Multiple genetic variants along candidate pathways influence plasma high-density lipoprotein cholesterol concentrations[S]

Yingchang Lu,[1,*,†] Martijn E. T. Dollé,[†] Sandra Imholz,[†] Ruben van 't Slot,[§] W. M. Monique Verschuren,[†] Cisca Wijmenga,[§] Edith J. M. Feskens,[*] and Jolanda M. A. Boer[†]

Division of Human Nutrition,* Wageningen University, Wageningen, The Netherlands; National Institute for Public Health and the Environment (RIVM),† Bilthoven, The Netherlands; and Complex Genetics Section,§ Department of Biomedical Genetics, University Medical Center Utrecht, Utrecht, The Netherlands

Abstract The known genetic variants determining plasma HDL cholesterol (HDL-C) levels explain only part of its variation. Three hundred eighty-four single nucleotide polymorphisms (SNPs) across 251 genes based on pathways potentially relevant to HDL-C metabolism were selected and genotyped in 3,575 subjects from the Doetinchem cohort, which was examined thrice over 11 years. Three hundred fifty-three SNPs in 239 genes passed the quality-control criteria. Seven SNPs [rs1800777 and rs5882 in cholesteryl ester transfer protein (CETP); rs3208305, rs328, and rs268 in LPL; rs1800588 in LIPC; rs2229741 in NRIP1] were associated with plasma HDL-C levels with false discovery rate (FDR) adjusted q values (FDR_q) < 0.05. Five other SNPs (rs17585739 in SC4MOL, rs11066322 in PTPN11, rs4961 in ADD1, rs6060717 near SCAND1, and rs3213451 in MBTPS2 in women) were associated with plasma HDL-C levels with FDR_q between 0.05 and 0.2. Two less well replicated associations (rs3135506 in APOA5 and rs1800961 in HNF4A) known from the literature were also observed, but their significance disappeared after adjustment for multiple testing (P = 0.008, FDR_q = 0.221 for rs3135506; P = 0.018, FDR_q = 0.338 for rs1800961, respectively). JLR **In addition to replication of previous results for candidate genes (CETP, LPL, LIPC, HNF4A, and APOA5), we found interesting new candidate SNPs (rs2229741 in NRIP1, rs3213451 in MBTPS2, rs17585739 in SC4MOL, rs11066322 in PTPN11, rs4961 in ADD1, and rs6060717 near SCAND1) for plasma HDL-C levels that should be evaluated further.—**Lu, Y., M. E. T. Dollé, S. Imholz, R. van 't Slot, W. M. M. Verschuren, C. Wijmenga, E. J. M. Feskens, and J. M. A. Boer. **Multiple genetic variants along candidate pathways influence plasma high-density lipoprotein cholesterol concentrations. *J. Lipid Res.* 2008.** 49: **2582–2589.**

Supplementary key words single nucleotide polymorphism • pathway-driven approach • random coefficient model

Numerous clinical and epidemiological studies have demonstrated an inverse and independent association between plasma concentrations of HDL cholesterol (HDL-C) and the risk of coronary heart disease (1). The most popular mechanistic explanation has been that HDL functions in reverse cholesterol transport, removing cholesterol from peripheral tissues and delivering it to the liver for biliary excretion and to steroidogenic organs for steroid hormones synthesis (2). Although efflux of cholesterol from macrophages represents only a tiny fraction of overall cellular cholesterol efflux, it is in fact, the most important with regard to antiatherogenic effects (3). More recently, a variety of other functions of HDL have been described, primarily based on in vitro assays, including anti-inflammatory, antioxidant, antithrombotic, and nitric oxide-inducing mechanisms that could also contribute to its antiatherogenic effects (4, 5).

Current evidence suggests that blood lipids are complex phenotypes, influenced by environmental and genetic factors. It has been well established that body weight (6), current smoking habits (7), exercise (8), alcohol use (9), and dietary fat intake (10) influence plasma HDL-C levels. Several twin and family studies indicate that heritability estimates for blood levels of HDL-C range from 24% to 83%, with most studies in the 40% to 60% range (11). Mutations in genes including ABC transporter A1 (ABCA1), apolipoprotein A1, and lecithin cholesterol transferase (LCAT) are implicated in rare mendelian forms of HDL deficiency and familial hypoalphalipoproteinemia (12, 13). A mutation in the cholesteryl ester transfer protein gene (CETP)

The Doetinchem Cohort Study was financially supported by the Ministry of Health, Welfare and Sport of The Netherlands and the National Institute for Public Health and the Environment. Y. Lu is funded by grant 2006B195 of the Netherlands Heart Foundation.

Manuscript received 6 May 2008 and in revised form 27 June 2008.

Published, JLR Papers in Press, July 25, 2008.
DOI 10.1194/jlr.M800232-JLR200

[1] To whom correspondence should be addressed:
e-mail: kevin.lu@wur.nl

[S] The online version of this article (available at http://www.jlr.org) contains supplementary data in the form of two tables.

found in two Japanese siblings causes CETP deficiency and extremely elevated levels of HDL-C (14). Findings from candidate gene studies suggest that genetic polymorphisms, including single nucleotide polymorphisms (SNPs) located in CETP, lipoprotein lipase (LPL), hepatic lipase (LIPC) (15, 16), and apolipoprotein A1/C3 (17) genes, are important sources of genetically determined variation in plasma HDL-C. However, they explain only a small part of the variation. Most of the DNA sequence variants that contribute to variation in plasma HDL-C levels in the general population are still largely unknown.

Since low plasma HDL-C levels always cluster with other kinds of dyslipidemia and also insulin resistance, in this study, we extended the candidate gene scope and emphasized on the intricate links within lipid metabolic pathways and also between glucose and lipid metabolic pathways (18). We postulated that more genuine signals from the genes that are involved in the metabolism of HDL-C could be captured. The purpose of this study was twofold: first, to survey genetic variants in a large number of candidate genes in relation to plasma HDL-C levels; and second, to investigate whether these associations could be modified by traditional environmental factors, such as drinking, smoking, and dietary fat intake.

