypes)§				
H1A (rs6031558-rs745975-rs3212198)	H1A (1-1-2)	0.27	0.02	0.05
rs6031558-rs745975	1-1	0.80	0.0009	0.1
H1B haplotype and its combinations (protective haplotypes)‖				
H1B (rs6031558-rs745975-rs3212198)	H1B (2-1-1)	0.06	0.006	0.1
rs6031558-rs745975	2-1	0.11	0.001	0.05

*Haplotype frequency in all family members; †*P* value obtained using the -o option of FBAT (for the single SNP) or HBAT (for haplotypes); ‡*P* value obtained using the -e option of the FBAT or HBAT program; §H1A, risk haplotype of 1-1-2 alleles, where 1 indicates the common allele; ‖H1B, protective haplotype of 2-1-1 alleles.

TABLE 5
Association analysis of high triglycerides and total cholesterol with rs6031558-rs745975-rs3212198 haplotypes in a combined analysis of 1,020 subjects from 60 Finnish and 24 Mexican FCHL families

Haplotype	Frequency*	Triglycerides *P* value†	Total cholesterol *P* value†
H1A‡	0.32	0.006	0.04
H1B§	0.13	NS (0.08)	0.005

*Haplotype frequency in all family members; †*P* value obtained using the -e option of HBAT; ‡H1A, risk haplotype of 1-1-2 alleles, where 1 indicates the common allele; §H1B, protective haplotype of 2-1-1 alleles.

Subset analyses to investigate whether the associated SNP or haplotype explain the linkage in Finns. We used subset analyses to examine whether the SNP rs2144908 and/or H1A haplotype explain the linkage at 20q12-q13.1 for high triglycerides (16). First, we used the "viewstat" and "viewhap" options of FBAT/HBAT to divide the 92 FCHL and low–HDL cholesterol families of the original linkage study (16) into at-risk and not-at-risk families based on whether the family contributed positively to the *z*-score statistics. For H1A, the 33 at-risk, 59 not-at-risk, and all 92 families produced logarithm of odds (LOD) scores of 1.8, 0.6, and 2.3 for the peak marker D20S102, suggesting that the H1A haplotype explains most of the linkage. For rs2144908, 22 at-risk and 70 not-at-risk families produced LOD scores of 1.1 and 1.2, respectively, indicating that this SNP cannot explain the original linkage.

Second, we divided the families into linked and nonlinked families, based on whether the family had a positive LOD score (>0.0) for the peak marker D20S102 (16), and we performed association analysis separately in each group, using the HHRR test for rs2144908 and HBAT for the H1A haplotype. Comparison of the association results for the H1A haplotype between the 27 linked, 65 nonlinked, and all 92 families showed that the haplotype association is obtained from the linked families ($P = 0.03$ and 0.4, 0.05, respectively). For rs2144908, the nonlinked families produced a *P* value of 0.01, and the linked and all 92 families produced *P* values of 0.8 and 0.02, respectively, suggesting the possibility of locus heterogeneity even in Finns. The results of the multipoint analyses were in very good agreement with these two-point results (data not shown). Taken together, these data suggest that the H1A haplotype explains most of the linkage (16).

Pairwise linkage disequilibrium between FCHL affected and unaffected individuals in Finnish and Mexican samples. We examined the pairwise linkage disequilibrium of the five implicated SNPs separately in the Finnish and Mexican FCHL probands and unaffected spouses (Fig. 2). In both the Finns and Mexicans, pairwise linkage disequilibrium was low in spouses, whereas for the SNPs forming the associated haplotypes, the pairwise linkage disequilibrium, measured by D′, appears to be higher in the Finnish and Mexican affected subjects than in spouses. (Fig. 2*A* and *B*; *upper triangle*). However, the r^2 values indicate no linkage disequilibrium between SNPs in either the Finnish or Mexican probands or spouses (Fig. 2). This difference between the D′ and r^2 results may reflect the dissimilarity of the allele frequencies of the SNPs because, unlike D′ statistics, the r^2 statistics depends on the allele frequencies of the two SNPs. Because small sample size (<100 chromosomes) is also known to affect the linkage disequilibrium estimates (38), the Mexican linkage disequilibrium estimates need to be interpreted with caution.

