ORIGINAL ARTICLE

Genetic Variants Associated with Lp(a) Lipoprotein Level and Coronary Disease

Robert Clarke, F.R.C.P., John F. Peden, Ph.D., Jemma C. Hopewell, Ph.D.,
Theodosios Kyriakou, Ph.D., Anuj Goel, M.Sc., Simon C. Heath, Ph.D.,
Sarah Parish, D.Phil., Simona Barlera, M.S., Maria Grazia Franzosi, Ph.D.,
Stephan Rust, Ph.D., Derrick Bennett, Ph.D., Angela Silveira, Ph.D.,
Anders Malarstig, Ph.D., Fiona R. Green, Ph.D., Mark Lathrop, Ph.D.,
Bruna Gigante, M.D., Karin Leander, Ph.D., Ulf de Faire, M.D.,
Udo Seedorf, Ph.D., Anders Hamsten, F.R.C.P., Rory Collins, F.R.C.P.,
Hugh Watkins, F.R.C.P., and Martin Farrall, F.R.C.Path.,
for the PROCARDIS Consortium*

ABSTRACT

BACKGROUND

An increased level of Lp(a) lipoprotein has been identified as a risk factor for coronary artery disease that is highly heritable. The genetic determinants of the Lp(a) lipoprotein level and their relevance for the risk of coronary disease are incompletely understood.

METHODS

We used a novel gene chip containing 48,742 single-nucleotide polymorphisms (SNPs) in 2100 candidate genes to test for associations in 3145 case subjects with coronary disease and 3352 control subjects. Replication was tested in three independent populations involving 4846 additional case subjects with coronary disease and 4594 control subjects.

RESULTS

Three chromosomal regions (6q26–27, 9p21, and 1p13) were strongly associated with the risk of coronary disease. The *LPA* locus on 6q26–27 encoding Lp(a) lipoprotein had the strongest association. We identified a common variant (rs10455872) at the *LPA* locus with an odds ratio for coronary disease of 1.70 (95% confidence interval [CI], 1.49 to 1.95) and another independent variant (rs3798220) with an odds ratio of 1.92 (95% CI, 1.48 to 2.49). Both variants were strongly associated with an increased level of Lp(a) lipoprotein, a reduced copy number in *LPA* (which determines the number of kringle IV–type 2 repeats), and a small Lp(a) lipoprotein size. Replication studies confirmed the effects of both variants on the Lp(a) lipoprotein level and the risk of coronary disease. A meta-analysis showed that with a genotype score involving both *LPA* SNPs, the odds ratios for coronary disease were 1.51 (95% CI, 1.38 to 1.66) for one variant and 2.57 (95% CI, 1.80 to 3.67) for two or more variants. After adjustment for the Lp(a) lipoprotein level, the association between the *LPA* genotype score and the risk of coronary disease was abolished.

CONCLUSIONS

We identified two *LPA* variants that were strongly associated with both an increased level of Lp(a) lipoprotein and an increased risk of coronary disease. Our findings provide support for a causal role of Lp(a) lipoprotein in coronary disease.

GENOMEWIDE ASSOCIATION STUDIES have identified several novel susceptibility loci for coronary artery disease,[1-4] but it is likely that only common variants can be detected in this way.[5,6] Moreover, loci that are identified with the use of genomewide association studies explain only a small amount of the expected contribution to the risk of coronary disease. The use of arrays of high-density single-nucleotide polymorphisms (SNPs) in candidate genes for cardiovascular disease may help elucidate the genetic contribution to the risk of coronary disease.