METHODS

Study population

Our study was conducted in Doetinchem, a town in a rural area in the east of the Netherlands. Random samples were taken from the municipal population register between 1987 and 1991. A total of 12,404 inhabitants aged 20–59 years were willing to participate and underwent a first measurement (response 62%). A subsample was reinvited for a second measurement between 1993 and 1997, and 6,100 participants were re-examined (response 79%). Between 1998 and 2002, a third measurement took place and 4,917 participants were reassessed (response 75%). For each survey, approval had been obtained from the Medical Ethics Committee. Informed consent was obtained from all participants. Overall, the Doetinchem Cohort consisted of 4,662 persons for which three measurements were available. Participants who changed their smoking habits (n = 872), who had missing data on smoking status (n = 11), or who were pregnant at the time of measurement (n = 122) were excluded. Finally, 3,779 participants met the inclusion criteria of this study. The design and detailed methods have been reported earlier (19). The subjects were surveyed on demographic, anthropometric, lifestyle information (smoking, alcohol use, physical activity, and dietary habits), disease history, and medications by questionnaires. A nonfasting blood sample was taken from all participants; fractionated into blood plasma, white blood cells, and erythrocytes; and subsequently stored. A validated semiquantitative food frequency questionnaire was used in the second and third surveys to assess the habitual consumption of 178 food items during the previous year (20). Nutrient and energy intake were quantified for each individual using an updated computerized Dutch food composition table.

Laboratory assessment of HDL-C

Plasma HDL-C levels were assayed in the Lipid Reference Laboratory (LRL) of the University Hospital Dijkzigt in Rotterdam using standardized enzymatic methods within 3 weeks after storage. The LRL Rotterdam is a permanent member of the International Cholesterol Reference Method Laboratory Network. It has been standardized to the Centers for Diseases Control and Prevention (CDC) through participating in the CDC/National Heart Lung and Blood Institute Lipid Standardization Program. HDL-C was determined in the supernatant after precipitation of apoB-containing lipoproteins with phophotungstic acid/$MgCl_2$ (Boehringer). The accuracy of HDL-C determination fulfilled National Cholesterol Education Program recommendations throughout the entire period. From December 1999, a homogen liquid method was used. From March 2002, the homogen liquid (2^{nd} generation) was used. Whenever a new method was introduced, a careful calibration was performed to make sure there was no difference between different assaying methods.

Gene and SNP selection and genotyping

Candidate genes were selected by a pathway-driven approach based on literature, with emphasis on regulatory pathways that control fatty acid, glucose, cholesterol, and bile salt homeostasis (for a recent review see reference 18). The selection procedure started from the master regulator genes encoding nuclear receptors (PPARs, LXR, and FXR) and transcription factors (SREBPs), and continued by selecting their coactivators, corepressors, and target genes. In addition, hormonal receptors (insulin receptor) and their downstream signaling proteins were selected. Furthermore, several candidate genes, described in literature to be associated with blood lipids or blood pressure, were added. The selection resulted in 251 candidate genes.

Where possible eligible SNPs for these genes were selected based on published associations with any disorder or functional parameter, using databases from NCBI (PubMed, Gene, and SNP), the Genetic Association Database (21) from CDC, and SNPper (22). In part limited by various constrains of the Golden Gate genotyping assay, a total of 153 SNPs across 91 genes were finally selected from the literature. Subsequently, we used the web-based program SNPselector (23) to query all 251 genes for potential candidate SNPs. We performed a "SNPs by gene" search, including 5kb 5′ and 1kb 3′ flanking sequences, with slightly modified default ranking settings. First, SNPs located in repeat regions were excluded (repeat_score $>$ 0), to avoid potential genotyping difficulties. Second, Caucasian minor allele frequencies (MAF_Caucasian $>$ 0) had to be available. Third, SNPs in predicted transcription factor binding sites (regulatory = TFBS) were preferred. Finally, subsequent ranking was based on the highest function score (function_score), followed by the highest regulatory score (regulatory_score), discriminating between SNPs that might affect gene transcript structure or protein product, and the regulatory potential of the SNP, respectively. A total of 226 SNPs across 178 genes, selected with SNPselector, were included on the Illumina array. Four of the 251 candidate genes remained without eligible SNPs in the SNP search described above. For each of these genes (SCAP, ACSL1, CEBPA, E2F4), a single SNP was handpicked based on Caucasian allele frequency, SNP location, and validation information in NCBI SNP. Together a final set of 383 SNPs across 251 candidate genes, with 1 to 7 SNPs per gene, passed the Illumina Assay Design Tool and were included. Besides the 383 selected SNPs, one Y chromosome marker was determined to serve as gender control. The complete list of SNPs and additional data for all SNPs are provided in supplementary Table I.

Genomic DNA was extracted from the buffy coat fraction with a salting out method. A total of 139 subjects were not eligible for genotyping, mainly because DNA extraction did not succeed or DNA was not available. For 3,639 subjects, high throughput SNP genotyping was performed with the Illumina Golden Gate assay using the Sentrix Array Matrix platform (Illumina Inc., San Diego,

CA). Illumina GenCall software (version 6.1.3.28) was used for automated genotype clustering and calling. Genotyping failed for 43 subjects due to an absence of any signal. In addition, 21 subjects were excluded because of discordance for the gender control. For 28 SNPs the genotype calling did not succeed due to low signal (n = 11), overlap between the genotype clusters (n = 13), multiple genotype clusters (n = 3), or scattering of clusters (n = 1). For the 8 SNPs, most out of Hardy Weinberg Equilibrium (HWE), genotyping of the particular SNP was verified in a random sample (n = 96) using Taqman, pyrosequencing, or sequencing. Two of these failed the verification and were excluded. Furthermore, genotype calling was not completely convincing for 42 SNPs. The latter were included in the analysis, but when an association was found with body mass index (BMI), waist circumference, or lipids levels, genotyping was verified. All SNPs passed the verification. Finally, for 3,575 participants, data was available of 353 SNPs in 239 genes.

