

ARTICLE

Fine mapping and association studies of a high-density lipoprotein cholesterol linkage region on chromosome 16 in French-Canadian subjects

Zari Dastani^{1,2}, Päivi Pajukanta³, Michel Marcil¹, Nicholas Rudzicz⁴, Isabelle Ruel¹, Swneke D Bailey², Jenny C Lee³, Mathieu Lemire^{5,9}, Janet Faith⁵, Jill Platko^{6,10}, John Rioux^{6,11}, Thomas J Hudson^{2,5,7,9}, Daniel Gaudet⁸, James C Engert^{*,2,7}, Jacques Genest^{1,2,7}

Low levels of high-density lipoprotein cholesterol (HDL-C) are an independent risk factor for cardiovascular disease. To identify novel genetic variants that contribute to HDL-C, we performed genome-wide scans and quantitative association studies in two study samples: a Quebec-wide study consisting of 11 multigenerational families and a study of 61 families from the Saguenay–Lac St-Jean (SLSJ) region of Quebec. The heritability of HDL-C in these study samples was 0.73 and 0.49, respectively. Variance components linkage methods identified a LOD score of 2.61 at 98 cM near the marker D16S515 in Quebec-wide families and an LOD score of 2.96 at 86 cM near the marker D16S2624 in SLSJ families. In the Quebec-wide sample, four families showed segregation over a 25.5-cM (18 Mb) region, which was further reduced to 6.6 Mb with additional markers. The coding regions of all genes within this region were sequenced. A missense variant in *CHST6* segregated in four families and, with additional families, we observed a *P* value of 0.015 for this variant. However, an association study of this single-nucleotide polymorphism (SNP) in unrelated Quebec-wide samples was not significant. We also identified a SNP (rs11646677) in the same region, which was significantly associated with a low HDL-C (*P*=0.016) in the SLSJ study sample. In addition, RT-PCR results from cultured cells showed a significant difference in the expression of *CHST6* and *KIAA1576*, another gene in the region. Our data constitute additional evidence for a locus on chromosome 16q23-24 that affects HDL-C levels in two independent French-Canadian studies.

European Journal of Human Genetics (2010) 18, 342–347; doi:10.1038/ejhg.2009.157; published online 21 October 2009

Keywords: HDL-C; family study; complex traits; coronary heart disease; gene identification

INTRODUCTION

Coronary artery disease (CAD) is the most common cause of death in Western societies, as well as in emerging market economies.^{1,2} Epidemiological studies have consistently shown an association between low levels of high-density lipoprotein cholesterol (HDL-C) and the prevalence of CAD.^{3–5} HDL-C has a key role in reverse cholesterol transport, mobilizing cholesterol from peripheral tissues to the liver, a mechanism contributing to the cardio-protective effect of HDL-C.

Several studies have shown modest to high estimates of heritability (0.24–0.83) for plasma HDL-C levels,^{6,7} underlining the importance of genetic factors. However, until recently, only a few specific genes had been identified (see Dastani *et al*⁶ for review). Rare mutations in the ATP-binding cassette A1 (*ABCA1*), lecithin:cholesterol acyltransferase (*LCAT*) and apolipoprotein A-I (*APOA1*) genes, and other genes involved in HDL metabolism, have been reported in individuals with low levels of HDL-C.^{8,9} However, these account for ~26% of

familial forms of HDL deficiency in Quebec.^{10,11} Identifying novel genes that regulate HDL-C levels may provide further insights into lipoprotein metabolism and potential pathways for pharmacological modulation of HDL-C.

In this study, we examined the genetic determinants of low HDL-C in well-characterized families and individuals of French-Canadian descent. Specifically, we performed genome-wide linkage followed by association analyses of single-nucleotide polymorphisms (SNPs) located on chromosome 16q23-24 with the HDL-C trait, and examined the expression of a number of candidate genes in this region.

METHODS

Subjects

Probands for the QUE study sample with HDL-C levels less than the 5th percentile, (age/sex specific) based on the Lipid Research Clinics population studies data book, as previously described,¹² were selected from the Preventive

¹Cardiovascular Genetics Laboratory, McGill University Health Centre, Royal Victoria Hospital, Montreal, Quebec, Canada; ²Department of Human Genetics, McGill University, Montreal, Quebec, Canada; ³Department of Human Genetics, David Geffen School of Medicine at UCLA, University of California, Los Angeles, CA, USA; ⁴Department of Computer Science, McGill University, Montreal, Quebec, Canada; ⁵McGill University and Genome Quebec Innovation Centre, Montreal, Quebec, Canada; ⁶Whitehead Institute for Biomedical Research, MIT, Cambridge, MA, USA; ⁷Department of Medicine, McGill University, Montreal, Quebec, Canada; ⁸Dyslipidemia, Diabetes and Atherosclerosis Group and Community Genomics Research Center, Université de Montréal and Complexe hospitalier de la Sagamie, Chicoutimi, Quebec, Canada

*Correspondence: Dr JC Engert, Departments of Medicine and Human Genetics, McGill University, Royal Victoria Hospital, H7.30, 687 Pine Avenue West, Montreal, Quebec, Canada H3A 1A1. Tel: +1 514 934-1934; Fax +1 514 843-2843; E-mail: jamie.engert@mcgill.ca ⁹Present address: Ontario Institute for Cancer Research, Toronto, Ontario, Canada. ¹⁰Present address: PHT Corporation, 500 Rutherford Avenue, Boston, MA, USA; ¹¹Present address: Broad Institute of MIT and Harvard, Cambridge, MA, USA; Université de Montréal, Montreal, Quebec, Canada; Montreal Heart Institute, Montreal, Quebec, Canada.