A recent genomewide association study showed that a cluster of genes — solute carrier family 22 member 3 (*SLC22A3*), lipoprotein(a)-like 2 (*LPAL2*), and lipoprotein(a) (*LPA*) — on chromosome 6q26–27 was strongly associated with coronary artery disease, but the investigators were unable to identify the precise variants at this locus.[7] The 6q26–27 region includes the *LPA* gene, which encodes the apolipoprotein(a) component of the Lp(a) lipoprotein particle. An increased level of Lp(a) lipoprotein has been associated with an increased risk of coronary disease,[8,9] carotid atherosclerosis, and stroke.[9-11] Plasma levels of Lp(a) lipoprotein vary substantially among persons, and most of this variation reflects the effects of genetic variation in *LPA*.[12] In particular, common copy-number variation within the *LPA* gene determines the number of kringle IV–type 2 repeats and hence the isoform size of apolipoprotein(a), and an inverse relationship has been reported between the number of repeats and Lp(a) levels. However, the genetic determinants of Lp(a) lipoprotein levels and of isoform size are incompletely understood, as is the relevance of both measures for coronary artery disease.[12]

Our multicenter case–control study, called the Precocious Coronary Artery Disease (PROCARDIS) study, had four aims. First, we examined genetic associations in coronary artery disease, using a newly available chip[13] that was specifically designed to assay SNPs in candidate genes selected for their putative relevance to cardiovascular disease. Second, we assessed the associations of *LPA* gene variants with Lp(a) lipoprotein levels and isoform size in a large case–control study. Third, we replicated the associations in three independent studies. Finally, we assessed the extent to which the observed associations were explained by their effects on Lp(a) lipoprotein levels.

METHODS

STUDY DESIGN AND OVERSIGHT

This study was designed and conducted by the authors and funded by the British Heart Foundation, the European Commission, and AstraZeneca. AstraZeneca had no role in the design of the study, in the data collection or analysis, in the writing of the manuscript, or in the decision to submit the results for publication. The authors vouch for the accuracy and completeness of the data and the analyses.

SUBJECTS

We recruited 3145 case subjects with coronary artery disease and 3352 control subjects from four European countries (United Kingdom, Italy, Sweden, and Germany), according to prespecified criteria.[14,15] All case subjects had received a diagnosis of coronary artery disease before the age of 66 years and also had a sibling in whom coronary disease had been diagnosed before the age of 66 years (see the Methods section in the Supplementary Appendix, available with the full text of this article at NEJM.org). Using the same infrastructure, we recruited population-matched control subjects with no personal or sibling history of coronary disease before the age of 66 years. The protocol was approved by the ethics committee at each participating center, and all subjects provided written informed consent. Blood samples were collected from all subjects.

ASSAYS

SNP genotyping was performed in case subjects and control subjects with the use of the HumanCVD BeadChip (Illumina)[13] with the Infinium II assay, which includes 48,742 markers designed to survey genetic variation in approximately 2100 candidate genes selected for their relevance in cardiovascular disease (for details, see the Methods section in the Supplementary Appendix). The chip also includes SNPs in regions that have been identified by genomewide association studies as susceptibility loci for coronary disease (e.g., chromosome 9p21).[1-4]