Statistical analysis

All analyses were performed with SAS version 9.1 software (SAS Institute, Cary, NC). Paired Student's *t*-tests and Chi-square tests were used for comparisons of means and proportions between measurement rounds. If data were not normally distributed, the Wilcoxon Signed-Rank test was used to make the comparison. Distributions of genotypes were tested for deviation from HWE by Chi-square analysis (PROC ALLELE). Random coefficient models (multilevel modeling) were used to study the relationship between SNPs and repeated measurements of plasma HDL-C levels (PROC MIXED). Men and women were analyzed separately for five X-linked SNPs (rs2073115, rs3213451, rs5969919, rs1403543, and rs3048). When the overall difference was statistically significant, the Tukey-Kramer method was used to identify significant differences between the genotype groups. The exact follow-up time in years was put into the model as a continuous variable. To adjust for the potential confounding effects and to improve model fitting, age, age^2, sex, current smoking (yes or no), alcohol use, and BMI were added to the model as covariates. The intercept and time were treated as random effects allowing unique baseline levels and unique changes of plasma HDL-C concentration over time for each individual. Potential gene-environment interactions in relation to plasma HDL-C level were also explored using MIXED models by including interaction terms into the model. All reported *P* values were two-tailed, and statistical significance without adjustment for multiple testing was defined at the $\alpha = 0.05$ level. The false discovery rate (FDR) was applied to take multiple testing into account (PROC MULTTEST). To date, there is no conventional q value threshold to categorize a discovery as significant. As in similar research, a q value threshold of 0.20 was used to define significance (24). Certain SNPs in previously confirmed candidate genes affecting plasma HDL-C levels were also reported ($P < 0.05$) although their q values exceeded 0.2.

RESULTS

The mean age of the subjects at the first survey was 40.8 years, ranging from 20 to 60 years (**Table 1**). More people took lipid-lowering medication in the second and the third round compared with the first examination. The average plasma HDL-C level increased from the first survey (1.26 ± 0.31 mmol/L) to the second survey (1.38 ± 0.38 mmol/L), but decreased slightly thereafter in the third survey (1.36 ± 0.39 mmol/L). Correlation coefficients between the three measurements of plasma HDL-C levels ranged from 0.76 to 0.82.

Twenty-eight SNPs were found to be significantly associated with variations in plasma HDL-C levels ($P < 0.05$, **Table 2**) after adjustment for age, age^2, sex, current smoking habit, alcohol use, and BMI. Eleven of them had a FDR_q value < 0.2 after adjustment for multiple testing; seven SNPs (rs1800777 and rs5882 in CETP; rs3208305, rs328 and rs268 in LPL; rs1800588 in LIPC, and rs2229741 in NRIP1) had a FDR_q value < 0.05, while four more SNPs (rs17585739 in SCMOL, rs11066322 in PTPN11, rs4961 in ADD1, and rs6060717 near SCAND1) had a FDR_q value between 0.05 and 0.2. Two of the SNPs for which the significance disappeared after adjustment for multiple testing concerned less well replicated associations in the literature (rs3135506 in APOA5, $P = 0.008$ and FDR_q = 0.221; rs1800961 in HNF4A, $P = 0.018$, and FDR_q = 0.338, respectively). With respect to the SNPs located on the X chromosome, a SNP in MBTPS2 (rs3213451) was associated with plasma HDL-C levels, but only in female subjects ($P = 0.02$ and FDR_q = 0.12). The *P* values for the above top SNPs (FDR_q value < 0.2) did not change

TABLE 1. General characteristics of the study population (n = 3575, 1,710 men and 1,865 women)

	First survey	Second survey	Third survey
Time since baseline (yr)[a]	—	6.0 (5.9–6.1)	11.0 (10.9–11.1)
Age (yrs)	40.8 ± 9.8	46.7 ± 9.8	51.7 ± 9.8
BMI (kg/m^2)[b]	24.5 ± 3.3	25.4 ± 3.6	26.1 ± 3.8
HDL cholesterol (mmol/L)[b]	1.26 ± 0.31	1.38 ± 0.38	1.36 ± 0.39
Lipid-lowering medication (%)[b]	6 (0.17)	52 (1.45)	149 (4.17)
Current smokers (%)	915 (25.6)	916 (25.6)	916 (25.6)
Alcohol consumption (glass/day)[a,b]	0.57 (0–1.43)	0.57 (0–1.57)	0.71 (0–1.71)
Dietary intake	Na		
Energy intake (kj/day)[b]	—	9,403 ± 2,570	9,081 ± 2,415
Total fat (% energy)[b]	—	35.0 ± 4.7	34.8 ± 4.8
Saturated fat (% energy)[b]	—	14.6 ± 2.3	14.4 ± 2.3
Monounsaturated fat (% energy)	—	13.3 ± 2.1	13.2 ± 2.2
Polyunsaturated fat (% energy)	—	6.8 ± 1.6	6.9 ± 1.6
Cholesterol (mg/day)[b]	—	240.1 ± 82.8	229.8 ± 79.6

BMI, body mass index; Na: not available. Data are expressed as mean ± standard deviation or N (%) unless otherwise indicated.

[a] Median (Q1–Q3).

[b] Significant difference between the surveys, $P < 0.05$.

TABLE 2. Association between SNPs and plasma HDL cholesterol levels

Nearest Gene	SNP	MAF	P^a	FDR_q[b]
LPL	rs3208305	0.30	<0.0001	<0.0001
LPL	rs328	0.11	<0.0001	<0.0001
CETP	rs1800777	0.03	<0.0001	<0.0001
CETP	rs5882	0.31	<0.0001	0.0002
NRIP1	rs2229741	0.41	<0.0001	0.0013
LIPC	rs1800588	0.20	<0.0001	0.0025
LPL	rs268	0.02	0.0003	0.0164
MBTPS2[c]	rs3213451	0.35	0.0231	0.1153
SC4MOL	rs17585739	0.06	0.0027	0.1190
PTPN11	rs11066322	0.19	0.0031	0.1196
ADD1	rs4961	0.21	0.0036	0.1264
SCAND1	rs6060717	0.20	0.0061	0.1924
PIAS1	rs1489599	0.43	0.0082	0.2211
APOA-V	rs3135506	0.08	0.0083	0.2211
NCOR2	rs2229840	0.17	0.0099	0.2451
LPL	rs1059507	0.16	0.0114	0.2645
MAP2K1[d]	rs17586159	0.02	0.0132	0.2862
PRKCA	rs7210446	0.40	0.0140	0.2862
HNF4A	rs1800961	0.04	0.0175	0.3383
NRIP1	rs2229742	0.12	0.0196	0.3532
ACADM	rs11549022	0.30	0.0203	0.3532
PRKCA	rs17633437	0.37	0.0234	0.3886
ILK	rs2288283	0.13	0.0254	0.4014
ADIPOQ	rs17300539	0.07	0.0312	0.4722
MYBBP1A	rs751670	0.16	0.0383	0.4978
MLYCD	rs11649200	0.17	0.0384	0.4978
NDN	rs850791	0.09	0.0386	0.4978
SAH	rs5716	0.08	0.0386	0.4978
NCOA2	rs10112498	0.39	0.0444	0.5518