Received 29 January 2009; revised 24 June 2009; accepted 10 August 2009; published online 21 October 2009

Cardiology/Lipid Clinic of the McGill University Health Center, Royal Victoria Hospital (Montreal, Quebec, Canada). The exclusion criteria for probands were the following: severe hypertriglyceridemia (>10 mmol/l), cellular cholesterol or phospholipid efflux defect in skin fibroblasts¹³ or the presence of known *ABCA1* or *ApoA1* mutations based on exonic sequencing. A total of 11 families out of 29 recruited multigenerational French-Canadian families, each exhibiting at least three individuals with a low HDL in consecutive generations ($n=412$ subjects), were selected for genome-wide linkage analysis. The size of the families varied between 4 and 109 individuals, with a median of 41. Family members were sampled after a 12-h fast and a discontinuation of lipid modifying medications for at least 2 weeks. A total of 136 cases (HDL $<5\%$) and 218 controls (HDL $>25\%$) derived from the QUE families and additional CAD patients were selected for an association study. These additional CAD patients consisted of unrelated French-Canadian patients with premature CAD.¹⁴ All subjects provided informed consent for plasma and DNA sampling, isolation and storage. The Research Ethics Board of the McGill University Health Centre approved the research protocol. The Saguenay-Lac St-Jean (SLSJ) sample composed of 410 phenotyped individuals from 61 nuclear families from the SLSJ region of Quebec who were ascertained as a part of ongoing family-based studies of CAD or type II diabetes. Both studies required that at least two siblings be affected. Biochemical measurements were determined after a 3-week medication washout period for all hypocholesterolemic and antihypertensive drugs. The SLSJ case-control sample was predominantly derived from the families and consisted of 94 individuals with an HDL-C <5 th percentile and 94 individuals with an HDL-C >40 th percentile. Table 1 summarizes the subjects for this study.

Biochemical measurements

High-density lipoprotein and triglycerides were measured using standardized techniques as described previously.¹³

Genome-scan markers and data quality control

The genotyping for QUE families was performed by DeCode Genetics (Reykjavik, Iceland) with 485 autosomal microsatellite markers from the Decode map (an average marker density of 6.8 cM) on 412 subjects from the QUE families ($n=11$, mean family size=41). The genome scan of 410 phenotyped individuals from 61 SLSJ families was executed as described in Rioux *et al*¹⁵ by using either 312 or 396 microsatellite markers, with an average marker density of 11.9 and 9 cM, respectively. These markers were modified versions of the Cooperative Human linkage Center Screening Set version 6.0 (<http://gai.nci.nih.gov/cgi-bin/CHLCMapsChr?v4.recmin.v6.weber-screen.sexave>), which were supplemented with additional Genethon markers to increase the information content.¹⁶ Overall, 34 markers were common between the two marker sets used; however, there were no markers in common on chromosome 16.

For both the SLSJ and QUE studies, the software GRR¹⁷ (<http://www.sph.umich.edu/csg/abecasis/GRR/>) was used to identify mispaternities and sample mix-ups. In total, three mispaternities and one sample mix-up were identified and appropriate adjustments were carried out before further analysis.

Sequencing and genotyping

Sequencing of all 39 known genes in the narrowed region was carried out using the Thermo-Sequenase BigDye Direct Cycle-Sequencing kit (Applied

Biosystems Inc., Streetsville, ON, Canada). We also selected 200 tagging SNPs (tSNPs) that spanned the region from 55574820 to 77244563 bp (NCBI, build 36.1) corresponding to the linkage peak on chromosome 16q13-23.1. The selected SNPs were based on $r^2 < 0.8$ between tSNPs and a minor allele frequency > 0.05 from the European samples of the HapMap project in genes that were differentially expressed after cholesterol or 22-hydroxycholesterol stimulation of human fibroblasts, using Affymetrix (Santa Clara, CA, USA) microarray data (Supplementary Table 1). The SNPs were genotyped in the SLSJ study sample using the SNPstream Genotyping System (Beckman Coulter, Inc., Mississauga, ON, Canada). We also genotyped 51 tSNPs in the QUE sample in a parallel study as previously described.¹⁸

Haplotype construction

Ten families with a positive LOD score (LOD > 0) on chromosome 16 were chosen for haplotype analysis. Haplotypes were constructed with the program Merlin¹⁹ (<http://www.sph.umich.edu/csg/abecasis/Merlin/index.html>) using both microsatellite markers and SNPs based on HDL-C $< 5\%$ as the affected status. The haplotype analysis for SNPs in unrelated subjects from the SLSJ and the QUE study sample was performed using Haploview v4.1^{18,20} (<http://www.broadinstitute.org/haploview/haploview>).