Functional investigation of the associated HNF4A variants. We hypothesized that the five implicated SNPs were either functional or in linkage disequilibrium with unknown causative variants. We did not choose to sequence the entire 77-kb region because of its size. Alternatively, we used experimental and computational tools to explore the potential functional significance of these five SNPs (online appendix Fig. 1). All five SNPs are intronic (Fig. 1 and online appendix Fig. 1*A*). Because the SNP rs745975 resides 2 bp apart from the consensus 3′ splice site in intron 2, it may alter splicing. The intronic tag SNP rs3212198 is in high linkage disequilibrium with SNPs within exons 6–9, possibly tagging a functional variant within this region. Next, we examined the conservation of each SNP and its surrounding area (~225 bp) across five species (online appendix Fig. 1*B*). The SNP rs6031558 is located in a conserved intronic region, and the tag SNP rs3212198 is in close proximity to a highly conserved intronic region, suggesting that these SNPs may tag variants in potential regulatory regions. We also examined both computationally and experimentally whether the five SNPs alter binding sites for DNA binding proteins. Using the Genomatix Suite's ElDorado tool, we identified several putative changes for four SNPs (online appendix Fig. 1*C*). None of the five investigated SNPs produced a specific electrophoretic mobility shift (data not shown). Because we have not exhaustively examined the variants under different conditions, it is still possible that one of the variants binds nuclear factors. To conclude, the possible regulatory role of the five SNPs cannot be fully excluded without more detailed functional analyses. However, because none of these SNPs represents a missense or nonsense variant with obvious functional consequences, it is more likely that the actual functional variants are in the context of the associated haplotypes formed by these SNPs.

DISCUSSION

In the current study, we demonstrated that common HNF4A variants and their haplotypes are associated with high serum lipid levels in Finnish and Mexican FCHL families. Importantly, both populations shared lipid-associated HNF4A risk and protective haplotypes. Furthermore, this is the first study implicating the HNF4A gene in the metabolic syndrome.

The metabolic syndrome is associated with increased risk of cardiovascular disease (39). Here, the connection between HNF4A and the metabolic syndrome was established by investigating two different dyslipidemic cohorts, Finnish FCHL and low–HDL cholesterol families, who differ significantly in the inclusion and exclusion criteria for lipids. With the exception of total cholesterol and the metabolic syndrome, no other traits showed evidence for association in this combined set of FCHL and low–HDL cholesterol families, suggesting that susceptibility variants of these phenotypic traits partially overlap among these clinically different dyslipidemic families.

The extent of linkage disequilibrium is relatively low at the HNF4A locus (4,5), making it challenging to capture most of the genetic variation. Thus, by genotyping the SNPs of the current study, we did not cover the majority of the common variations in the 77-kb gene region, making it

unlikely that the investigated SNPs represent the actual causative variant(s). Interestingly, a gene called R3HDML (R3H domain containing-like), with an unknown function, forms an additional putative regional candidate gene that may explain the observed association because R3HDML resides 5′ to HNF4A in the same large linkage disequilibrium block with rs2144908 and rs6031558. In the current study, only two SNPs showed evidence of association as singletons. However, because all four of the SNPs (rs6031558, rs2425640, rs745975, and rs3212198) implicated in the haplotypes (H1A-B and H2A-B) represent nonredundant SNPs based on the HapMap data, additional genetic information was captured by analyzing these haplotypes in the association analyses. Because the functionally relevant polymorphism(s) can be in the context of these haplotypes, sequencing of the subjects with the risk/protective haplotypes for the 77-kb HNF4A gene region is warranted in future studies.