SNP, single nucleotide polymorphism; MAF, minor allele frequency; FDR, false discovery rate.

[a] Adjusted for age, age^2, sex, current smoking habits, alcohol use, and body mass index.

[b] Adjusted for multiple testing (348 SNPs) with the FDR method.

[c] Statistical analysis was conducted in women only and was adjusted for multiple testing (5 SNPs) with FDR method.

[d] There was only one subject with the genotype AA. This subject was incorporated into the genotype GA group during the statistical analysis.

too much after we excluded the subjects who had started lipid-lowering medication in each round of survey (data not shown).

Compared with the ancestral allele, rs1800777 in CETP, rs3208305 in LPL, rs11066322 in PTPN11, rs4961 in ADD1, rs3135506 in APOA5, and rs1800961 in HNF4A were associated with decreased plasma HDL-C levels, while rs268 and rs328 in LPL, rs5882 in CETP, rs2229741 in NRIP1, rs1800588 in LIPC, rs17585739 in SC4MOL, rs6060717 near SCAND1, and rs3213451 in MBTPS2 were associated with increased plasma HDL-C levels (**Table 3** and supplementary Table II). Rs328 and rs3208305 in LPL, rs1800588 in LIPC, and rs17585739 in SC4MOL had big effects on plasma HDL-C levels compared with other SNPs ($\geq$ 0.1 mmol/L between the two homozygote genotypes).

Considering that variation in plasma HDL-C level also relates to lifestyle factors, including smoking, body weight, alcohol use, and dietary fat intake, we further examined whether these factors could modulate the observed associations between the 14 SNPs identified above (SNPs with FDR_q $<$ 0.2, rs3135506 and rs1800961) and plasma HDL-C level. For dietary fat intake (saturated fat, monounsaturated fat, and polyunsaturated fat), only information in the second and third surveys was included in the analysis, as information in the first survey is less valid. Interactions with the other lifestyle factors were based on three rounds of measurements. Before adjustment for multiple testing, there were possible interactions between NRIP1 genotype (rs2229741) and BMI, APOA5 genotype (rs3135506) and alcohol use (low vs. high), and APOA5 genotype (rs3135506) and saturated fat intake (low vs. high) (P for interaction = 0.017, 0.023, and 0.042, respectively). However, the significance of the interactions disappeared after adjustment for multiple testing (FDR_q $>$ 0.2).

DISCUSSION

In this longitudinal genetic association study, with a pathway-driven approach, we confirmed several genetic variants in genes along known HDL metabolic pathway to be associated with plasma HDL-C levels. At the same time, we found some new additional candidate genes whose genetic variants may be associated with plasma HDL-C levels. For these genes, the results need to be replicated in other population studies.

Genetic variants among CETP, LPL, and LIPC have previously been found to be associated with plasma HDL-C levels in candidate gene studies (15, 16). Recently, in a series of genome-wide association analyses (25–27), the associations between genetic variants from these genes and plasma HDL-C levels were also confirmed (rs711752, rs7205804, rs5880 and rs3764261 in CETP; rs326 and rs12678919 in LPL; rs11858164 and rs10468017 in LIPC). The influence from these genetic variants (rs1800777 and rs5882 in CETP; rs3208305, rs328 and rs268 in LPL; and rs1800588 in LIPC) on plasma HDL-C levels was once again observed in our study. These above-mentioned significant loci in LPL or LIPC are in fact located in one haplotype block (28), which makes them more frequently inherited together. All the loci in CETP are located in two connected haplotype blocks (28). The roles these genes play in the metabolic process of HDL-C have well been elucidated either from experimentation in cell culture and in animal models or from observational and interventional studies in humans (3).

Two less well replicated associations (rs1800961 in HNF4A and rs3135506 in APOA5) with plasma HDL-C levels were also observed in our study, although their statistical significance disappeared after adjustment for multiple testing. In line with our findings, serum HDL-C concentration was lower in subjects with the T130I (rs1800961) mutation compared with those without this mutation in a study among Japanese patients with late-onset type 2 diabetes (29). A large European Caucasian collection of MODY (maturity onset diabetes of the young) patients due to HNF4A mutations (30) also presented with lower fasting apolipoprotein A-II, A-I, and HDL-C levels than control subjects. Hepatic nuclear factor 4 α (HNF4A) regulates hepatic expression of a number of genes associated with lipoprotein metabolism, including genes encoding apoA-I, A-II, A-IV, B, C-II, C-III, E, microsomal triglyceride transfer protein, cholesterol 7 α-hydroxylase, SR-BI, and PPARα

TABLE 3. Distribution of plasma HDL cholesterol levels according to the genotypes of 14 SNPs significantly associated with plasma HDL cholesterol levels