RNA extraction and RT-PCR

Skin fibroblasts were available for three out of four probands from the QUE families with the highest linkage scores. These and three control cell lines were cultured on the basis of a standard protocol.¹³ Total RNA was prepared from cultured fibroblasts using the RNeasy mini RNA extraction kit (Qiagen, Mississauga, ON, Canada) according to the manufacturer's instructions. Total RNA (100 ng) in a 14- μ l reaction was reverse transcribed using the cDNA Qiagen Kit according to the manufacturer's protocols. Real-time quantitative PCR assays were carried out using the Quantitect SYBR Green PCR kit (Qiagen) on an ABI PRISM 7300 Sequence Detection System (Applied Biosystems). Amplifications were carried out in a 96-well plate with 50 μ l reaction volumes and 40 amplification cycles (94°C, 15 s; 55°C, 30 s; 72°C, 34 s). Experiments were carried out in triplicate and the mRNA expression was taken as the mean of three different experiments. The expression of each gene was normalized to succinate dehydrogenase expression levels. Fold changes relative to controls were determined using the $\Delta\Delta C_t$ method. A mean fold change ($2^{-\Delta\Delta C_t}$) and the SEM value were determined and significance was determined with a two-tailed *t*-test.

Statistical analyses

SIMWALK2 (version 2.91) (<http://www.genetics.ucla.edu/software/download?package=2>) was used to estimate the multipoint inheritance vectors in QUE families.^{21,22} Genome scans were performed using the variance components approach as implemented in SOLAR (version 4.2.0) (http://www.sfn.org/departments/genetics_detail.aspx?p=37). For multipoint linkage analysis, the HDL-C phenotype was adjusted for triglycerides, sex and age \times sex using SOLAR. Association tests in families were conducted using the Mendel (version 9.0.0²³) (<http://www.genetics.ucla.edu/software/mendel>) with the option for association given linkage. Case-control association analyses were carried out with Haploview. Heritability of HDL-C was analyzed with the method implemented in the SOLAR software package with the 'ascertainment correction' option. Although we realize that using this correction may still lead to biased results, the heritability we report for our samples is consistent with previous reports (see refs. 6, 7 and results section).

RESULTS

In QUE families, heritability of HDL-C was estimated to be 73% ($P=9.89 \times 10^{-16}$) in genotyped individuals. Sex, age \times sex and triglycerides were significant contributors ($P < 0.05$) to the variance of HDL-C in this study sample. The heritability estimate for SLSJ families was 49% ($P=6.55 \times 10^{-22}$). Sex and triglycerides were significant covariates in this study sample. The results of the linkage analysis for HDL-C in the SLSJ and QUE studies are shown in Supplementary Figure 1.

Table 1 Subjects in this study

Subjects	QUE	SLSJ
<i>Family</i>		
No. of families	11	61
No. of individuals	412	410
<i>Case-control</i>		
No. of cases	136	94
No. of controls	218	94

QUE: samples from French-Canadian families from across Quebec.
SLSJ: samples from the SLSJ region of Quebec.

Table 2 Whole genome-scan results with LOD score > 1

Cohort	Chromosome	Location (cM)	LOD score
QUE	16	98	2.61
	15	8	1.68
	5	217	1.24
	1	209	1.08
SLSJ	16	86	2.96

In the genome scan for the QUE study sample, we identified four chromosomal regions that may harbor QTL for HDL-C (LOD scores > 1). We observed an LOD score of 2.61 at 98 cM (based on the Decode Map²⁴) near the marker D16S515 in chromosome 16q23-24. In the SLSJ study sample, the highest multipoint LOD score (2.96) was observed at 86 cM on chromosome 16, located 1.25 cM away from D16S2624 in chromosome 16q22-23. Chromosomal regions with a multipoint LOD score ≥ 1 in either of the two genome scans are summarized in Table 2. Figure 1 shows the multipoint LOD scores on chromosome 16 for the SLSJ and QUE study samples, as well as two previously reported findings in the region.

Haplotype construction identified a 25.5 cM (18 Mb) region on 16q23-24 that segregated with the low HDL-C phenotype in the four QUE families with the highest HDL-C LOD scores. The segregating region was further restricted by additional microsatellite markers from 25.5 cM to 18.1 cM (~7.8 Mb) in the family with the highest LOD score (Figure 2a). Using the UCSC genome browser (<http://genome.ucsc.edu/>), we identified 39 genes residing within this 7.8-Mb region.

In the SLSJ study sample, we tested for the association between HDL-C affection and 200 tSNPs selected from genes in the region, which were differentially expressed in human fibroblasts treated with cholesterol or 22-hydrocholesterol and thus seemed to be sterol regulated (microarray data not shown). One SNP, rs11646677, which was located within the shared haplotype from QUE families, showed evidence of association with HDL-C affection ($P=0.016$) (Supplementary Figure 2).

We sequenced all coding regions of the 39 genes located in this region (Figure 2a). The *LCAT* gene was excluded from the candidate region by a recombination event (subject no. 407, Supplementary Figure 3a) and by sequencing in all QUE probands. We identified several nonsynonymous variants within the 39 genes. These were genotyped in family members and checked for segregation. We found three nonsynonymous variants within the *CHST6* gene and its homolog *CHST5*, located in the same region. One of the variants in the *CHST6* gene, R162G, segregated with HDL-C in the four QUE families that had the highest linkage score (Supplementary Figure 3a, b, c and d). In the association analysis of these families and of additional families recruited later using Mendel with the option for association given linkage, we observed a P value of 0.015 for this variant. However, an association study in 136 cases and 218 controls did not confirm this finding (data not shown). This SNP and four other telomeric SNPs (spanning from 74070744 to 74822026 bp build 36.1) comprise a haplotype (Figure 2b) that is found in the four families but that had a haplotype frequency of only 0.05 in our unrelated samples from Quebec.