The two SNPs implicated as singletons in the FCHL families have previously been associated with type 2 diabetes (1,2,4). The SNP rs2144908 was associated with type 2 diabetes in Finns and Ashkenazi Jews (1,4) and the rs2425640 in Finns and the Amish (2,4). However, the association of rs1884613, in complete linkage disequilibrium with rs2144908, was not confirmed in a recent large study (5). The H1A and H1B haplotypes have not yet been tested for type 2 diabetes. The haplotype SNP rs6031558 seemed to be individually associated with type 2 diabetes in Pima Indians (40), although the variant was rare in Pima Indians. The haplotype SNP rs745975 was not individually associated with type 2 diabetes in Finns (4). These discrepancies may imply that the SNPs do not represent the actual functional variants.

Many of the genes regulated by HNF4A are involved in glucose and lipid metabolism (41,42). Accordingly, HNF4A is an excellent candidate gene for the clinical characteristics of FCHL and type 2 diabetes. The association of HNF4A variants with several lipid and glucose traits provides further support for the hypothesis that HNF4A is involved in the complex FCHL phenotype.

The upstream transcription factor 1 (USF1) gene has been recently associated with FCHL in several populations (28,43,44). Similar to HNF-4α, USF1 is a transcription factor, and many of the genes regulated by USF1 are also involved in lipid and glucose metabolism. A recent study demonstrates that sequences from sites bound by HNF-4α and USF1 showed significant overlap in HepG2 cells (45). Interactions between DNA sequence variants are expected to play a crucial role in complex traits, such as FCHL. Based on this (45) and the current study, it is tempting to propose that variants in USF1 and HNF4A may interactively confer susceptibility to FCHL. In future studies, we intend to investigate the possible interaction between USF1 and HNF4A risk alleles both in Finnish and Mexican dyslipidemic families.

In conclusion, we observed that common DNA sequence variants in the HNF4A gene are associated with elevated serum lipid levels and the metabolic syndrome in two independent Finnish study samples. Importantly, we also observed the same common HNF4A haplotypes to be associated with elevated serum lipid levels in Mexican FCHL families. Because HNF4A variants have now been implicated in type 2 diabetes, FCHL, and the metabolic syndrome, this association may partly explain the well-known phenotypic overlap between these common disorders.

ACKNOWLEDGMENTS

This research was supported by a UC-MEXUS CONACYT grant, National Institutes of Health Grants HL-28481 and HL-70150, and American Heart Association Grant 0430180N.

We thank the patients and family members for their participation in this study. L. Peltonen is thanked for the Finnish sample collection; A. Lopez for the Mexican sample collection; and L. Riba, E. Nikkola, A. Jokiaho, H. Heikelä, and S. Ramirez for laboratory technical assistance.