Nearest Gene	SNP	Nucleotide Substitution[a]	*P*(HWE)	Frequency	HDL-C (mmol/L)
LPL	rs3208305	26696 T→A	0.97		
	TT			315	1.39a
	AT			1490	1.32b
	AA			1767	1.27c
	rs328 (Ser447Ter)	22772 C→G	0.90		
	GG			40	1.43a
	CG			683	1.36a
	CC			2851	1.28b
	rs268 (Asn291Ser)	16577 A→G	0.76		
	GG			2	1.39
	AG			148	1.21a
	AA			3423	1.30b
CETP	rs1800777(Arg468Gln)	21427 G→A[b]	0.60		
	AA			3	1.22
	AG			233	1.20a
	GG			3332	1.31b
	rs5882 (Ile422Val)	20200 A→G	0.20		
	GG			328	1.35a
	AG			1563	1.31a
	AA			1682	1.28b
NRIP1	rs2229741	224 A→G	0.05		
	AA			634	1.26a
	AG			1674	1.30b
	GG			1263	1.32b
LIPC	rs1800588	−514 C→T	0.63		
	TT			149	1.39a
	CT			1134	1.30b
	CC			2267	1.29b
SC4MOL	rs17585739	3968 G→A	0.55		
	AA			15	1.52a
	AG			401	1.32b
	GG			3159	1.30b
PTPN11	rs11066322	65613 A→G	0.02		
	AA			150	1.36a
	AG			1049	1.28b
	GG			2342	1.30b
ADD1	rs4961 (Gly460Trp)	29064 G→T	0.67		
	TT			146	1.27
	TG			1172	1.28a
	GG			2252	1.31b
SCAND1	rs6060717	−2441 C→T	0.81		
	CC			134	1.23a
	CT			1127	1.31b
	TT			2309	1.30b
APOA-V	rs3135506 (Ser19Trp)	169 G→C	0.66		
	CC			19	1.24
	CG			508	1.27a
	GG			3042	1.31b
HNF4A	rs1800961(Thr117/139Ile)	12351 C→T	0.44		
	TT			3	1.17
	CT			250	1.26a
	CC			3321	1.30b
MBTPS2[c]	rs3213451	3581 A→G	0.36		
	GG			214	1.49a
	AG			860	1.43b
	AA			787	1.45

SNP, single nucleotide polymorphism; *P*(HWE), *P* value from x^2 test of Hardy-Weinberg Equilibrium; HDL-C, high density lipoprotein cholesterol. Genotypes with different letters marked with (a, b, and c) differ significantly from each other in plasma HDL cholesterol levels.

[a] Human mutation nomenclature (50).

[b] Information on ancestral allele is not available from chimpanzee but from macaque.

[c] Statistical analysis and calculation were only conducted in women.

(31). Mice with targeted mutations of HNF4A have a dramatic decrease in LDL and HDL cholesterol levels, and the HDL particles in these mice are small and lipid-poor (31). These results altogether show that HNF4A is important in HDL-C metabolism.

The APOA5 gene is located ~27 kb upstream of the well characterized APOA1/C3/A4 gene cluster. It is suggested that HDL is a reservoir for apoA-V based on the observation that there was a greater than 20-fold higher level of apoA-V in HDL than in VLDL of APOA5 transgenic mice

(32). There are numerous reports on the association of genetic variants of APOA5 with low HDL-C levels, such as rs662799, rs651821, and rs2075291 in Japanese populations (33) and Chinese populations (34). The SNP of rs3135506 is significantly associated with blood HDL-C levels in northern American (35) and Austrian Caucasian subjects (36). In our Dutch population, rs3135506 was associated with plasma HDL-C levels before adjustment for multiple testing. This SNP, substituting Trp for Ser at residual 19 within the predicted signal peptide, is the only common variant with known influence on APOA-V expression (37). Furthermore, another SNP (rs28927680) in linkage disequilibrium with rs3135506 ($r^2 = 0.98$) is recently reported to be associated with HDL-C concentrations in a genome-wide association analysis (27). Therefore, we think rs3135506 is an important determinant of HDL-C levels.

A SNP (rs2229741) in nuclear receptor interacting protein 1 (NRIP1) was found to be associated with plasma HDL-C levels in our study. Another SNP (rs2229742), ~1kb downstream of rs2229741 was also found to be associated with plasma HDL-C levels, although the significance disappeared after adjustment for multiple testing (Table 2). These two SNPs are correlated with each other ($r^2 = 0.44$). Two other SNPS (rs1297214 and rs2142450) in this gene are shown to be associated with age-sex/multivariable adjusted NMR intermediate HDL concentration in the Framingham Offspring study (Exam 4) using a generalized estimating equation regression approach ($P = 0.02/0.04$ for rs1297214; $P = 0.02/0.04$ for rs2142450) (38). NRIP1 is reported to be involved in the estrogen regulation of apoA-I enhancer activity. At low ratios of NRIP1 to estrogen receptor α, estradiol repressed apoA-I enhancer activity, whereas at high ratios this repression was reversed (39). Recent studies have identified a hepatocyte specific role for NRIP1 as a cofactor for LXR in different ways, namely serving as a coactivator in lipogenesis and as a corepressor in gluconeogenesis (40). Since NRIP1 is a widely expressed cofactor for nuclear receptors involved in the regulation of metabolic gene expression (41), its exact role in HDL-C metabolism warrants further research.

Rs17585739 in sterol C4-methyloxidase-like (SC4MOL) that catalyzes sterol-4α-methyl oxidation process in cholesterol biosynthesis (42), was associated with plasma HDL-C levels in our population. Another SNP (rs1550270) in this gene was also associated with multivariable adjusted NMR large HDL concentration ($P = 0.03$) and HDL size ($P = 0.04$) in the Offspring Exam 4 of the Framingham SNP Health Association Resource (38). Interestingly, also along the cholesterol biosynthetic pathway, MVK (mevalonate kinase), encoding mevalonate kinase to convert mevalonate into 5-phosphomevalonate (2), is reported to possibly affect HDL-C concentrations in a recent genome-wide association study (25). However, the SNP (rs7957619) in MVK we studied was not associated with plasma HDL-C concentrations ($P = 0.45$). Another SNP in a gene regulating this pathway, rs3213451 in MBTPS2, a metalloprotease required for intramembrane proteolysis of sterol regulatory element-binding proteins (2), was associated with plasma HDL-C concentrations in women in our study. Although this finding needs replication in other studies, it is tempting to speculate that de novo biosynthesis of cholesterol is an important contributor to plasma HDL-C. However, this remains to be confirmed.