The nonsynonymous variants found in all the sequenced genes are listed in Table 3. Segregation analyses with these SNPs narrowed down the region to 6.6 Mb in the family with the highest LOD score.

As none of the other nonsynonymous variants identified by sequence analysis segregated in our families, we investigated the

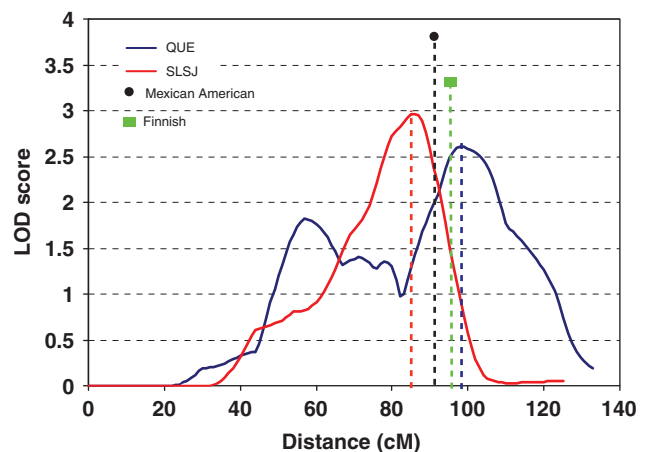


Figure 1 Multipoint LOD scores for HDL-C on chromosome 16 in SLSJ and QUE study samples, as well as previously reported findings in the region. The filled line and dashed line correspond to results from QUE samples near marker D16S515 at 16q23-24 and to results from SLSJ samples near marker D16S2624 at 16q22-23. ● reported a peak in Mexican-American families near markers D16S2624-D16S518 at 16q22.3-23.1.²⁶ ■ reported a peak in Finnish and Dutch families near marker D16S3096 at 16q22.1-24.3.²⁷

potential influence of regulatory variants on HDL-C levels using RT-PCR analyses of nine genes within the 6.6-Mb region, between the genes *CHST6* and *WVVOX*. Three of those genes (*ADAMTS18*, *CLEC3A* and *CNTNAP4*) were not significantly expressed in fibroblasts. The results of those genes that are expressed in fibroblasts are shown in Figure 3. The expression of *CHST6* and *KIAA1576* genes was found to be increased in cases when compared with controls ($P<0.001$). *CHST5* and *WVVOX* were not differentially expressed, whereas the *NUDT7* gene had a lower expression in cases compared with controls, but this was not statistically significant.

DISCUSSION

The inverse relation between HDL-C plasma levels and CAD has been well described in multiple epidemiological studies. Although environmental influences have an important role in determining serum HDL-C levels, genetics has a major role. This has been underscored by the discovery and identification of several genes involved in HDL metabolism.⁶ We performed two genome-wide scans in large families with probands ascertained for type II diabetes, CHD or low HDL-C. We obtained suggestive evidence for linkage²⁵ on chromosome 16 from both of these French-Canadian study samples. The region on chromosome 16q22-24 has been associated with HDL-C in several studies. Specifically, two previous studies, one involving Mexican Americans²⁶ and the other using Finnish and Dutch family samples,²⁷ provided evidence of linkage to this region on chromosome 16 and both were located <12 cM from the peaks observed in our analysis of Quebec families. Indeed, at least eight previous genome-wide scan studies have identified chromosomal loci for HDL-C on chromosome 16 with LOD scores > 1.²⁸⁻³⁵ In addition, Mehrabian *et al*³⁴ found a QTL for HDL-C on mouse chromosome 8 at genetic markers D8Mit12 (LOD=3.6) and D8Mit14 (LOD=3.5), which is syntenic to human chromosomal region 16q22.1-16q24. Moreover, in a study of mouse strains C57BL/6J and 129S1/SvImJ, Ishimori *et al*³⁵ identified a suggestive QTL on chromosome 8 (44 cM, LOD 2.6), which is consistent with the locus identified by Mehrabian *et al*.³⁴ Thus, these mouse QTLs for HDL-C provide additional confirmation of