REFERENCES

1. Love-Gregory LD, Wasson J, Ma J, Jin CH, Glaser B, Suarez BK, Permutt MA: A common polymorphism in the upstream promoter region of the hepatocyte nuclear factor 4 alpha gene on chromosome 20q is associated with type 2 diabetes and appears to contribute to the evidence for linkage in an Ashkenazi Jewish population. *Diabetes* 53:1134–1140, 2004
2. Damcott CM, Hoppman N, Ott SH, Reinhart LJ, Wang J, Pollin TI, O'Connell JR, Mitchell BD, Shuldiner AR: Polymorphisms in both promoters of hepatocyte nuclear factor 4-α are associated with type 2 diabetes in the Amish. *Diabetes* 53:3337–3341, 2004
3. Hansen SK, Rose CS, Glumer C, Drivsholm T, Borch-Johnsen K, Jorgensen T, Pedersen O, Hansen T: Variation near the hepatocyte nuclear factor (HNF)-4alpha gene associates with type 2 diabetes in the Danish population. *Diabetologia* 48:452–458,2005
4. Silander K, Mohlke KL, Scott LJ, Peck EC, Hollstein P, Skol AD, Jackson AU, Deloukas P, Hunt S, Stavrides G, Chines PS, Erdos MR, Narisu N, Conneely KN, Li C, Fingerlin TE, Dhanjal SK, Valle TT, Bergman RN, Tuomilehto J, Watanabe RM, Boehnke M, Collins FS: Genetic variation near the hepatocyte nuclear factor-4α gene predicts susceptibility to type 2 diabetes. *Diabetes* 53:1141–1149, 2004
5. Winckler W, Graham RR, de Bakker PI, Sun M, Almgren P, Tuomi T, Gaudet D, Hudson TJ, Ardlie KG, Daly MJ, Hirschhorn JN, Groop L, Altshuler D: Association testing of variants in the hepatocyte nuclear factor-4α gene with risk of type 2 diabetes in 7,883 people. *Diabetes* 54:886–892, 2005
6. Yamagata K, Furuta H, Oda N, Kaisaki PJ, Menzel S, Cox NJ, Fajans SS, Signorini S, Stoffel M, Bell GI: Mutations in the hepatocyte nuclear factor 4, alpha gene in maturity-onset diabetes of the young (MODY1). *Nature* 384:458–460, 1996
7. Sladek FM, Seidel S: Hepatocyte nuclear factor 4a. In *Nuclear Receptors and Genetic Disease*. Burris TP, McCabeed, Eds. London, Academic Press, 2001
8. Thomas H, Jaschkowitz K, Bulman M, Frayling TM, Mitchell SM, Roosen S, Lingott-Frieg A, Tack CJ, Ellard S, Ryffel GU, Hattersley AT: A distant upstream promoter of the HNF-4 alpha gene connects the transcription factors involved in maturity-onset diabetes of the young. *Hum Mol Genet* 10:2089–2097, 2001
9. Nakhei H, Lingott A, Lemm I, Ryffel GU: An alternative splice variant of the tissue specific transcription factor HNF-4 alpha predominates in undifferentiated murine cell types. *Nucleic Acid Res* 26:497–504, 1998
10. Grundy SM, Chait A, Brunzell JD: Familial combined hyperlipidemia workshop. *Arteriosclerosis* 7:203–207, 1989
11. Goldstein JL, Schrott HG, Hazzard WR, Bierman EL, Motulsky AG: Hyperlipidemia in coronary heart disease. II. Genetic analysis of lipid levels in 176 families and delineation of a new inherited disorder, combined hyperlipidemia. *J Clin Invest* 52:1544–1568, 1973
12. Vakkilainen J, Porkka KV, Nuotio I, Pajukanta P, Suurinkeroinen L, Ylitalo K, Viikari JS, Ehnholm C, Taskinen MR: Glucose intolerance in familial combined hyperlipidaemia: EUFAM Study Group. *Eur J Clin Invest* 28:24–32, 1998
13. Nikkilä EA, Aro A: Family study of serum lipids and lipoproteins in coronary heart-disease. *Lancet* 1:954–959, 1973
14. Pajukanta P, Nuotio I, Terwilliger JD, Porkka KV, Ylitalo K, Pihlajamaki J, Suomalainen AJ, Syvanen AC, Lehtimaki T, Viikari JS, Laakso M, Taskinen MR, Ehnholm C, Peltonen L: Linkage of familial combined hyperlipidaemia to chromosome 1q21–q23. *Nat Genet* 18:369–373, 1998
15. Soro A, Pajukanta P, Lilja HE, Ylitalo K, Hiekkalinna T, Perola M, Cantor RM, Viikari JS, Taskinen MR, Peltonen L: Genome scans provide evidence for low-HDL-C loci on chromosomes 8q23, 16q24.1–24.2, and 20q13.11 in Finnish families. *Am J Hum Genet* 70:1333–1340, 2002
16. Lilja HE, Suviolahti E, Soro-Paavonen A, Hiekkalinna T, Day A, Lange K, Sobel E, Taskinen MR, Peltonen L, Perola M, Pajukanta P: Locus for quantitative HDL-cholesterol on chromosome 10q in Finnish families with dyslipidemia. *J Lipid Res* 45:1876–1884, 2004