Also for other SNPs, our results are more speculative. The results are inconsistent with those found in other studies, as in those studies no associations with HDL-C levels have been reported but with other (lipid) parameters. Also no association has been found between these genes and HDL-C in recent genome-wide association studies. Therefore, these findings need replication. The genetic variant of PTPN11 (rs11066322) is reported to be associated with serum apoB levels and LDL cholesterol levels in a study conduced in Caucasian female twins (43), but in our study, this genetic variant was associated with plasma HDL-C levels. Furthermore, the genetic variant of ADD1 (rs4961) has been associated with susceptibility to hypertension, but not with plasma HDL-C levels (44). However, we did find a significant association between rs4961 and plasma HDL-C levels in our study. SCAN-domain-containing protein 1 (SCAND1) is the nearest gene to the locus of rs6060717 that was found to be associated with plasma HDL-C levels in our study. SCAND1 is a widely expressed nuclear protein that may function as a key regulator of zinc finger transcription factor function (45). The capability of SCAND1 to interact with ZNF202 and PPARγ2 as two crucial transcription factors involved in lipid metabolism suggests that SCAND1 could function as important coregulator in lipid metabolism (46). In view of the above, a role of these genes in lipid metabolism is possible, but the association with HDL-C needs to be studied further.

There are some limitations to acknowledge in our study, which could have lead to either type I (false-positive) or type II (false-negative) errors. We adopted the FDR method to adjust for multiple testing, as we took a candidate gene approach (and not random markers on the genome), and Bonferroni correction would have been too stringent. We think it is justified to take a more liberal threshold of FDR_q value < 0.2. Although we took this approach to avoid false-positive results, we cannot exclude the possibility that the novel SNPs are spuriously associated with plasma HDL-C levels; therefore, the results for novel SNPs need further replication in other populations. A number of factors could have resulted in a type II error, leading to the inability to detect a true underlying association. First, only a limited number of SNPs within a candidate gene has been studied. Failure to find an association with these SNPs does not exclude the possibility that other SNPs in the gene are related to plasma HDL-C levels. Second, less-common variants with low minor allele frequencies (between 0.001 and 0.01) have been proven to affect plasma HDL-C levels (47). Although our sample size is large enough to allow us to detect effects as low as 0.4% variation in plasma HDL-C levels with 80% power ($P = 0.005$, FDR_q = 0.2) for such less-common genetic variants, we did not take these less-common variants into consideration in our study. Recently, some authors argued that with the findings from the previous linkage studies, candidate gene studies, and also current genome-wide association studies, we still can-

not explain the high heritability of complex phenotypes. In fact, all the significant 13 SNPs (except rs3213451 in MBTPS2) identified in our study only explain 4% of the inter-individual variation in average HDL-C concentrations over the three measurements in an additive model in our study. It is suggested that deep resequencing of some candidate genes to find these culprit (rare) variants may be a solution (48). Finally, our blood samples were taken from the subjects in a nonfasting condition. However, there was no clear association between the time to last meal and plasma HDL-C levels in the second and third surveys (data not shown). Also others found that fasting does not influence HDL-C levels to a large extent (49). Therefore, we do not think the nonfasting condition has influenced our results.

In conclusion, in addition to replication of previous results for candidate genes (SNPs among CETP, LPL, LIPC, HNF4A, and APOA5), we found interesting new candidate SNPs for plasma HDL-C levels (rs2229741 in NRIP1, rs3213451 in MBTPS2, rs17585739 in SC4MOL, rs11066322 in PTPN11, rs4961 in ADD1, and rs6060717 near SCAND1). The validity and relevance of these novel associations with plasma HDL-C levels require further validation in other study populations.

The authors thank the epidemiologists and fieldworkers of the Municipal Health Service in Doetinchem for their contribution to the data collection for this study. Logistic management was provided by J. Steenbrink and P. Vissink, and administrative support by E.P. van der Wolf. Data management was provided by A. Blokstra, A.W.D. van Kessel, and P.E. Steinberger. The authors would also like to thank B. Hoebee for her role in the conception of the study, C. Strien for DNA preparation, K. Siezen for her assistance in SNP selection, H.M. Hodemakers for genotyping assistance, and S.W. van den Berg and A.G. Heidema for their statistical advice.