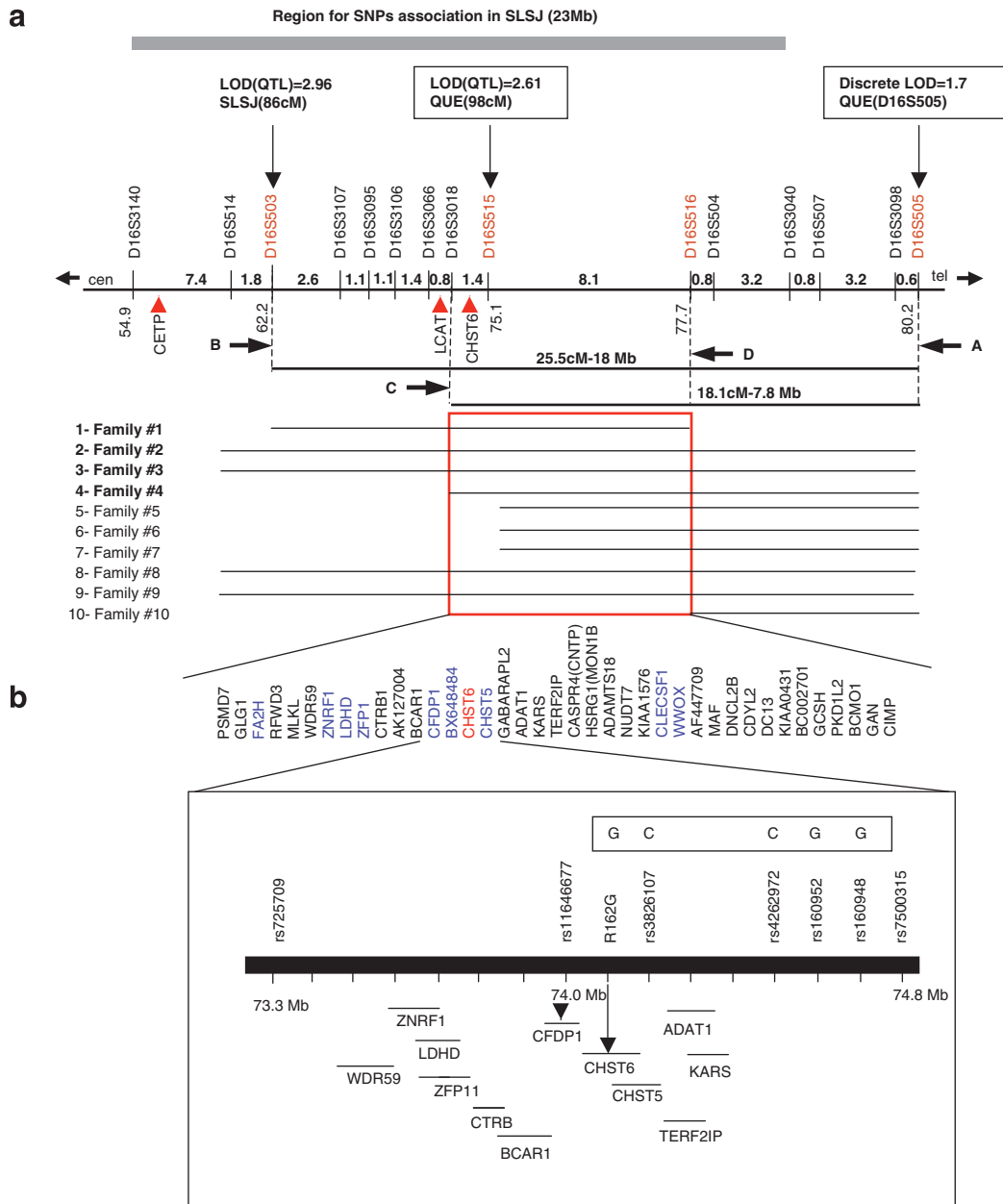


Figure 2 Linked region for HDL-C on chromosome 16q21-24. (a) The position of microsatellite markers with the interval distance (cM) between them is indicated and below the line are the actual physical locations (Mb). Red markers are genome-wide scan markers and the other markers are from subsequent fine mapping. Position of the LOD score peak is shown in the box. Family identification numbers are indicated on the left side and segregating haplotypes for each family are shown. The positions of *CETP*, *LCAT* and *CHST6* are indicated by arrowheads. The distance between arrows A and B corresponds to the initially identified linked region, which was narrowed down to the region between A and C. (b) Other known genes in the region of *CHST6* are shown. Inside the box is the five SNP haplotype that had an increased frequency. Families indicated in bold carry the R162G mutation in *CHST6*.

the involvement of this region of human chromosome 16 in HDL-C metabolism.

The *LCAT* and *CETP* genes, two well-known genes involved in HDL metabolism, are located near chromosome 16q23-24. *LCAT* mediates the formation of cholesteryl esters within HDL particles, which allows the maintenance of a gradient of free cholesterol from the plasma membrane to HDL particles and ensures a net movement of cholesterol from the cell to HDL. We sequenced the *LCAT* gene in the probands of QUE families and no coding variants were identified. *CETP* promotes the transfer of cholesteryl esters from HDL particles

to triglyceride-rich (VLDL and IDL) particles in exchange for triglycerides.³⁶⁻³⁸ Many association studies have revealed a significant association between *CETP* gene polymorphisms and levels of plasma HDL-C (reviewed in de Grooth *et al*³⁹ Thompson *et al*⁴⁰). Fine mapping of the 16q23-24 region with 12 additional microsatellite markers excluded both the *LCAT* and *CETP* genes from the linked region.

Our previous analysis of common variants in the region showed an association between an intronic SNP in the *WWOX* gene (rs2548861) and HDL-C in the QUE sample ($P=0.02$ using a dominant model).¹⁸

