Genome-wide linkage analyses and candidate gene fine mapping for HDL_3 cholesterol: the Framingham Study

Qiong Yang,[1,*] Chao-Qiang Lai,[†] Laurence Parnell,[†] L. Adrienne Cupples,[*] Xian Adiconis,[†] Yueping Zhu,[†] Peter W. F. Wilson,[§] David E. Housman,[**] Amanda M. Shearman,[**] Ralph B. D'Agostino,[††] and Jose M. Ordovas[†]

Department of Biostatistics,[*] Department of Endocrinology, Nutrition, and Diabetes,[§] and Department of Mathematics and Statistics,[††] Boston University, Boston, MA 02118; Nutrition and Genomics Laboratory,[†] Jean Mayer-United States Department of Agriculture Human Nutrition Research Center on Aging, Tufts University, Boston, MA 02111; and Center for Cancer Research,[**] Massachusetts Institute of Technology, Cambridge, MA 02139

Abstract **High density lipoprotein cholesterol (HDL-C) is inversely associated with coronary heart disease and has a genetic component; however, linkage to HDL-C is not conclusive. Subfractions of HDL, such as HDL_3-C, may be better phenotypes for linkage studies. Using HDL_3-C levels measured on 907 Framingham Heart Study subjects from 330 families around 1987, we conducted a genome-wide variance components linkage analysis with 401 microsatellite markers spaced ~10 centimorgan (cM) apart. Nine candidate genes were identified and annotated using a bioinformatics approach in the region of the highest linkage peak. Twenty-eight single nucleotide polymorphisms (SNPs) were selected from these candidate genes, and linkage and family-based association fine mapping were conducted using these SNPs. The highest multipoint log-of-the-odds (LOD) score from the initial linkage analysis was 3.7 at 133 cM on chromosome 6. Linkage analyses with additional SNPs yielded the highest LOD score of 4.0 at 129 cM on chromosome 6. Family-based association analysis revealed that SNP rs2257104 in PLAGL1 at ~143 cM was associated with multivariable adjusted HDL_3 ($P = 0.03$). Further study of the linkage region and exploration of other variants in PLAGL1 are warranted to define the potential functional variants of HDL-C metabolism.**—Yang, Q., C-Q. Lai, L. Parnell, L. A. Cupples, X. Adiconis, Y. Zhu, P. W. F. Wilson, D. E. Housman, A. M. Shearman, R. B. D'Agostino, and J. M. Ordovas. **Genome-wide linkage analyses and candidate gene fine mapping for HDL_3 cholesterol: the Framingham Study. *J. Lipid Res.* 2005.** 46: **1416–1425.**

Supplementary key words high density lipoprotein • high density lipoprotein 3 • microsatellite marker • single nucleotide polymorphism • PLAGL1

The inverse association of high density lipoprotein cholesterol (HDL-C) concentrations and risk of coronary heart disease was initially observed in the 1960s and 1970s and is now well established (1). Several studies have provided significant evidence supporting strong familial aggregation of HDL-C levels, but inconsistent results have been reported on whether this is attributable to a gene of major effect (2–6).

HDL is a heterogeneous mixture of particles of different sizes, densities, and compositions (7, 8). Differential chemical precipitation separates HDL particles into two major subfractions: HDL_2 and HDL_3. As thoroughly reviewed by Wilson (9), a number of studies have found that lower concentrations of HDL_2-C and HDL_3-C are significantly associated with increased coronary heart disease risk, suggesting an important role of these two subfractions in coronary heart disease. However, there have been inconsistent findings across populations. To obtain a better understanding of the genetic influences on HDL levels, some studies have examined the more metabolically homogenous subfractions. A segregation study using families of patients who underwent coronary arteriography found that a major gene explains 34% of the total variation in HDL_3-C and 9% in HDL-C, but the variation in HDL_2-C is likely mostly the result of environmental factors (10). Using gradient gel electrophoresis to further differentiate three subclasses (HDL_{3a}, HDL_{3b}, and HDL_{3c}) within HDL_3 and two subclasses (HDL_{2a} and HDL_{2b}) within HDL_2, a genome-wide linkage study (11) was performed on the five subclasses, and significant linkage [log-of-the-odds (LOD) > 3] to HDL_{2a} and suggestive linkage (LOD > 2) to HDL_{3a} and HDL_{2b} was found. These studies suggest that the subfractions may provide better phenotypes than HDL-C levels for genetic studies.

Other investigators (12–18) have examined the association between HDL and HDL subfractions and candidate

Manuscript received 4 October 2004 and in revised form 2 March 2005.

Published, JLR Papers in Press, April 1, 2005.
DOI 10.1194/jlr.M400382-JLR200

[1] To whom correspondence should be addressed.
e-mail: qyang@bu.edu

genes such as hepatic lipase, LPL, cholesteryl ester transfer protein (CETP), LCAT, and apolipoproteins A-I, A-II, A-IV, B, C-III, and E. Similar to the outcomes of the segregation studies, there is considerable heterogeneity regarding the significance of these associations in different populations.

To identify chromosome regions likely to contain quantitative trait loci (QTLs) affecting HDL_3-C concentration, we performed a genome-wide linkage study using 401 microsatellite markers and 330 extended families from the Framingham Heart Study. Because this analysis provided significant evidence of linkage to a region on chromosome 6, we identified a number of candidate genes within this promising region of the highest LOD score. A number of single nucleotide polymorphisms (SNPs) in these candidate genes were selected and genotyped to further map possible variants that may affect HDL_3-C concentration.

METHODS

Subjects

The study subjects are members of the 330 largest extended families in the Framingham Heart Study. The selection criteria and study design of the Framingham Heart Study have been described in detail previously (19, 20). In brief, the study began in 1948 with the enrollment of 5,209 men and women, referred to as the original cohort, from Framingham, Massachusetts, who have undergone biennial examinations. Starting in 1971, 5,124 individuals, adult children of the original cohort, and their spouses, were recruited and are referred to as the offspring cohort. The members of the offspring cohort were examined every 4 years (except for an 8 year gap between the first and second examinations). Subjects in the Framingham Heart Study were ascertained without regard to any trait values. In the mid to late 1990s, 1,702 subjects from the 330 largest extended pedigrees were genotyped for a set of 401 microsatellite markers spaced ~10 centimorgan (cM) apart.

The analysis reported here was conducted on a subset of 907 offspring subjects who had HDL_3-C levels measured at examination cycle 4 around 1987 and genotypic data of the CETP polymorphism. The sample included 729 sibling pairs, 37 half-sibling pairs, 41 avuncular pairs, and 489 first cousin pairs. CETP was genotyped on study participants who attended examinations 4 or 5 between 1987 and 1995 and for whom we had DNA. Participants may have missed both exams as a result of death, serious illness, moving out of Massachusetts, or other unknown reasons. Of those 1,089 subjects who attended examination 4 or 5, had DNA, and were members of the families that were included in the genome scan, 282 (23%) subjects were excluded because their CETP genotypes were not available.

All subjects provided informed consent before each clinic visit and were examined under the standard protocol for the Framingham Offspring Study approved by the Institutional Review Board at Boston Medical Center (Boston, MA).

Genotyping of microsatellite markers

DNA specimens were obtained from blood samples routinely collected during examinations of original and offspring subjects between 1987 and 1991 and between 1995 and 2000. DNA was extracted from the buffy coat of whole blood specimens using a Qiagen Blood and Cell Culture DNA Maxi Kit. A genome-wide scan with 401 microsatellite markers over the 22 chromosomes, at an average of 1 marker every 10 cM, was completed by Mammalian Genotyping Service (Marshfield, WI) in 2000. The marker screen set 9 and genotyping protocols are available at the website of the Center for Medical Genetics, Marshfield Medical Research Foundation (http://research.marshfieldclinic.org/genetics/).