REFERENCES

1. Gordon, D. J., and B. M. Rifkind. 1989. High-density lipoprotein–the clinical implications of recent studies. *N. Engl. J. Med.* **321:** 1311–1316.
2. Lodish, H., A. Berk, P. Matsudaira, C. A. Kaiser, M. Krieger, M. P. Scott, S. Zipursky, and J. Darnell. 2004. Molecular cell biology. 5th ed. W. H. Freeman and Company, New York. 750–767.
3. Lewis, G. F., and D. J. Rader. 2005. New insights into the regulation of HDL metabolism and reverse cholesterol transport. *Circ. Res.* **96:** 1221–1232.
4. Barter, P. J., S. Nicholls, K. A. Rye, G. M. Anantharamaiah, M. Navab, and A. M. Fogelman. 2004. Antiinflammatory properties of HDL. *Circ. Res.* **95:** 764–772.
5. Mineo, C., H. Deguchi, J. H. Griffin, and P. W. Shaul. 2006. Endothelial and antithrombotic actions of HDL. *Circ. Res.* **98:** 1352–1364.
6. Dattilo, A. M., and P. M. Kris-Etherton. 1992. Effects of weight reduction on blood lipids and lipoproteins: a meta-analysis. *Am. J. Clin. Nutr.* **56:** 320–328.
7. Moffatt, R. J. 1988. Effects of cessation of smoking on serum lipids and high density lipoprotein-cholesterol. *Atherosclerosis.* **74:** 85–89.
8. Kokkinos, P. F., J. C. Holland, P. Narayan, J. A. Colleran, C. O. Dotson, and V. Papademetriou. 1995. Miles run per week and high-density lipoprotein cholesterol levels in healthy, middle-aged men. A dose-response relationship. *Arch. Intern. Med.* **155:** 415–420.
9. McConnell, M. V., I. Vavouranakis, L. L. Wu, D. E. Vaughan, and P. M. Ridker. 1997. Effects of a single, daily alcoholic beverage on lipid and hemostatic markers of cardiovascular risk. *Am. J. Cardiol.* **80:** 1226–1228.
10. Mooradian, A. D., M. J. Haas, and N. C. Wong. 2006. The effect of select nutrients on serum high-density lipoprotein cholesterol and apolipoprotein A-I levels. *Endocr. Rev.* **27:** 2–16.
11. Peacock, J. M., D. K. Arnett, L. D. Atwood, R. H. Myers, H. Coon, S. S. Rich, M. A. Province, and G. Heiss. 2001. Genome scan for quantitative trait loci linked to high-density lipoprotein cholesterol: The NHLBI Family Heart Study. *Arterioscler. Thromb. Vasc. Biol.* **21:** 1823–1828.
12. Brooks-Wilson, A., M. Marcil, S. M. Clee, L. H. Zhang, K. Roomp, M. van Dam, L. Yu, C. Brewer, J. A. Collins, H. O. Molhuizen, et al. 1999. Mutations in ABC1 in Tangier disease and familial high-density lipoprotein deficiency. *Nat. Genet.* **22:** 336–345.
13. Hovingh, G. K., E. de Groot, W. van der Steeg, S. M. Boekholdt, B. A. Hutten, J. A. Kuivenhoven, and J. J. Kastelein. 2005. Inherited disorders of HDL metabolism and atherosclerosis. *Curr. Opin. Lipidol.* **16:** 139–145.
14. Brown, M. L., A. Inazu, C. B. Hesler, L. B. Agellon, C. Mann, M. E. Whitlock, Y. L. Marcel, R. W. Milne, J. Koizumi, H. Mabuchi, et al. 1989. Molecular basis of lipid transfer protein deficiency in a family with increased high-density lipoproteins. *Nature.* **342:** 448–451.
15. Knoblauch, H., A. Bauerfeind, M. R. Toliat, C. Becker, T. Luganskaja, U. P. Gunther, K. Rohde, H. Schuster, C. Junghans, F. C. Luft, et al. 2004. Haplotypes and SNPs in 13 lipid-relevant genes explain most of the genetic variance in high-density lipoprotein and low-density lipoprotein cholesterol. *Hum. Mol. Genet.* **13:** 993–1004.
16. Isaacs, A., F. A. Sayed-Tabatabaei, O. T. Njajou, J. C. Witteman, and C. M. van Duijn. 2004. The -514 C->T hepatic lipase promoter region polymorphism and plasma lipids: a meta-analysis. *J. Clin. Endocrinol. Metab.* **89:** 3858–3863.
17. Brown, C. M., T. J. Rea, S. C. Hamon, J. E. Hixson, E. Boerwinkle, A. G. Clark, and C. F. Sing. 2006. The contribution of individual and pairwise combinations of SNPs in the APOA1 and APOC3 genes to interindividual HDL-C variability. *J. Mol. Med.* **84:** 561–572.
18. Desvergne, B., L. Michalik, and W. Wahli. 2006. Transcriptional regulation of metabolism. *Physiol. Rev.* **86:** 465–514.
19. Verschuren, W., A. Blokstra, H. Picavet, and H. Smit. 2008. Cohort Profile: The Doetinchem Cohort Study. *Int. J. Epidemiol.* 10.1093/ije/dym292.
20. Ocke, M. C., H. B. Bueno-de-Mesquita, M. A. Pols, H. A. Smit, W. A. van Staveren, and D. Kromhout. 1997. The Dutch EPIC food frequency questionnaire. II. Relative validity and reproducibility for nutrients. *Int. J. Epidemiol.* **26 (Suppl 1):** S49–S58.
21. Becker, K. G., K. C. Barnes, T. J. Bright, and S. A. Wang. 2004. The genetic association database. *Nat. Genet.* **36:** 431–432.
22. Riva, A., and I. S. Kohane. 2004. A SNP-centric database for the investigation of the human genome. *BMC Bioinformatics.* **5:** 33.
23. Xu, H., S. G. Gregory, E. R. Hauser, J. E. Stenger, M. A. Pericak-Vance, J. M. Vance, S. Zuchner, and M. A. Hauser. 2005. SNPselector: a web tool for selecting SNPs for genetic association studies. *Bioinformatics.* **21:** 4181–4186.
24. Smith, N. L., L. A. Hindorff, S. R. Heckbert, R. N. Lemaitre, K. D. Marciante, K. Rice, T. Lumley, J. C. Bis, K. L. Wiggins, F. R. Rosendaal, et al. 2007. Association of genetic variations with nonfatal venous thrombosis in postmenopausal women. *JAMA.* **297:** 489–498.
25. Willer, C. J., S. Sanna, A. U. Jackson, A. Scuteri, L. L. Bonnycastle, R. Clarke, S. C. Heath, N. J. Timpson, S. S. Najjar, H. M. Stringham, et al. 2008. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat. Genet.* **40:** 161–169.
26. Kooner, J. S., J. C. Chambers, C. A. Aguilar-Salinas, D. A. Hinds, C. L. Hyde, G. R. Warnes, F. J. Gomez Perez, K. A. Frazer, P. Elliott, J. Scott, et al. 2008. Genome-wide scan identifies variation in MLXIPL associated with plasma triglycerides. *Nat. Genet.* **40:** 149–151.
27. Kathiresan, S., O. Melander, C. Guiducci, A. Surti, N. P. Burtt, M. J. Rieder, G. M. Cooper, C. Roos, B. F. Voight, A. S. Havulinna, et al. 2008. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat. Genet.* **40:** 189–197.
28. Frazer, K. A., D. G. Ballinger, D. R. Cox, D. A. Hinds, L. L. Stuve, R. A. Gibbs, J. W. Belmont, A. Boudreau, P. Hardenbol, S. M. Leal, et al. 2007. A second generation human haplotype map of over 3.1 million SNPs. *Nature.* **449:** 851–861.
29. Zhu, Q., K. Yamagata, A. Miura, N. Shihara, Y. Horikawa, J. Takeda, J. Miyagawa, and Y. Matsuzawa. 2003. T130I mutation in HNF-4alpha gene is a loss-of-function mutation in hepatocytes and is as-