Table 3 Non-synonymous variants identified by sequencing

Gene	Amino acid change	rs number	Physical distance ^a
FA2H	Pro97Ala	rs35874850	73331496
RFWD3	Thr90Asn	rs8058922	73252580
RFWD3	Ile564Val	rs7193541	73222244
MLKL	Pro132Ser	rs35589326	73286763
MLKL	Met364Thr	rs34389205	73267111
CFDP1	Asn204Asp		73986529
BX648484	Met 44Ile		74043126
CHST6	Arg162Gly		74070744
CHST5	thr297met	rs3826107	74120831
CHST5	Leu 6Val	rs3743601	74147593
TERF21P	Lys324Glu	rs4888444	74247780
CASPR4	Leu279Val	rs34251012	75040248
CASPR4	Gln789Arg	rs12933808	75090084
CASPR4	Ala922Pro		75126951
CASPR4	Asp1158Glu	rs7202925	75144703
ADAMTS18	Tyr191His	rs11643211	75959046
ADAMTS18	Leu625Ile	rs11640912	75917420
ADAMTS18	Leu769Ile	rs9930984	75911474
ADAMTS18	Ala946Ser	rs12935394	75886491
ADAMTS18	Ser1080Arg	rs35478105	75882826
ADAMTS18	Ser1159Thr	rs3743749	75880736
NUDT7	Are183His	rs308925	76327335
CLEC3A	Gln197 Lys	rs2072663	76622234
WWOX	Ala179Thr	rs12918952	76978276
KIAA0431	Ser 600Thr		79635403
PKD1L2	Val20Ala	rs9924530	79811418
PKD1L2	Trp73Arg	rs9924371	79811260
PKD1L2	Gln120Leu	rs7191351	79807455
PKD1L2	Gly129Asp	rs7185774	79807428
PKD1L2	Lue173Ser	rs8060294	79806246
PKD1L2	Val183Ile	rs12933806	79806217
PKD1L2	Pro301Ala	rs11150370	79798601
PKD1L2	Lys416Gln	rs7194871	79790065
PKD1L2	Pro512Lue	rs7205673	79789776
PKD1L2	Leu711Pro	rs4889261	76770879
PKD1L2	Gly785Cys	rs9935113	79768997
PKD1L2	Arg849His	rs1869349	79766748
PKD1L2	Ala863Val	rs12596941	79766016
PKD1L2	Leu1036Pro	rs12597040	79757056
BCMO1	Arg267Ser	rs12934922	79859195
BCMO1	Ala379Val	rs7501331	79871997
PLCG2	His244Arg	rs11548656	80474413
PLCG2	Arg268Trp	rs17537869	80480314

^aPhysical distance is based on positions from the UCSC genome browser (<http://genome.ucsc.edu>) using the March 2006 assembly from NCBI's Build 36.1.

However, this SNP explains only 1.5% of the variance in HDL-C concentrations at the population level¹⁸ and is not likely to be a major determinant of HDL-C in families that show a strong pattern of segregation.

Using both microsatellite markers and SNPs, we were able to identify a minimal haplotype in one linked family of ~6.6 Mb on chromosome 16q23–24. It is noteworthy that a genome-wide association study from the Diabetes Genetics Initiative (<http://www.broad.mit.edu/diabetes/>) also showed evidence of association (P value < 0.0001) with six SNPs (rs8057477, rs4888535, rs6564359, rs7404386, rs3924497 and rs8049365) between positions 75307621–75376819, all residing within the linked region. We sequenced all coding and exon–intron boundaries of the 39 genes found within this

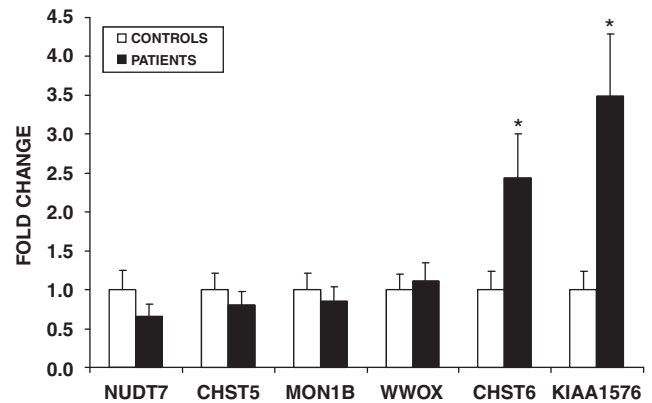


Figure 3 mRNA expression levels of select candidate genes found in the linked region. The mRNA expression of the genes (normalized to succinate dehydrogenase) from the skin fibroblasts of three probands are compared with the mean mRNA expression from three controls. The mRNA expression level was taken as the mean fold change ($2^{-\Delta\Delta C_t}$). The SEM values of triplicate experiments are shown. *indicates significant P value < 0.05.

region. One variant in *CHST6* showed evidence of co-segregation with low HDL-C and was investigated further. *CHST6* belongs to the sulfotransferase family. It catalyzes the transfer of sulfate to position 6 of nonreducing *N*-acetylglucosamine (GlcNAc) residues of keratin.⁴¹ It is located on the Golgi apparatus membrane and is mainly expressed in the cornea and brain but also in the spinal cord and trachea. Defects in the *CHST6* gene are the cause of macular corneal dystrophy (MIM: 217800). In this study, *CHST6* was considered a good candidate gene for low HDL-C, as ocular manifestations have been previously associated with other gene defects causing a low HDL-C phenotype (for example, *LCAT*, *APOA1*). However, there are currently no published data characterizing the lipid profile of patients with macular corneal dystrophy or directly relating the function of this gene to lipid metabolism.

We resequenced all other genes in the 6.6-Mb region and found no segregating variants but could not rule out variants found in the noncoding regions of these loci, which could have an effect on the transcription of these genes. Therefore, we examined the expression of genes in three probands of families linked to this region. An increase in the level of *CHST6* expression was found in these probands and we hypothesize that it might contribute to HDL metabolism. However, our RT-PCR results should be interpreted with caution as fibroblasts may not be the most appropriate cell model. Furthermore, we cannot exclude the possibility that the observed linkage and association at this locus may be caused by another nearby gene.