Measurement of lipids

Fasting venous blood samples were collected, and plasma was separated from blood cells by centrifugation and immediately used for the measurement of lipids. HDL_3-C concentrations were measured as described by Gidez et al. (21).

Definition of traditional risk factors

Body mass index (BMI; in kg/m^2) was calculated using the measurements of weight and height. Alcohol consumption was reported as the usual number of drinks (of comparable ethanol content) per week. Cigarette smoking (yes/no) was defined by whether a person had a history of regularly smoking cigarettes. Physical activity index was derived by taking a weighted average of the number of hours spent in five different energy expenditure categories (sleep, 1.0; sedentary, 1.1; slight, 1.5; moderate, 2.4; and heavy, 5.0). Medication treatment variables (yes/no; including anticholesterol therapy, β-blockers, and oral contraceptives) were defined by whether a person took such treatment in the past year. Menopause status (yes/no) was defined by whether a woman's menstrual periods had stopped for 1 year. Estrogen therapy (yes/no) was defined by whether a woman was taking estrogen therapy after menopause.

Heritability and linkage analyses

Linkage information content at each locus was calculated in GENEHUNTER. Linkage information content at a locus for a pedigree is defined as how closely the exact allele identity-by-descendent sharing at this locus can be determined for every relative pair (22). Linkage information content at a locus for a sample of pedigrees is a sum of information content at this locus over all of the pedigrees in the sample. The higher the linkage information content, the better the power to detect linkage if the locus is close to the QTL.

To reduce the variability caused by known risk factors, we calculated the standardized residuals from sex-specific multiple linear regression models adjusted for the traditional risk factors. We further calculated normalized deviate scores transformed from ranked standardized residuals for heritability and linkage analyses to avoid potentially inflated false-positive rates attributable to nonnormality. The linear regressions were conducted separately for men and women to account for different associations and to further adjust for oral contraceptive medicine, estrogen therapy, and menopause in women. The heritability and linkage analyses were conducted using MERLIN (23), which is based on a variance components methodology. A measure of the evidence for linkage, the LOD score, is log base 10 of the likelihood ratio of the model with a QTL effect to that without such effect associated with the marker locus under evaluation.

Bioinformatic identification of candidate genes in the linkage region

We conducted a bioinformatic search of candidate genes within or slightly beyond the 2-LOD support interval of the linkage peak (125–150 cM or 6q22.33–6q24.3) on chromosome 6 (**Fig. 1**), the only chromosome on which significant linkage was found. MapViewer at the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/) was used to obtain predicted genes. Similarity searches, using these predicted genes and genomic sequence as input, were performed with the BLAST algorithm (24) (http://www.ncbi.nlm.nih.gov/BLAST/).

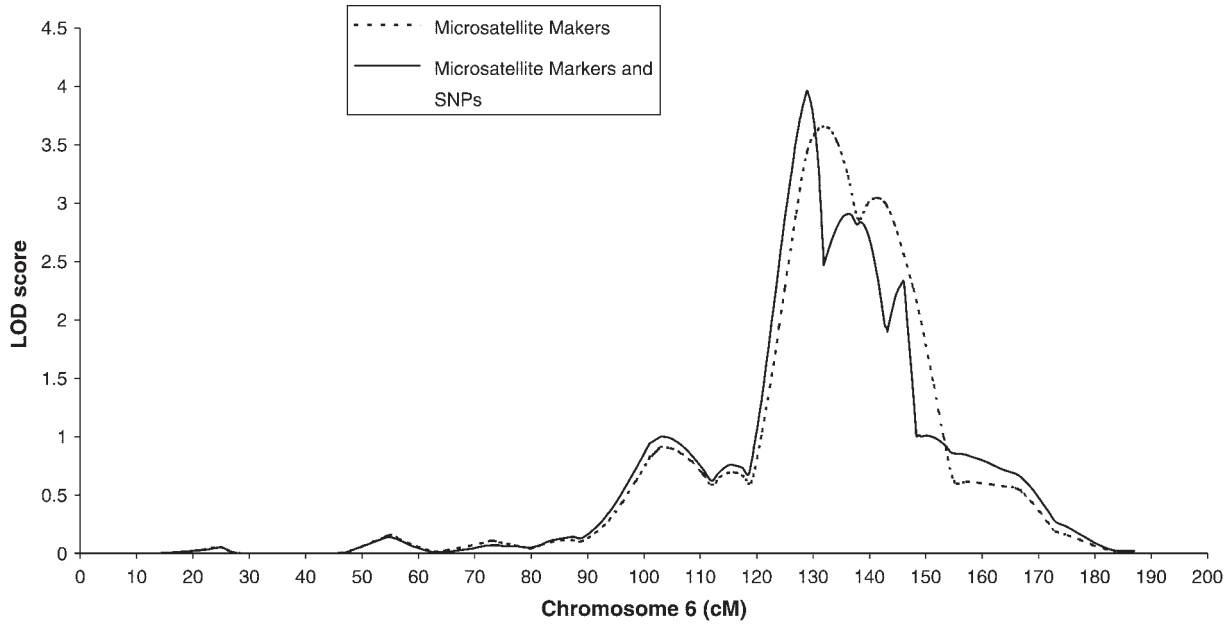

Fig. 1. Multipoint log-of-the-odds (LOD) scores on chromosome 6 for high density lipoprotein cholesterol (HDL_3-C) with existing microsatellite markers and additional single nucleotide polymorphisms (SNPs). cM, centimorgan.

Protein sequences of encoded genes were searched with BLASTP against a "nonredundant" protein set and with TBLASTN against the mouse genome as described above. To provide further functional characterization, protein sequences were also analyzed for a range of functional domains [Pfam release 8.0 (25), http://pfam.wustl.edu; InterPRO release 5.3 (26), http://dip.doe-mbi.ucla.edu/], including putative transmembrane domains [TM-HMM version 2.0 (27), http://www.cbs.dtu.dk/services/TMHMM/], for putative subcellular localization [TargetP version 1.01 (28), http://www.cbs.dtu.dk/services/TargetP/], for potential N-linked glycosylation sites (NetNglyc version 1.0; R. Gupta, E. Jung, and S. Brunak, unpublished data, 2002; www.cbs.dtu.dk/services/NetNGlyc/), and for the presence and location of putative signal peptide cleavage sites [SignalP version 2.0.b2 (29), www.cbs.dtu.dk/services/SignalP/], all as a means to further characterize and understand the functions of these proteins, especially as they relate to lipid biochemistry.

SNP selection and genotyping

Twenty-eight SNPs in the candidate genes were selected from the public SNP database (http://www.ncbi.nlm.nih.gov/SNP) based on the following rationale. First, even distribution across a gene was sought with three, four, or five SNPs. Second, SNPs were selected based on their potential to be functional: altering protein sequence (particularly in functional domains), affecting mRNA splicing, or altering promoter activity. Third, attention was given to SNPs in regions of either protein-coding or noncoding sequence that is conserved in other species, notably the mouse. Fourth, SNPs must have a minor allele frequency of at least 5%. SNP genotyping was conducted using the ABI Prism SnapShot multiplex system (Applied Biosystems, Foster City, CA) and was described previously (30). The primers and probes used for genotyping are displayed in **Table 1**.

Determination of the genetic map for the SNPs and microsatellite markers

Genetic locations in centimorgan of the SNPs were extrapolated from the existing Marshfield genetic map for microsatellite markers (31) and then were merged into the Marshfield genetic map to form a new map for linkage analysis. The extrapolation was as follows. First, the ratio of centimorgan distance to physical distance was calculated between two adjacentmicrosatellite markers; second, for each SNP between two adjacent markers, the genetic distance between this SNP and one of the markers was the physical distance times the ratio of centimorgan distance to physical distance between the two markers.

Linkage disequilibrium structure at the candidate genes

Each SNP was evaluated for deviation from Hardy-Weinberg equilibrium (HWE) by randomly selecting one subject from each family and comparing the genotype frequencies of those unrelated subjects with the expected frequencies under HWE in a Chi-square statistic of 2 degrees of freedom.