sociated with late-onset Type 2 diabetes mellitus in Japanese subjects. *Diabetologia.* **46:** 567–573.
30. Pearson, E. R., S. Pruhova, C. J. Tack, A. Johansen, H. A. Castleden, P. J. Lumb, A. S. Wierzbicki, P. M. Clark, J. Lebl, O. Pedersen, et al. 2005. Molecular genetics and phenotypic characteristics of MODY caused by hepatocyte nuclear factor 4alpha mutations in a large European collection. *Diabetologia.* **48:** 878–885.
31. Hayhurst, G. P., Y. H. Lee, G. Lambert, J. M. Ward, and F. J. Gonzalez. 2001. Hepatocyte nuclear factor 4alpha (nuclear receptor 2A1) is essential for maintenance of hepatic gene expression and lipid homeostasis. *Mol. Cell. Biol.* **21:** 1393–1403.
32. Nelbach, L., X. Shu, R. J. Konrad, R. O. Ryan, and T. M. Forte. 2008. Effect of apolipoprotein A-V on plasma triglyceride, lipoprotein size and composition in genetically engineered mice. *J. Lipid Res.* **49:** 572–580.
33. Yamada, Y., H. Matsuo, S. Warita, S. Watanabe, K. Kato, M. Oguri, K. Yokoi, N. Metoki, H. Yoshida, K. Satoh, et al. 2007. Prediction of genetic risk for dyslipidemia. *Genomics.* **90:** 551–558.
34. Chien, K. L., M. F. Chen, H. C. Hsu, T. C. Su, W. T. Chang, C. M. Lee, and Y. T. Lee. 2008. Genetic association study of APOA1/C3/A4/A5 gene cluster and haplotypes on triglyceride and HDL cholesterol in a community-based population. *Clin. Chim. Acta.* **388:** 78–83.
35. Lai, C. Q., S. Demissie, L. A. Cupples, Y. Zhu, X. Adiconis, L. D. Parnell, D. Corella, and J. M. Ordovas. 2004. Influence of the APOA5 locus on plasma triglyceride, lipoprotein subclasses, and CVD risk in the Framingham Heart Study. *J. Lipid Res.* **45:** 2096–2105.
36. Grallert, H., E. M. Sedlmeier, C. Huth, M. Kolz, I. M. Heid, C. Meisinger, C. Herder, K. Strassburger, A. Gehringer, M. Haak, et al. 2007. APOA5 variants and metabolic syndrome in Caucasians. *J. Lipid Res.* **48:** 2614–2621.
37. Talmud, P. J., J. Palmen, W. Putt, L. Lins, and S. E. Humphries. 2005. Determination of the functionality of common APOA5 polymorphisms. *J. Biol. Chem.* **280:** 28215–28220.
38. Resource, Framingham SNP Health Association. Accessed February 19, 2008, at http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/analysis.cgi?study_id=phs000007.v2.p1&phv=&phd=&pha=547&phsf=0&phvf=&phdf=&phaf=747.
39. Harnish, D. C., M. J. Evans, M. S. Scicchitano, R. A. Bhat, and S. K. Karathanasis. 1998. Estrogen regulation of the apolipoprotein AI gene promoter through transcription cofactor sharing. *J. Biol. Chem.* **273:** 9270–9278.
40. Herzog, B., M. Hallberg, A. Seth, A. Woods, R. White, and M. G. Parker. 2007. The nuclear receptor cofactor, receptor-interacting protein 140, is required for the regulation of hepatic lipid and glucose metabolism by liver X receptor. *Mol. Endocrinol.* **21:** 2687–2697.
41. Parker, M. G., M. Christian, and R. White. 2006. The nuclear receptor co-repressor RIP140 controls the expression of metabolic gene networks. *Biochem. Soc. Trans.* **34:** 1103–1106.
42. Rodriguez, C., B. Raposo, J. Martinez-Gonzalez, V. Llorente-Cortes, G. Vilahur, and L. Badimon. 2003. Modulation of ERG25 expression by LDL in vascular cells. *Cardiovasc. Res.* **58:** 178–185.
43. Jamshidi, Y., S. B. Gooljar, H. Snieder, X. Wang, D. Ge, R. Swaminathan, T. D. Spector, and S. D. O'Dell. 2007. SHP-2 and PI3-kinase genes PTPN11 and PIK3R1 may influence serum apoB and LDL cholesterol levels in normal women. *Atherosclerosis.* **194:** e26–e33.
44. Yazdanpanah, M., F. A. Sayed-Tabatabaei, A. Hofman, Y. S. Aulchenko, B. A. Oostra, B. H. Stricker, H. A. Pols, S. W. Lamberts, J. C. Witteman, J. A. Janssen, et al. 2006. The alpha-adducin gene is associated with macrovascular complications and mortality in patients with type 2 diabetes. *Diabetes.* **55:** 2922–2927.
45. Sander, T. L., and J. F. Morris. 2002. Characterization of the SCAN box encoding RAZ1 gene: analysis of cDNA transcripts, expression, and cellular localization. *Gene.* **296:** 53–64.
46. Schmitz, G., S. Heimerl, and T. Langmann. 2004. Zinc finger protein ZNF202 structure and function in transcriptional control of HDL metabolism. *Curr. Opin. Lipidol.* **15:** 199–208.
47. Cohen, J. C., R. S. Kiss, A. Pertsemlidis, Y. L. Marcel, R. McPherson, and H. H. Hobbs. 2004. Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science.* **305:** 869–872.
48. Petretto, E., E. T. Liu, and T. J. Aitman. 2007. A gene harvest revealing the archeology and complexity of human disease. *Nat. Genet.* **39:** 1299–1301.
49. Craig, S. R., R. V. Amin, D. W. Russell, and N. F. Paradise. 2000. Blood cholesterol screening influence of fasting state on cholesterol results and management decisions. *J. Gen. Intern. Med.* **15:** 395–399.
50. den Dunnen, J. T., and S. E. Antonarakis. 2001. Nomenclature for the description of human sequence variations. *Hum. Genet.* **109:** 121–124.