In a proteomic analysis of HDL by Rezaee *et al*⁴², C-type lectin domain family 3 member A (*CLEC3A*) has been identified as an HDL-associated protein. The gene coding for the *CLEC3A* protein also resides in the fine-mapped region on chromosome 16q23–24. Sequencing of this gene in probands revealed one nonsynonymous (Q197L) variant in one family. However, this variant did not segregate with low HDL-C in this family and, as *CLEC3A* is not expressed in fibroblasts, we were unable to investigate the differential expression level of its mRNA.

In summary, we have fine-mapped a region on chromosome 16q23–24 that is likely to harbor a gene for low plasma levels of HDL-C. Recently, genome-wide association studies have been carried out for lipid traits, including HDL-C.^{43,44} Although these studies have revealed polymorphisms associated with the variation in HDL-C level in a number of European populations, such variants do not

account for a significant fraction of their variance and are not likely to be the cause of extreme levels of HDL-C that segregate in some families. More likely, multiple rare genetic variants that have a large effect lead to the extreme levels of HDL-C observed in families. These may sometimes be found in the same genes as the common variants that affect HDL-C, but perhaps not always. Therefore, the identification of the causal genes for more severe forms of low HDL-C may contribute to our understanding of the regulation of HDL-C levels and subsequently provide new therapeutic targets for CAD prevention. With recent advances in sequencing technology, it should be feasible to resequence this particular region in its entirety to identify the variants that affect HDL-C levels.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank all patients and their families for participating in the studies. SNP genotyping in the SLSJ study sample was performed by Katia Desbiens. This work was supported by Grant CIHR MOP 62834 from the Canadian Institutes of Health Research (CIHR). Jacques Genest holds the McGill University-Novartis Chair in Medicine. Zari Dastani is a recipient of a Doctoral research award from the Heart and Stroke Foundation of Canada (HSFC). Michel Marcil is a research scholar from the *Fonds de la recherche en santé du Québec* (FRSQ). Päivi Pajukanta is supported by NIH grants HL 28481, HL08276 and HL095056.