For each pair of SNPs within the same candidate gene or located less than 500 kb apart, a measure of the strength of linkage disequilibrium (LD), D′ (32), was calculated using Haploview (http://www.broad.mit.edu/personal/jcbarret/haploview/).

Family-based association analyses with SNPs in the candidate genes

A score statistic that captures the covariance between marker genotype and trait was calculated using FBAT (33, 34) to test for association. The score statistic divided by its variance was distributed as standard normal (Z). This statistic conditions on parental genotypes to control for bias caused by potential population admixture. Population admixture can be characterized as differential marker allele and trait distribution among unobserved subpopulations. When parental genotypes were incomplete, the distribution of the score statistic was conditioned on the genotypes of available parents and the whole sibship. Because the association was tested with markers in the linkage region, an empirical variance of the score statistic that took into account the dependence among relatives was used in computing the conditional distribution (35).

In haplotype analyses using FBAT (36), multiple marker haplotypes instead of single marker alleles were used in the score statistic. For markers with ambiguous phase, the test statistic was a weighted average of the score statistics over all compatible phase configurations. Two types of hypothesis tests were performed:

TABLE 1. Probes and primers used for SnapShot genotyping for 28 SNPs

SNP name	Probe	Forward Primer	Reverse Primer
rs11963811	TGCAAAATCAGGTCATTCTC	AACCTTATGTCCTCCCTGTCTCCAG	CAACCCCACTTTCAGAATGGTACAG
rs1567	GAGGTGTGGGCATTTGA	AAGATGGGGGCAAGAGAGAATAGTG	TTCTTTTTCCAGCCAGCTTTAGGAG
rs15960	CAGGGTGGTGGTTCTGT	TCATTGTCTCCGGGACAGTTGTAAT	ACCAGCATGAAGACATACCGAGCTA
rs1883617	GGACCATTTGGTGCATAAAT	TTTTCCCAACTCTGTCTTCATGTCA	GGGAAATTGGCAGTTTTGGTTTTT
rs1884087	TTTCCTTATTTGGTTATTTGTCT	GAGGGGATGTCGAAAGTCCTAGTGT	AACTGGTCAGCAACAGCTGGATTAG
rs2064495	AACAAAAATCACCCACACTTA	CACTATAGCTGGGGCATGTCCTG	AGCTATCAGCTTTGTGCCACCAC
rs2065168	TTCATTCTCTTGATGTCATAGC	GGACCTCTGTTCAGCATTTCACAGT	AAGAATGGGGTACCTGTGTTGACCT
rs2076684	CAGGTGCGATTCAGGA	TTCTACCCAGACATGGACCTCTCAG	AGCACCATTTAAAGGAACGTGTTGA
rs2095376	AAAAGTACTTAGATGGGCAGTAG	TCCACCTTCATGCCCTAATCATC	GCCCTTCTAAGCACTTCACATGC
rs2257104	AATTTGGAATTTCAAAATAGCT	GCTTCAGGGACAAGCCGTAGTCT	ATCTGCCCTGAGGGATATGAAGC
rs2272996	ATTTGCCACAACATAGATAGAG	ATCACTGGTATCGCATGGCTTCTT	GCCAGACCCCAGTACAAGAAAGACT
rs2294757	TGCTCATAAACAGCTGCA	GATTCATTAATGCCAAAGCCTCCTC	CGTGGCAATTTTGCTTTTCTATGTC
rs2294759	TGTATTTCTCTTAAGCTTGGTG	CCAGCCCAATTTTCAAAGAAAGG	GAGTGGAATCTACGCCCCAGAAG
rs2745427	TCACTGTCAAAAAGATTAGCC	CATGTGGGATATGTGAATTTGCTTG	CAGATTACAGGGGTTGCCTCTAGGT
rs2876549	GTTTGCAAAGAGAAATCACA	AGATCTGGAAAGTAGGAAGGCAAGG	AATTCGCCATGAAGAAGGCTAAAAA
rs3827797	AGTCAGGAATCTCAGGCTC	TAGCCACAGTTGGATTCCATATTCC	TACTGGAATGACCCTCATACCCAAG
rs4870052	CTAAAAATGTCTTGAACAACTCA	TCTGCTGTCTTCTTTCTCACCCTAAAA	TGCAATGAATGACCCAGATGTTATG
rs490361	CACAGTGTTATCTGCTCTATACTTC	ATATTACAGACATGAGCCGCCACAC	TGAAATGATTCCCTAGTGCCTTTGA
rs508534	TGGGGATTTGCTTAAAGAA	TTCTCCCACTTGACAGAACCTGTTT	TGATGTGAATTATAACGGCTGCTGA
rs567049	TGAAAGGAGTCCATACCATT	GAAAGACCCAATGTCAGACTGAGGA	TTTTTGCCAAGTTGAATCCAATACG
rs652070	CCTTTCTCCCTCACCCC	TTGGCCTCTCATCCTTTTACATCTG	GTCAGTCATCACTCGGCAAAGGTAT
rs764262	GCAAGTCTCCTTTACAAGTATAAAA	AATGCTCTCCCTGGTGAAGATCC	ATCCAATTACGTCCTGGCTGTCA
rs764263	GCAGCGAAGCAGGTATTA	AATGCTCTCCCTGGTGAAGATCC	ATCCAATTACGTCCTGGCTGTCA
rs892713	GAGAACTCTGAGAAGGATGC	GATCATTTCCTTGAGACTCCCCAGT	AGGGTGGACTTGGTCTCCTAACAAG
rs928501	TGACTCAGCAAATGTTTGTGT	ACAGACACATACATGCACACACACC	GTGCATCTTAACCATTTTGCACACA
rs9483497	TCAATTTGCTCATGAGATTCT	AGTCCAGCCACCTTTTAGCTACTGG	AGGCCATAAAATCCAAAGATCAGGA
rs960711	CCCTAGTTCCAGGAAATGTA	CTTCATTCTCCCAGCCCCTAGTTC	GATTTGGGGAAAGCCAGTAAACTCTT
rs9908	TTATTGCACTTCTTTTTGAGAA	TTGTTTCCAAAAGAACTCACCCTCA	TGAGGCAACTGTGACATTATGCAAT

SNP, single nucleotide polymorphism.

one was haplotype specific for a difference between a specific haplotype against all other haplotypes combined; the other was a global test for a difference among all of the haplotypes that generalized the score statistic to a Chi-square statistic. The latter test was robust to false positives that could arise from multiple testing using the first type of test.

RESULTS

Characteristics of the study subjects

Means and standard deviations for relevant anthropometric, behavioral, and biochemical risk factors are presented in **Table 2**. The mean HDL_3-C concentrations were 38.5 and 46.2 mg/dl in men and women, respectively. Men had higher values for BMI, percentage of cigarette smokers, alcohol consumption, physical activity index, percentage on anticholesterol treatment, and percentage using β-blockers.

TABLE 2. Anthropometric, behavioral, and biochemical characteristics of men and women in the 330 families

Variable	Mean ± SD	
	Men (n = 695)	Women (n = 678)
HDL_3 (mg/dl)	38.5 ± 8.9	46.2 ± 10.6
HDL_2 (mg/dl)	5.0 ± 3.8	9.4 ± 5.4
Age (years)	47.9 ± 10.2	48.9 ± 10.3
Body mass index (kg/m^2)	27.6 ± 3.8	26.1 ± 5.6
Cigarette smoker (%)	71	62
Alcohol consumption (ounces/week)	4.1 ± 5.5	1.8 ± 2.6
Physical activity index	37.8 ± 7.7	36.8 ± 6.0
Anticholesterol therapy (%)	3	2
β-Blockers (%)	11	6
Oral contraceptive (%)	NA	5
Estrogen therapy (%)	NA	6
Menopausal status (%)	NA	48

NA, not applicable.