17. Yu Y, Wyszynski DF, Waterworth DM, Wilton SD, Barter PJ, Kesaniemi YA, Mahley RW, McPherson R, Waeber G, Bersot TP, Ma Q, Sharma SS, Montgomery DS, Middleton LT, Sundseth SS, Mooser V, Grundy SM, Farrer LA: Multiple QTLs influencing triglyceride and HDL and total cholesterol levels identified in families with atherogenic dyslipidemia. *J Lipid Res* 46:2202–2213, 2005
18. Li WD, Dong C, Li D, Garrigan C, Price RA: A genome scan for serum triglyceride in obese nuclear families. *J Lipid Res* 46:432–438, 2005
19. Bowden DW, Sale M, Howard TD, Qadri A, Spray BJ, Rothschild CB, Akots G, Rich SS, Freedman BI: Linkage of genetic markers on human chromosomes 20 and 12 to NIDDM in Caucasian sib pairs with a history of diabetic nephropathy. *Diabetes* 46:882–886, 1997
20. Ji L, Malecki M, Warram JH, Yang Y, Rich SS, Krolewski AS: New susceptibility locus for NIDDM is localized to human chromosome 20q. *Diabetes* 46:876–881, 1997
21. Lembertas AV, Perusse L, Chagnon YC, Fisler JS, Warden CH, Purcell-Huynh DA, Dionne FT, Gagnon J, Nadeau A, Lusis AJ, Bouchard C: Identification of an obesity quantitative trait locus on mouse chromosome 2 and evidence of linkage to body fat and insulin on the human homologous region 20q. *J Clin Invest* 100:1240–1247, 1997
22. Zouali H, Hani EH, Philippi A, Vionnet N, Beckmann JS, Demenais F, Froguel P: A susceptibility locus for early-onset non-insulin dependent (type 2) diabetes mellitus maps to chromosome 20q, proximal to the phosphoenolpyruvate carboxykinase gene. *Hum Mol Genet* 6:1401–1408, 1997
23. Ghosh S, Watanabe RM, Hauser ER, Valle T, Magnuson VL, Erdos MR, Langefeld CD, Balow J Jr, Ally DS, Kohtamaki K, Chines P, Birznieks G, Kaleta HS, Musick A, Te C, Tannenbaum J, Eldridge W, Shapiro S, Martin C, Witt A, So A, Chang J, Shurtleff B, Porter R, Kudelko K, Unni A, Segal L, Sharaf R, Blaschak-Harvan J, Eriksson J, Tenkula T, Vidgren G, Ehnholm C, Tuomilehto-Wolf E, Hagopian W, Buchanan TA, Tuomilehto J, Bergman RN, Collins FS, Boehnke M: Type 2 diabetes: evidence for linkage on chromosome 20 in 716 Finnish affected sib pairs. *Proc Natl Acad Sci U S A* 96:2198–2203, 1999
24. Permutt MA, Wasson JC, Suarez BK, Lin J, Thomas J, Meyer J, Lewitzky S, Rennich JS, Parker A, DuPrat L, Maruti S, Chayen S, Glaser B: A genome scan for type 2 diabetes susceptibility loci in a genetically isolated population. *Diabetes* 50:681–685, 2001
25. Olefsky JM, Farquhar JW, Reaven GM: Reappraisal of the role of insulin in hypertriglyceridemia. *Am J Med* 57:551–560, 1974
26. Shih DQ, Dansky HM, Fleisher M, Assmann G, Fajans SS, Stoffel M: Genotype/phenotype relationships in HNF-4α/MODY1: haploinsufficiency is associated with reduced apolipoprotein (AII), apolipoprotein (CIII), lipoprotein(a), and triglyceride levels. *Diabetes* 49:832–837, 2000
27. Pajukanta P, Terwilliger JD, Perola M, Hiekkalinna T, Nuotio I, Ellonen P, Parkkonen M, Hartiala J, Ylitalo K, Pihlajamaki J, Porkka K, Laakso M, Viikari J, Ehnholm C, Taskinen MR, Peltonen L: Genomewide scan for familial combined hyperlipidemia genes in Finnish families, suggesting multiple susceptibility loci influencing triglyceride, cholesterol, and apolipoprotein B levels. *Am J Hum Genet* 64:1453–1463, 1999
28. Huertas-Vazquez A, Aguilar-Salinas C, Lusis 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. *Arterioscler Thromb Vasc Biol* 25:1985–1991, 2005