- Murray CJ, Lopez AD: Alternative projections of mortality and disability by cause 1990–2020: Global Burden of Disease Study. *Lancet* 1997; **349**: 1498–1504.
- Yusuf S, Reddy S, Ounpuu S *et al*: Global burden of cardiovascular diseases: Part II: variations in cardiovascular disease by specific ethnic groups and geographic regions and prevention strategies. *Circulation* 2001; **104**: 2855–2864.
- Despres JP, Lemieux I, Dagenais GR *et al*: HDL-cholesterol as a marker of coronary heart disease risk: the Quebec cardiovascular study. *Atherosclerosis* 2000; **153**: 263–272.
- Gordon T, Castelli WP, Hjortland MC *et al*: High density lipoprotein as a protective factor against coronary heart disease. The Framingham Study. *Am J Med* 1977; **62**: 707–714.
- Assmann G, Schulte H, von Eckardstein A *et al*: High-density lipoprotein cholesterol as a predictor of coronary heart disease risk. The PROCAM experience and pathophysiological implications for reverse cholesterol transport. *Atherosclerosis* 1996; **124** (Suppl): S11–S20.
- Dastani Z, Engert JC, Genest J *et al*: Genetics of high-density lipoproteins. *Curr Opin Cardiol* 2006; **21**: 329–335.
- Breslow JL: Familial disorders of high density lipoprotein metabolism. In: Scriver CR, Beaudet AL, Sly WS, Valle D (eds): *The Molecular and Metabolic Basis of Inherited Diseases*. New York: McGraw-Hill, 2001.
- Cohen JC, Kiss RS, Pertsemlidis A *et al*: Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science* 2004; **305**: 869–872.
- Boes E, Coassin S, Kollerits B *et al*: Genetic-epidemiological evidence on genes associated with HDL cholesterol levels: a systematic in-depth review. *Exp Gerontol* 2008; **44**: 136–160.
- Alrasadi K, Ruel IL, Marcil M *et al*: Functional mutations of the ABCA1 gene in subjects of French-Canadian descent with HDL deficiency. *Atherosclerosis* 2005; **188**: 281–291.
- Dastani Z, Dangoisse C, Boucher B *et al*: A novel nonsense apolipoprotein A-I mutation (apoA-I(E136X)) causes low HDL cholesterol in French Canadians. *Atherosclerosis* 2006; **185**: 127–136.
- NIH Washington DC: *Lipid Research Clinics Population Studies Databook*. Washington, DC: Department of health and human services, Public Health Services. NIH publication 80–1527 1980. pp 28–81.
- Marcil M, Yu L, Krimbou L *et al*: Cellular cholesterol transport and efflux in fibroblasts are abnormal in subjects with familial HDL deficiency. *Arterioscler Thromb Vasc Biol* 1999; **19**: 159–169.
- Weber M, McNicoll S, Marcil M *et al*: Metabolic factors clustering, lipoprotein cholesterol, apolipoprotein B, lipoprotein (a) and apolipoprotein E phenotypes in premature coronary artery disease in French Canadians. *Can J Cardiol* 1997; **13**: 253–260.
- Rioux JD, Stone VA, Daly MJ *et al*: Familial eosinophilia maps to the cytokine gene cluster on human chromosomal region 5q31-q33. *Am J Hum Genet* 1998; **63**: 1086–1094.
- Laitinen T, Daly MJ, Rioux JD *et al*: A susceptibility locus for asthma-related traits on chromosome 7 revealed by genome-wide scan in a founder population. *Nat Genet* 2001; **28**: 87–91.
- Abecasis GR, Cherny SS, Cookson WO *et al*: GRR: graphical representation of relationship errors. *Bioinformatics* 2001; **17**: 742–743.
- Lee JC, Weissglas-Volkov D, Kyttala M *et al*: WW-domain-containing oxidoreductase is associated with low plasma HDL-C levels. *Am J Hum Genet* 2008; **83**: 180–192.
- Abecasis GR, Cherny SS, Cookson WO *et al*: Merlin – rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 2002; **30**: 97–101.
- Barrett JC, Fry B, Maller J *et al*: Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; **21**: 263–265.
- Heath SC: Markov chain Monte Carlo segregation and linkage analysis for oligogenic models. *Am J Hum Genet* 1997; **61**: 748–760.
- Sobel E, Lange K: Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. *Am J Hum Genet* 1996; **58**: 1323–1337.
- Lange K, Cantor RM, Horvath S *et al*: Mendel version 4.0: a complete package for the exact genetic analysis of discrete traits in pedigree and population data sets. *Am J Hum Genet* 2001; **504**.
- Kong A, Gudbjartsson DF, Sainz J *et al*: A high-resolution recombination map of the human genome. *Nat Genet* 2002; **31**: 241–247.
- Lander E, Kruglyak L: Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 1995; **11**: 241–247.
- Mahoney MC, Almasy L, Rainwater DL *et al*: A quantitative trait locus on chromosome 16q influences variation in plasma HDL-C levels in Mexican Americans. *Arterioscler Thromb Vasc Biol* 2003; **23**: 339–345.
- Pajukanta P, Allayee H, Krass KL *et al*: Combined analysis of genome scans of Dutch and Finnish families reveals a susceptibility locus for high-density lipoprotein cholesterol on chromosome 16q. *Am J Hum Genet* 2003; **72**: 903–917.
- Shearman AM, Ordoval JM, Cupples LA *et al*: Evidence for a gene influencing the TG/HDL-C ratio on chromosome 7q32.3-qter: a genome-wide scan in the Framingham Study. *Hum Mol Genet* 2000; **9**: 1315–1320.
- Amos CI, Wilson AF, Rosenbaum PA *et al*: An approach to the multivariate analysis of high-density-lipoprotein cholesterol in a large kindred: the Bogalusa Heart Study. *Genet Epidemiol* 1986; **3**: 255–267.
- Aouizerat BE, Allayee H, Cantor RM *et al*: Linkage of a candidate gene locus to familial combined hyperlipidemia: lecithin:cholesterol acyltransferase on 16q. *Arterioscler Thromb Vasc Biol* 1999; **19**: 2730–2736.
- Soro A, Pajukanta P, Lilja HE *et al*: 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* 2002; **70**: 1333–1340.
- Gagnon F, Jarvik GP, Motulsky AG *et al*: Evidence of linkage of HDL level variation to APOC3 in two samples with different ascertainment. *Hum Genet* 2003; **113**: 522–533.
- Yip AG, Ma Q, Wilcox M *et al*: Search for genetic factors predisposing to atherogenic dyslipidemia. *BMC Genet* 2003; **4** (Suppl 1): S100.
- Mehrabian M, Castellani LW, Wen PZ *et al*: Genetic control of HDL levels and composition in an interspecific mouse cross (CAST/Ei x C57BL/6J). *J Lipid Res* 2000; **41**: 1936–1946.
- Ishimori N, Li R, Kelmenson PM *et al*: Quantitative trait loci analysis for plasma HDL-cholesterol concentrations and atherosclerosis susceptibility between inbred mouse strains C57BL/6J and 129S1/SvImJ. *Arterioscler Thromb Vasc Biol* 2004; **24**: 161–166.
- Thompson JF, Durham LK, Lira ME *et al*: CETP polymorphisms associated with HDL cholesterol may differ from those associated with cardiovascular disease. *Atherosclerosis* 2005; **181**: 45–53.
- Kondo I, Berg K, Drayna D *et al*: DNA polymorphism at the locus for human cholesteryl ester transfer protein (Cetp) is associated with high-density lipoprotein cholesterol and apolipoprotein levels. *Clin Genet* 1989; **35**: 49–56.
- Inazu A, Jiang XC, Haraki T *et al*: Genetic cholesteryl ester transfer protein deficiency caused by two prevalent mutations as a major determinant of increased levels of high density lipoprotein cholesterol. *J Clin Invest* 1994; **94**: 1872–1882.
- de Grooth GJ, Klerkx AH, Stroes ES *et al*: A review of CETP and its relation to atherosclerosis. *J Lipid Res* 2004; **45**: 1967–1974.
- Thompson JF, Lira ME, Durham LK *et al*: Polymorphisms in the CETP gene and association with CETP mass and HDL levels. *Atherosclerosis* 2003; **167**: 195–204.
- Hasegawa N, Torii T, Kato T *et al*: Decreased GlcNAc 6-O-sulfotransferase activity in the cornea with macular corneal dystrophy. *Invest Ophthalmol Vis Sci* 2000; **41**: 3670–3677.
- Rezaee F, Casetta B, Levels JH *et al*: Proteomic analysis of high-density lipoprotein. *Proteomics* 2006; **6**: 721–730.
- Willer CJ, Sanna S, Jackson AU *et al*: Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet* 2008; **40**: 161–169.
- Kathiresan S, Melander O, Guiducci C *et al*: Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat Genet* 2008; **40**: 189–197.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (<http://www.nature.com/ejhg>)