Heritability and genome-wide linkage analysis with microsatellite markers

The heritability estimate of fully adjusted HDL_3-C concentration was 0.43 with a standard error of 0.08. Maximum multipoint LOD scores that were greater than 1 from the genome-wide linkage analyses using the scan with microsatellite markers are presented in **Table 3**. The highest multipoint LOD score was 3.7 at 133 cM on chromosome 6 (Fig. 1), nearest to marker GATA23F08. The next highest LOD scores are on chromosome 3 (LOD = 1.8 at 5 cM) and chromosome 8 (LOD = 1.6 at 8 cM).

TABLE 3. Maximum multipoint LOD scores > 1 from the genome-wide linkage analyses

Chromosome	Location (Centimorgan)	Multipoint LOD Score
2	18	1.3
3	5	1.8
6	133	3.7
8	8	1.6
15	5	1.0
16	33	1.3
16	100	1.1

We conducted additional linkage analyses for HDL_2-C and HDL-C. Linkage signals for HDL-C were found on the same location as HDL_3-C, but none of the LOD scores was >2. There was only suggestive evidence of linkage on chromosome 11 (LOD = 2.2 at 91 cM) for fully adjusted HDL_2. We did not proceed with fine mapping for these two traits.

Identification of candidate genes in the linkage region

Nine genes that code for proteins that are candidates, either because of a likely direct role in regulating HDL_3-C level or a more indirect mechanism mediated via a role in some aspect of lipid biochemistry, were identified between 125 and 150 cM. These nine candidate genes are *NCOA7*, *CTGF*, *VNN1*, *VNN3*, *VNN2*, *AIG-1*, *PLAGL1*, *LOC285746*, and *FLJ14735* (*LRP11*). Locations and annotations of those candidate genes are listed in **Table 4**.

Determination of the genetic map for the SNPs and microsatellite markers

The physical and genetic locations for both SNPs and existing microsatellite markers in the 2-LOD region are presented in Table 4.

In **Fig. 2**, we plotted the multipoint LOD scores as well as FBAT *P* values from association analyses with both physical and genetic locations on the same figure. Locations of the SNPs and microsatellite markers are also marked on the figure.

LD structure at the candidate genes

Twenty-eight SNPs in these candidate genes were selected and typed for further analyses (Table 4). All of the SNPs were in HWE or close to HWE ($P > 0.02$). LD structure of SNPs in the nine candidate genes is displayed in **Fig. 3**.

TABLE 4. Candidate genes, their functions and locations, SNPs, and nearest microsatellite markers of the 2-LOD support interval of the highest linkage peak on chromosome 6

			SNPs and Microsatellites					
Gene	Our Annotation	Position on NT_025741	Name	Physical Location	Genetic Location	Minor Allele Frequency[a]	FBAT *P*[a]	Multipoint LOD Score[b]
GATA31	Microsatellite marker	17048614	GATA31	17048614	118.600		0.21	0.7
NCOA7	Endoplasmic reticulum-associated protein 140	30206736–30357605	rs490361	30257682	126.126	0.31	0.64	3.3
			rs2065168	30295622	126.147	0.33	0.66	3.3
			rs11963811	30319530	126.161	0.2	0.58	3.3
			rs1567	30354343	126.181	0.15	0.27	3.3
GATA23F08	Microsatellite marker	35090320	GATA23F08	35090320	128.900		0.43	4.0
CTGF	Connective tissue growth factor	36373745–36376941	rs15960	36374997	130.799	0.04	0.20	3.4
			rs928501	36379866	130.806	0.25	0.35	3.4
VNN1	Vanin 1, C-N hydrolase, glycosylphosphatidylinositol-anchored regulator of hematopoietic precursor cell entry into thymus	37107158–37139617	rs2745427	37110488	131.808	0.28	0.83	2.6
			rs2272996	37119700	131.821	0.25	0.92	2.6
			rs2294757	37139527	131.848	0.34	0.34	2.5
VNN3	Vanin 3, C-N hydrolase of uncharacterized specificity	37148357–37160332	rs2294759	37150331	131.863	0.33	0.76	2.5
			rs764262	37160039	131.876	0.31	0.68	2.5
			Rs764263	37160041	131.877	0.3	0.78	2.5
			Rs960711	37161975	131.878	0.5	0.90	2.5
VNN2	Vanin 2, C-N hydrolase of uncharacterized specificity	37169439–37189027	rs9483497	37173894	131.895	0.16	0.77	2.5
			rs1883617	37177079	131.899	0.35	0.81	2.5
GATA32B03	Microsatellite marker		GATA32B03	41406690	137.700		0.72	2.8
AIG-1	Androgen-induced gene	47484929–47768403	rs3827797	47489394	142.468	0.11	0.67	2.0
			rs2876549	47585456	142.543	0.5	0.76	2.0
			rs892713	47690436	142.626	0.28	0.26	2.0
			rs9908	47765595	142.685	0.2	0.72	2.0
PLAGL1	Putative coactivator of hormone-dependent nuclear receptors	48365866–48433982	rs1884087	48363048	143.153	0.39	0.66	1.9
			rs2257104	48369369	143.158	0.1	0.03	1.9
			rs2076684	48373970	143.161	0.45	0.45	1.9
			rs2064495	48374430	143.162	0.21	0.53	1.9
GATA184A08	Microsatellite marker	52071660	GATA184A08	52071660	146.060		0.39	2.3
LOC285746	Putative fatty acid binding protein	53698950–53736572	rs508534	53724589	148.340	0.34	0.68	1.0
			rs567049	53726379	148.343	0.41	1.00	1.0
			rs652070	53736438	148.357	0.46	0.21	1.0
FLJ14735	Lipoprotein-type receptor	54294784–54340330	rs4870052	54315964	149.156	0.35	0.83	1.0
			rs2095376	54344024	149.195	0.41	0.63	1.0
GATA165G02	Microsatellite marker	58291068	GATA165G02	58291068	154.640		0.72	0.9

[a] For SNPs only.
[b] For both microsatellite markers and SNPs.