29. Aguilar-Salinas CA, Olaiz G, Valles V, Torres JM, Gomez Perez FJ, Rull JA, Rojas R, Franco A, Sepulveda J: High prevalence of low HDL cholesterol concentrations and mixed hyperlipidemia in a Mexican nationwide survey. *J Lipid Res* 42:1298–307, 2002
30. World Health Organization: *Diabetes Mellitus*: *Report of a WHO Study Group*. Geneva, World Health Org., 1985 (Tech. Rep. Ser., no. 727), p. 1–113
31. Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III): Executive summary of the Third Report of the National Cholesterol Education Program (NCEP). *JAMA* 285:2486–2497, 2001
32. Terwilliger JD, Ott J: A haplotype-based 'haplotype relative risk' approach to detecting allelic associations. *Hum Hered* 42:337–346, 1992
33. Horvath S, Xu X, Lake SL, Silverman EK, Weiss ST, Laird NM: Family-based tests for associating haplotypes with general phenotype data: application to asthma genetics. *Genet Epidemiol* 26:61–69, 2004
34. Göring HH, Terwilliger JD: Linkage analysis in the presence of error III: marker loci and their map as nuisance parameters. *Am J Hum Genet* 66:1298–1309, 2000
35. O'Connell JR, Weeks DE: PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 63:259–266, 1998
36. Lathrop GM, Lalouel J-M, Julier CA, Ott J: Strategies for multilocus linkage analysis in humans. *Proc Natl Acad Sci U S A* 81:3443–3446, 1984
37. Sobel E, Lange K: Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. *Am J Hum Genet* 58:1323–1337, 1996
38. Terwilliger JD, Haghighi F, Hiekkalinna TS, Göring HH: A biased assessment of the use of SNPs in human complex traits. *Curr Opin Genet Dev* 12:726–734, 2002
39. Eckel RH, Grundy SM, Zimmet PZ: The metabolic syndrome. *Lancet* 365:1415–1428, 2005
40. Muller YL, Infante AM, Hanson RL, Love-Gregory L, Knowler W, Bogardus C, Baier LJ: Variants in hepatocyte nuclear factor 4α are modestly associated with type 2 diabetes in Pima Indians. *Diabetes* 54:3035–3039, 2005
41. Wang H, Maechler P, Antinozzi PA, Hagenfeldt KA, Wollheim CB: Hepatocyte nuclear factor 4alpha regulates the expression of pancreatic beta-cell genes implicated in glucose metabolism and nutrient-induced insulin secretion. *J Biol Chem* 275:35953–35959, 2000
42. Rhee J, Inoue Y, Yoon JC, Puigserver P, Fan M, Gonzalez FJ, Spiegelman BM: Regulation of hepatic fasting response by PPARgamma coactivator-1alpha (PGC-1): requirement for hepatocyte nuclear factor-4 alpha in gluconeogenesis. *Proc Natl Acad Sci U S A* 100:4012–4017, 2003
43. Pajukanta P, Lilja HE, Sinsheimer JS, Cantor RM, Lusis AJ, Gentile M, Duan XJ, Soro-Paavonen A, Naukkarinen J, Saarela J, Laakso M, Ehnholm C, Taskinen MR, Peltonen L: Familial combined hyperlipidemia is associated with upstream transcription factor 1 (USF1). *Nat Genet* 36:371–376, 2004
44. 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. *Hum Genet* 117:444–451, 2005
45. Rada-Iglesias A, Wallerman O, Koch C, Ameur A, Enroth S, Clelland G, Wester K, Wilcox S, Dovey OM, Ellis PD, Wraight VL, James K, Andrews R, Langford C, Dhami P, Carter N, Vetrie D, Ponten F, Komorowski J, Dunham I, Wadelius C: Binding sites for metabolic disease related transcription factors inferred at base pair resolution by chromatin immunoprecipitation and genomic microarrays. *Hum Mol Genet* 14:3435–3447, 2005