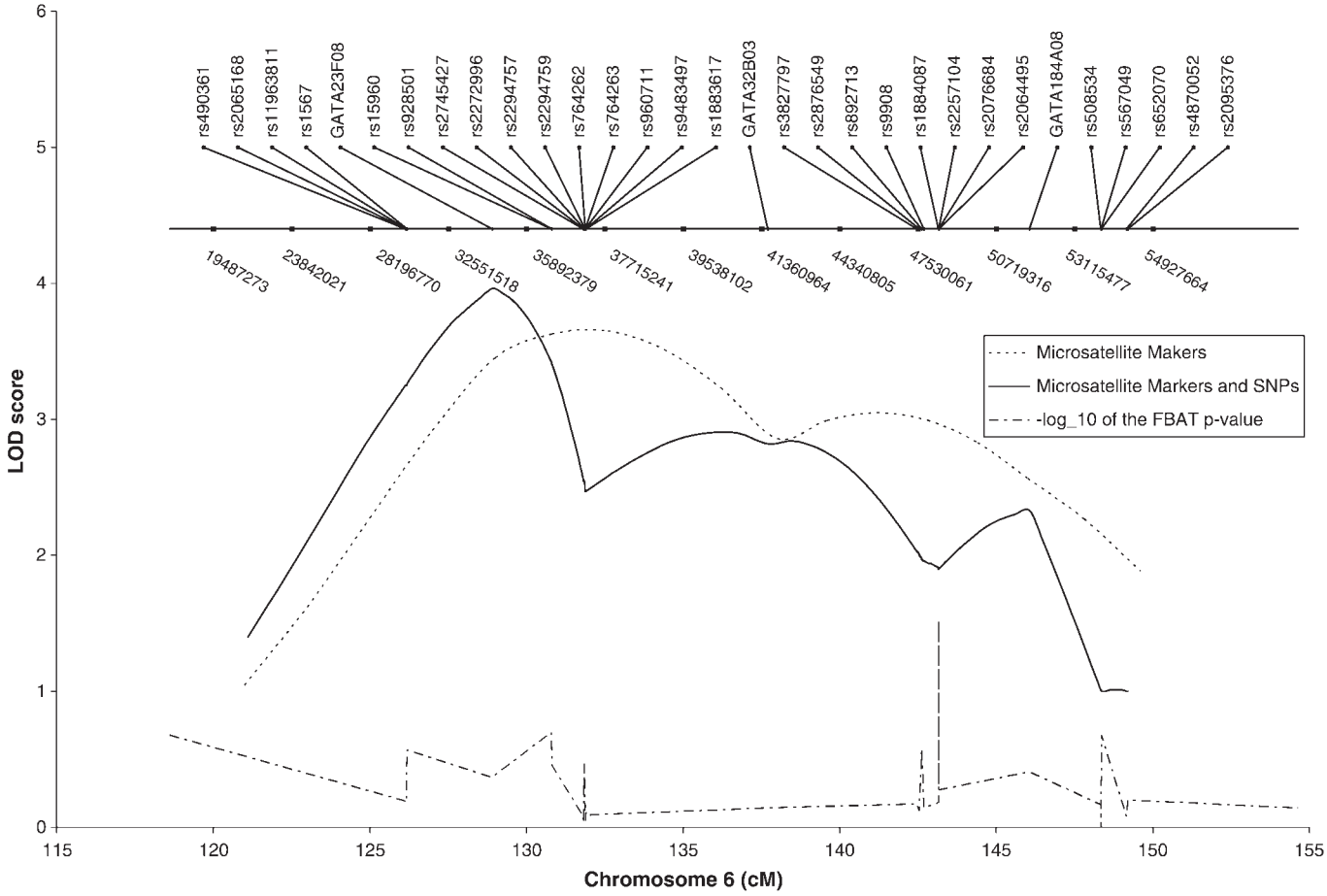

Fig. 2. Multipoint log-of-the-odds (LOD) scores, $-\log_{10}$ of FBAT *P* values in the 2-LOD region on chromosome 6 with both physical and genetic locations. The scale above the LOD score curves is the physical map. The position of each SNP or microsatellite marker is marked on the physical map by a line connecting the name of the marker and the position on the physical map.

The LD between SNPs within the same candidate gene was generally strong ($D' > 0.8$) except for *CTGF* and *AIG-1*. For *VNN1*, *VNN3*, and *VNN2*, which were located within 82 kb of each other, strong LD existed only between SNPs in *VNN3* and *VNN2*.

Fine mapping added 28 SNPs in the candidate genes to the existing map

Compared with existing microsatellite markers, incorporating the 28 SNPs into the existing 10 cM map increased the linkage information content of genotype data in the region (**Fig. 4**). The largest increases occurred at ~134 and 140 cM, and each was ~0.1.

The multipoint LOD scores on chromosome 6 after incorporating the SNPs into the existing map are presented in Fig. 1. The highest LOD score on chromosome 6 was 4.0 at 129 cM on the microsatellite marker GATA23F08. The 2-LOD support interval spans between 124 and 144 cM and covers seven of the nine candidate genes: *NCOA7*, *CTGF*, *VNN1*, *VNN3*, *VNN2*, *AIG-1*, and *PLAGL1*.

Significant association ($P < 0.05$) was found with two SNPs, rs15960 (*CTGF*) and rs2257104 (*PLAGL1*), in family-based association analyses. The A allele of rs15960 was associated with lower levels of HDL_3-C ($Z = -2.95$, $P = 0.0032$) in crude analyses without adjusting for traditional risk factors. The association was not significant after adjusting for traditional risk factors ($Z = -1.3$, $P = 0.2$). The A allele of rs2257104 has an additive effect associated with lower levels of HDL_3-C ($Z = -2.2$, $P = 0.026$) in a fully adjusted model.

However, little of the linkage signal was explained by rs2257104. The proportion of total phenotypic variation explained by rs2257104 is 0.007. After adjusting for rs2257104, the difference in the multivariate LOD score at the SNP location is ~0.04.

In haplotype analyses of SNPs within a candidate gene, significant global *P* values ($P < 0.05$) were only found with SNPs in PLAGL1 for fully adjusted HDL_3-C (**Table 5**) by assuming a dominant effect with each haplotype. There were six haplotypes with frequencies >0.01. The haplotype h6 containing C, A, C, and T alleles at rs1884087, rs2257104, rs2076684, and rs2064495, respectively, had a frequency of 0.02 and was associated with lower HDL_3-C levels (haplotype-specific $P = 0.02$). A borderline significant global *P* value ($P = 0.07$) was found with haplotypes in CTGT that contained rs15960 and rs928501, again only in the crude model.

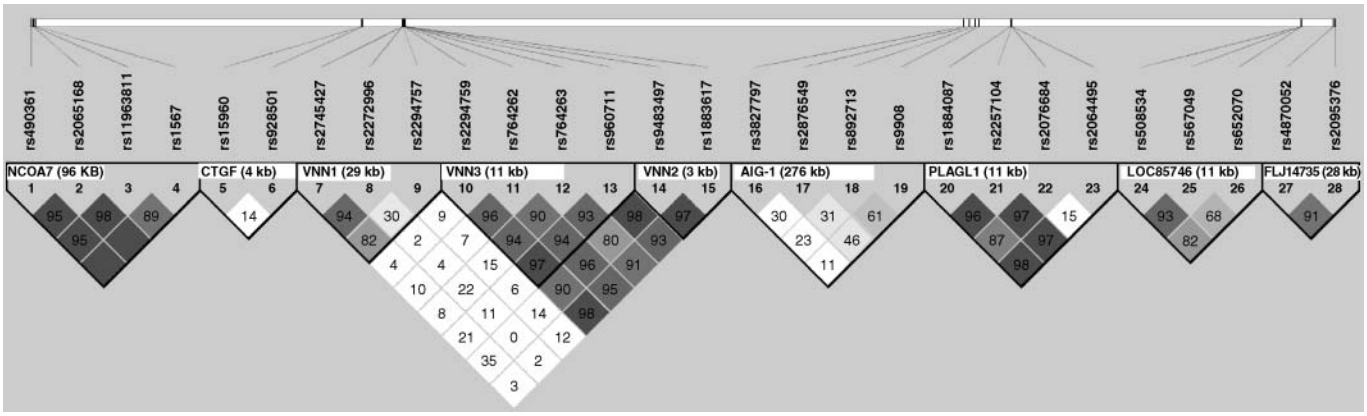

Fig. 3. Linkage disequilibrium (LD) structure at candidate genes. The horizontal white bar at the top indicates the total physical length of the genomic region, and the small vertical black lines in the bar indicate the positions of the SNP markers. The SNP makers are numbered consecutively according to their physical locations, with their SNP name on top. The number in each square is $D' \times 100$ between two markers, and the color indicates the significance of $D' > 0$. For squares without numbers, the pairwise D' is 1. Black or shades of gray indicate log-of-the-odds (LOD) ≥ 2, and white indicates LOD < 2. Only LD of pairs of SNPs within a candidate gene or within 500 kb apart are displayed.

DISCUSSION

We have found significant evidence of linkage at 133 cM on chromosome 6 for HDL_3-C concentration using a 10 cM microsatellite marker map. This result was further confirmed by fine mapping using SNPs discovered among candidate genes located within the QTL region. Only two previously published studies found evidence of linkage to HDL-C or its subclasses on chromosome 6. Coon et al. (37) reported a maximum LOD score of 1.82 at 107 cM on chromosome 6 for HDL-C. Canizales-Quinteros et al. (38) recently identified a linkage region between 73 and 80 cM on chromosome 6 predisposing to increased HDL-C levels. These two regions were ~30 and 60 cM, respectively, from our HDL_3-C linkage peak on chromosome 6.

The region of chromosome 6 in which we obtained the highest LOD is close to the region where a 4.64 LOD score at 144 cM was found for BMI measured at examination cycle 1 in the Framingham Heart Study (39). Our results were based on the HDL_3-C measured during examination cycle 4, 16 years after the time of examination cycle 1. For BMI measured at examination cycle 4, the highest LOD score on chromosome 6 was also obtained at the same region, but the value of the LOD scores was 1.43. BMI was used as a covariate in our analysis, so our linkage result on chromosome 6 is not likely to be confounded by BMI.

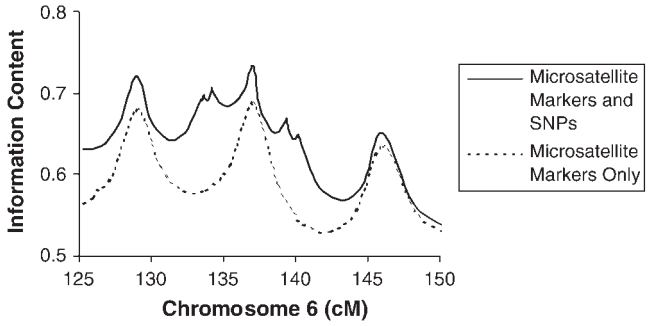

Fig. 4. Information content of chromosome 6 marker genotype data between 125 and 149 cM with microsatellite markers only and with both microsatellite and SNP markers.

There has been a great deal of heterogeneity in the linkage findings of HDL-C and its subparticles. Significant findings (LOD > 3) were found on chromosomes 8 and 15 using 477 Mexican Americans of the San Antonio Family Heart Study (11), on chromosome 5 using 1,027 Caucasians of the Family Heart Study (6), on chromosome 12 using 534 pairs of siblings of the Quebec Family Study, on chromosome 10q11 using 1,109 individuals from 92 low HDL-C or hyperlipidemia families, and on chromosome 6 between 73 and 80 cM for increased HDL-C levels (38). It is difficult to explain the lack of replication across these studies. Multiple factors, including phenotype definition, ethnicity, gene-environment interaction, power, and false positives, could contribute to the lack of replication of significant results. Although there is no simple solution to all of these problems, fine mapping in the linkage region may prove an effective tool to discover the existence of any functional variants as the cost of genotyping declines.

The genetic linkage data for HDL_3-C was used to select a region on chromosome 6 from 6q22.33 to 6q24.3, and a bioinformatics analysis produced nine candidate genes. Each of these genes encodes a protein with the likelihood of playing a role in lipid homeostasis by physical interaction with lipid moieties, by regulating lipid biosynthesis or catabolism, by stearoyl hormone-dependent regulation of gene expression, or by association with manifestations of cardiovascular disease. By including the 28 SNPs in the nine candidate genes into the linkage analyses, the linkage information content of the marker data has been improved, thereby providing stronger information with which to evaluate linkage. Furthermore, the increase in the LOD score after incorporating the SNPs provides additional evidence that at least one QTL exists in this region.

Family-based association analyses revealed that two

TABLE 5. Haplotype analyses of SNPs at PLAGL1 with fully adjusted high density lipoprotein-3 cholesterol

Label of Haplotype	rs1884087	rs2257104	rs2076684	rs2064495	Frequency	Number of Informative Families	Haplotype-Specific Z Statistic	Haplotype-Specific *P*	Global Chi-Square Statistic	Global *P*
h1	G	C	G	T	0.359	137	0.9	0.390	12.7	0.049
h2	A	A	G	T	0.329	125	−1.5	0.145		
h3	A	C	G	T	0.094	57	1.7	0.085		
h4	A	C	G	C	0.086	72	1.2	0.212		
h5	A	A	C	C	0.085	61	−1.3	0.181		
h6	G	A	G	T	0.022	18	−2.4	0.017		

SNPs in two of the candidate genes were associated with HDL_3-C. However, only rs2257104 in *PLGAL1* was significant after adjusting for all of the traditional risk factors. *PLAGL1*, or pleiomorphic adenoma gene-like 1, encodes a putative coactivator of hormone-dependent nuclear receptors with several C_2H_2-type zinc finger domains. SNP rs2257104 is located at ~143 cM, which had a LOD score of 1.9 in the linkage analyses incorporating SNPs into the current map. For the SNPs near the maximum LOD score at 128 cM, we did not find any association for fully adjusted HDL_3-C. Although it is possible that the SNPs we typed near the location of the maximum LOD score are not in linkage disequilibrium with a QTL at that location, it is also possible that the QTL was not at the location of the maximum LOD score. Simulation studies (40) have shown that compared with a 10 cM map, further fine mapping using linkage analysis did not result in much reduction in the average location error of maximum LOD score from the QTL. Although only nine candidate genes in the 25 cM interval were investigated and there was considerable distance between these candidate genes, we feel that these nine genes represent the best candidates based on extensive bioinformatics analysis. Nonetheless, we may have missed genes that are located between or outside of these candidate genes but are associated with HDL_3-C. Increasing the density of SNP by genotyping more SNPs within and outside of the nine candidate genes in the linkage region would increase the likelihood of detecting a true association.

We noted that the A allele of rs2257104 resided on all three haplotypes (h2, h5, and h6) with negative Z statistics (Table 5), which was consistent with the result from single SNP association analyses of this marker. However, only the h6 haplotype was statistically significant compared with all others combined. None of the other SNPs in PLAGL1 had an allele that resided on haplotypes that corresponded only to positive or only to negative Z statistics. In sum, these results suggested that the significance of the global haplotype test is most likely attributable to rs2257104.

The estrogen receptor initiates transcription of select genes after binding of ligand and translocation of the receptor-ligand complex to the nucleus (41). Important to lipid physiology, it has been demonstrated that estrogen, via estrogen receptor and coactivators, promotes transcription of the *ApoE* (42) and *ApoA1* genes (43–45). However, for the following reasons, we do not believe that the gene encoding the estrogen receptor (*ESR1*) is directly responsible for the genetic determinant of HDL_3-C levels: *ESR1* maps to 154.8 cM (6q25.1) or ~7 Mb beyond our 2-LOD interval; *ESR1* is at a LOD score of 0.9, whereas the maximum LOD score seen here is 4.0 after incorporating the SNPs in the linkage analyses. Furthermore, a reanalysis of the linkage data in which we adjusted for either or both of the genotypes of two common variants in *ESR1*, *PvuII* and *C1335G*, generated a similar maximum LOD score (difference in LOD score of <0.17) on chromosome 6.

The study subjects are those with *CETP* polymorphism typed, because we were also interested in testing whether some of the known lipid candidate genes, such as *CETP* and hepatic lipase, could explain the linkage peaks. We did not find that adjusting for *CETP* or hepatic lipase changed the linkage signals much on chromosome 6, which suggested lack of evidence for interaction between QTLs on chromosome 6 and these two known lipid candidate genes. However, when we used all of the subjects with HDL_3-C without regard to whether they were typed for *CETP* genotype, the maximum LOD score decreased to 1.8 on chromosome 6 in the analyses with the microsatellite markers. Subjects with CETP genotyped were younger on average (48 vs. 50 years) than subjects with missing CETP, which may have resulted in a more homogeneous group to yield better linkage signals. Otherwise, there were no significant differences between those with CETP genotyping and those without.

Regulation of HDL metabolism, lipid homeostasis, and the determinants of cardiovascular health is complex and undoubtedly involves the functions and interactions of many different genes. On the long arm of chromosome 6 are several genes, which were not considered in this study, but whose roles in regulating cardiovascular health via lipid levels ought to be considered as candidates in affecting lipid metabolism. The LPA gene cluster, including APOARGC (apolipoprotein A-related gene C), LPA [lipoprotein Lp(a)], and possible pseudogenes LPAL1 and LPAL2, encodes genes in which certain variations are known to be risk factors for cardiovascular disease (46) but that are not related to the HDL_3-C trait presented in this study. In addition, this cluster maps to 6q26-q27, ~10 Mb downstream of LETAL. Several IDDM (insulin-dependent diabetes mellitus) loci have been identified but not narrowed to a specific gene. IDDM5, IDDM8, and IDDM15, although mapping to 6q24-q27, 6q25-q27, and 6q21, respectively, are not considered further because the gene or genetic element responsible for the observed phenotype

has not been discovered and because, unlike type II diabetes, which has been shown to have important links to cardiovascular disease via HDL, such a connection to type I diabetes has not been established.

A number of previous genome-wide linkage analyses of cardiovascular anthropometric and lipid phenotypes were based on the same genome scan used in this study. The phenotypes studied included triglycerides (47), BMI (39), waist circumference (48), and blood pressure (49, 50). Our finding of significant linkage on chromosome 6 may be subject to an inflated false-positive rate as a result of multiple testing. Because these anthropometric or lipid phenotypes may not be independent, however, a Bonferroni correction would be too conservative and is thus not suitable for correction of multiple testing here. As the LOD score increased in the fine mapping, the likelihood of a false positive is reduced for our study (51). Furthermore, the chromosome 6 location of the peak LOD score is close to the locations of peak LOD scores of BMI (39) that may be in the lipid metabolism pathway. Considering all of these factors, we believe that our results, although not completely conclusive, can still be a valuable reference for further research in HDL_3 or lipid metabolism in general.

In summary, a genome-wide scan of 401 microsatellite markers in 330 extended families of the Framingham Heart Study has revealed a promising and statistically significant linkage with HDL_3-C concentrations on chromosome 6 (multipoint LOD = 3.7 at 133 cM). None of the classical genes associated with HDL metabolism are located within this region. Bioinformatic analyses of the region between 125 and 150 cM suggested the presence of nine interesting genes in which we further genotyped 28 SNPs. Linkage analyses incorporating those SNPs into the current marker map increased the linkage information of marker data in that region and resulted in a multipoint LOD score of 4.0 at 129 cM. Family-based association analyses revealed that SNP rs2257104 in PLAGL1 was associated with fully adjusted HDL_3-C. Further study of variants in PLAGL1 and increased SNP density in the linkage region are warranted to more clearly define the potential functional variants.

The authors thank all of the individuals who participated in the Framingham Heart Study. This work was supported in part by Contracts N01-HC-25195 and 1-38038 and Grants HL-54776, P50-HL-63494, and RO1-HL-65230 from the National Heart, Lung, and Blood Institute and by Contracts 53-K06-5-10 and 58-1950-9-001 from the U. S. Department of Agriculture Research Service.

REFERENCES

1. Miller, N. E. 1987. Associations of high-density lipoprotein subclasses and apolipoproteins with ischemic heart disease and coronary atherosclerosis. *Am. Heart J.* **113:** 589–597.
2. Cupples, L. A., and R. H. Myers. 1993. Segregation analysis for high density lipoprotein in the Berkeley data. *Genet. Epidemiol.* **10:** 629–634.
3. Xu, J. F., V. L. Prenger, J. Coresh, and T. H. Beaty. 1993. Two-locus and bivariate segregation analysis of HDL-C and apo AI. *Genet. Epidemiol.* **10:** 605–610.
4. Williams, R. R., S. C. Hunt, P. N. Hopkins, L. L. Wu, S. J. Hasstedt, T. D. Berry, G. K. Barlow, B. M. Stults, M. C. Schumacher, and E. H. Ludwig. 1993. Genetic basis of familial dyslipidemia and hypertension: 15-year results from Utah. *Am. J. Hypertens.* **6 (Suppl. 1):** 319–327.
5. Mahaney, M. C., J. Blangero, D. L. Rainwater, A. G. Comuzzie, J. L. VandeBerg, M. P. Stern, J. W. MacCluer, and J. E. Hixson. 1995. A major locus influencing plasma high-density lipoprotein cholesterol levels in the San Antonio Family Heart Study. Segregation and linkage analyses. *Arterioscler. Thromb. Vasc. Biol.* **15:** 1730–1739.
6. 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.
7. Delalla, O. F., H. A. Elliot, and J. W. Gofman. 1954. Ultracentrifugal studies of high density serum lipoproteins in clinically healthy adults. *Am. J. Physiol.* **179:** 333–337.
8. Handbook of Lipoprotein Testing. American Association of Clinical Chemistry, Washington, 1997.
9. Wilson, P. W. 1995. Relation of high-density lipoprotein subfractions and apolipoprotein E isoforms to coronary disease. *Clin. Chem.* **41:** 165–169.
10. Coresh, J., T. H. Beaty, V. L. Prenger, J. Xu, and P. O. Kwiterovich, Jr. 1995. Segregation analysis of HDL3-C levels in families of patients undergoing coronary arteriography at an early age. *Arterioscler. Thromb. Vasc. Biol.* **15:** 1307–1313.
11. Almasy, L., J. E. Hixson, D. L. Rainwater, S. Cole, J. T. Williams, M. C. Mahaney, J. L. VandeBerg, M. P. Stern, J. W. MacCluer, and J. Blangero. 1999. Human pedigree-based quantitative-trait-locus mapping: localization of two genes influencing HDL-cholesterol metabolism. *Am. J. Hum. Genet.* **64:** 1686–1693.
12. Gerdes, C., L. U. Gerdes, P. S. Hansen, and O. Faergeman. 1995. Polymorphisms in the lipoprotein lipase gene and their associations with plasma lipid concentrations in 40-year-old Danish men. *Circulation.* **92:** 1765–1769.
13. Minnich, A., G. DeLangavant, J. Lavigne, G. Roederer, S. Lussier-Cacan, and J. Davignon. 1995. GÆA substitution at position −75 of the apolipoprotein A-I gene promoter. Evidence against a direct effect on HDL cholesterol levels. *Arterioscler. Thromb. Vasc. Biol.* **15:** 1740–1745.
14. Kamboh, M. I., C. E. Aston, C. M. Nestlerode, A. E. McAllister, and R. F. Hamman. 1996. Haplotype analysis of two APOA1/MspI polymorphisms in relation to plasma levels of apo A-I and HDL-cholesterol. *Atherosclerosis.* **127:** 255–262.
15. McPherson, R., S. M. Grundy, R. Guerra, and J. C. Cohen. 1996. Allelic variation in the gene encoding the cholesteryl ester transfer protein is associated with variation in the plasma concentrations of cholesteryl ester transfer protein. *J. Lipid Res.* **37:** 1743–1748.
16. Guerra, R., J. Wang, S. M. Grundy, and J. C. Cohen. 1997. A hepatic lipase (LIPC) allele associated with high plasma concentrations of high density lipoprotein cholesterol. *Proc. Natl. Acad. Sci. USA.* **94:** 4532–4537.
17. Devlin, C. M., V. L. Prenger, and M. Miller. 1998. Linkage of the apo CIII microsatellite with isolated low high-density lipoprotein cholesterol. *Hum. Genet.* **102:** 273–281.
18. Ordovas, J. M., L. A. Cupples, D. Corella, J. D. Otvos, D. Osgood, A. Martinez, C. Lahoz, O. Coltell, P. W. Wilson, and E. J. Schaefer. 2000. Association of cholesteryl ester transfer protein-TaqIB polymorphism with variations in lipoprotein subclasses and coronary heart disease risk: the Framingham Study. *Arterioscler. Thromb. Vasc. Biol.* **20:** 1323–1329.
19. Dawber, T. R., G. F. Meadors, and F. E. J. Moore. 1951. Epidemiological approaches to heart disease: the Framingham Study. *Am. J. Public Health.* **41:** 279–286.
20. Kannel, W. B., M. Feinleib, P. M. McNamara, R. J. Garrison, and W. P. Castelli. 1979. An investigation of coronary heart disease in families. The Framingham Offspring Study. *Am. J. Epidemiol.* **110:** 281–290.
21. Gidez, L. I., G. J. Miller, M. Burstein, S. Slagle, and H. A. Eder. 1982. Separation and quantitation of subclasses of human plasma high density lipoproteins by a simple precipitation procedure. *J. Lipid Res.* **23:** 1206–1223.
22. Kruglyak, L., M. J. Daly, M. P. Reeve-Daly, and E. S. Lander. 1996. Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am. J. Hum. Genet.* **58:** 1347–1363.

23. Abecasis, G. R., S. S. Cherny, W. O. Cookson, and L. R. Cardon. 2002. Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. *Nat. Genet.* **30:** 97–101.
24. Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25:** 3389–3402.
25. Bateman, A., E. Birney, L. Cerruti, R. Durbin, L. Etwiller, S. R. Eddy, S. Griffiths-Jones, K. L. Howe, M. Marshall, and E. L. Sonnhammer. 2002. The Pfam protein families database. *Nucleic Acids Res.* **30:** 276–280.
26. Mulder, N. J., R. Apweiler, T. K. Attwood, A. Bairoch, D. Barrell, A. Bateman, D. Binns, M. Biswas, P. Bradley, P. Bork, et al. 2003. The InterPro Database, 2003 brings increased coverage and new features. *Nucleic Acids Res.* **31:** 315–318.
27. Krogh, A., B. Larsson, G. von Heijne, and E. L. Sonnhammer. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305:** 567–580.
28. Emanuelsson, O., H. Nielsen, S. Brunak, and G. von Heijne. 2000. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.* **300:** 1005–1016.
29. Nielsen, H., J. Engelbrecht, S. Brunak, and G. von Heijne. 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* **10:** 1–6.
30. Tai, E. S., J. M. Ordovas, D. Corella, M. Deurenberg-Yap, E. Chan, X. Adiconis, S. K. Chew, L. M. Loh, and C. E. Tan. 2003. The TaqIB and −629C>A polymorphisms at the cholesteryl ester transfer protein locus: associations with lipid levels in a multiethnic population. The 1998 Singapore National Health Survey. *Clin. Genet.* **63:** 19–30.
31. Broman, K. W., J. C. Murray, V. C. Sheffield, R. L. White, and J. L. Weber. 1998. Comprehensive human genetic maps: individual and sex-specific variation in recombination. *Am. J. Hum. Genet.* **63:** 861–869.
32. Weiss, K. M., and A. G. Clark. 2002. Linkage disequilibrium and the mapping of complex human traits. *Trends Genet.* **18:** 19–24.
33. Rabinowitz, D., and N. Laird. 2000. A unified approach to adjusting association tests for population admixture with arbitrary pedigree structure and arbitrary missing marker information. *Hum. Hered.* **50:** 211–223.
34. Laird, N. M., S. Horvath, and X. Xu. 2000. Implementing a unified approach to family-based tests of association. *Genet. Epidemiol.* **19 (Suppl. 1):** 36–42.
35. Lake, S. L., D. Blacker, and N. M. Laird. 2000. Family-based tests of association in the presence of linkage. *Am. J. Hum. Genet.* **67:** 1515–1525.
36. Horvath, S., X. Xu, S. L. Lake, E. K. Silverman, S. T. Weiss, and N. M. Laird. 2004. Family-based tests for associating haplotypes with general phenotype data: application to asthma genetics. *Genet. Epidemiol.* **26:** 61–69.
37. Coon, H., M. F. Leppert, J. H. Eckfeldt, A. Oberman, R. H. Myers, J. M. Peacock, M. A. Province, P. N. Hopkins, and G. Heiss. 2001. Genome-wide linkage analysis of lipids in the Hypertension Genetic Epidemiology Network (HyperGEN) Blood Pressure Study. *Arterioscler. Thromb. Vasc. Biol.* **21:** 1969–1976.
38. Canizales-Quinteros, S., C. A. Aguilar-Salinas, E. Reyes-Rodriguez, L. Riba, M. Rodriguez-Torres, S. Ramirez-Jimenez, A. Huertas-Vazquez, V. Fragoso-Ontiveros, A. Zentella-Dehesa, J. L. Ventura-Gallegos, et al. 2003. Locus on chromosome 6p linked to elevated HDL cholesterol serum levels and to protection against premature atherosclerosis in a kindred with familial hypercholesterolemia. *Circ. Res.* **92:** 569–576.
39. Atwood, L. D., N. L. Heard-Costa, L. A. Cupples, C. E. Jaquish, P. W. Wilson, and R. B. D'Agostino. 2002. Genomewide linkage analysis of body mass index across 28 years of the Framingham Heart Study. *Am. J. Hum. Genet.* **71:** 1044–1050.
40. Atwood, L. D., and N. L. Heard-Costa. 2003. Limits of fine-mapping a quantitative trait. *Genet. Epidemiol.* **24:** 99–106.
41. Metzger, D., J. H. White, and P. Chambon. 1988. The human oestrogen receptor functions in yeast. *Nature.* **334:** 31–36.
42. Srivastava, R. A., N. Srivastava, M. Averna, R. C. Lin, K. S. Korach, D. B. Lubahn, and G. Schonfeld. 1997. Estrogen up-regulates apolipoprotein E (ApoE) gene expression by increasing ApoE mRNA in the translating pool via the estrogen receptor alpha-mediated pathway. *J. Biol. Chem.* **272:** 33360–33366.
43. 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.
44. Lamon-Fava, S., J. M. Ordovas, and E. J. Schaefer. 1999. Estrogen increases apolipoprotein (apo) A-I secretion in Hep G2 cells by modulating transcription of the apo A-I gene promoter. *Arterioscler. Thromb. Vasc. Biol.* **19:** 2960–2965.
45. Zhang, X., J. J. Jiao, B. R. Bhavnani, and S. P. Tam. 2001. Regulation of human apolipoprotein A-I gene expression by equine estrogens. *J. Lipid Res.* **42:** 1789–1800.
46. Rainwater, D. L., and C. M. Kammerer. 1998. Lipoprotein Lp(a): effects of allelic variation at the LPA locus. *J. Exp. Zool.* **282:** 54–61.
47. Shearman, A. M., J. M. Ordovas, L. A. Cupples, E. J. Schaefer, M. D. Harmon, Y. Shao, J. D. Keen, A. L. DeStefano, O. Joost, P. W. Wilson, et al. 2000. Evidence for a gene influencing the TG/HDL-C ratio on chromosome 7q32.3-qter: a genome-wide scan in the Framingham study. *Hum. Mol. Genet.* **9:** 1315–1320.
48. Fox, C. S., N. L. Heard-Costa, P. W. Wilson, D. Levy, R. B. D'Agostino, Sr., and L. D. Atwood. 2004. Genome-wide linkage to chromosome 6 for waist circumference in the Framingham Heart Study. *Diabetes.* **53:** 1399–1402.
49. Levy, D., A. L. DeStefano, M. G. Larson, C. J. O'Donnell, R. P. Lifton, H. Gavras, L. A. Cupples, and R. H. Myers. 2000. Evidence for a gene influencing blood pressure on chromosome 17. Genome scan linkage results for longitudinal blood pressure phenotypes in subjects from the Framingham Heart Study. *Hypertension.* **36:** 477–483.
50. Meigs, J. B., C. I. Panhuysen, R. H. Myers, P. W. Wilson, and L. A. Cupples. 2002. A genome-wide scan for loci linked to plasma levels of glucose and HbA(1c) in a community-based sample of Caucasian pedigrees: the Framingham Offspring Study. *Diabetes.* **51:** 833–840.
51. Wiltshire, S., A. P. Morris, M. I. McCarthy, and L. R. Cardon. 2005. How useful is the fine-scale mapping of complex trait linkage peaks? Evaluating the impact of additional microsatellite genotyping on the posterior probability of linkage. *Genet. Epidemiol.* **28:** 1–10.