The HumanCVD BeadChip includes 40 SNPs from the *LPA* region. After SNP exclusions for a low call rate or minimum allele frequency, 27 SNPs were available for association analysis in the *LPA* region (Fig. 1, and the Methods section in the Supplementary Appendix).
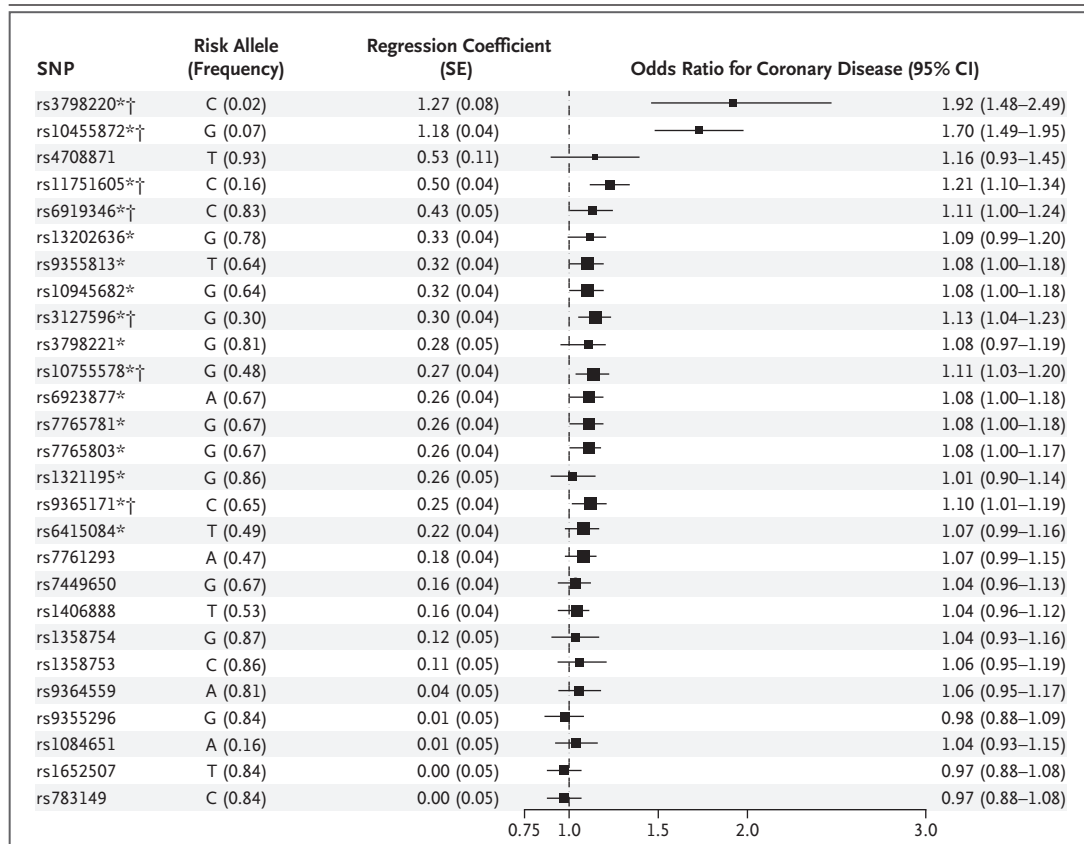
2519

SNP	Risk Allele (Frequency)	Regression Coefficient (SE)	Odds Ratio for Coronary Disease (95% CI)	
rs3798220*†	C (0.02)	1.27 (0.08)		1.92 (1.48–2.49)
rs10455872*†	G (0.07)	1.18 (0.04)		1.70 (1.49–1.95)
rs4708871	T (0.93)	0.53 (0.11)		1.16 (0.93–1.45)
rs11751605*†	C (0.16)	0.50 (0.04)		1.21 (1.10–1.34)
rs6919346*†	C (0.83)	0.43 (0.05)		1.11 (1.00–1.24)
rs13202636*	G (0.78)	0.33 (0.04)		1.09 (0.99–1.20)
rs9355813*	T (0.64)	0.32 (0.04)		1.08 (1.00–1.18)
rs10945682*	G (0.64)	0.32 (0.04)		1.08 (1.00–1.18)
rs3127596*†	G (0.30)	0.30 (0.04)		1.13 (1.04–1.23)
rs3798221*	G (0.81)	0.28 (0.05)		1.08 (0.97–1.19)
rs10755578*†	G (0.48)	0.27 (0.04)		1.11 (1.03–1.20)
rs6923877*	A (0.67)	0.26 (0.04)		1.08 (1.00–1.18)
rs7765781*	G (0.67)	0.26 (0.04)		1.08 (1.00–1.18)
rs7765803*	G (0.67)	0.26 (0.04)		1.08 (1.00–1.17)
rs1321195*	G (0.86)	0.26 (0.05)		1.01 (0.90–1.14)
rs9365171*†	C (0.65)	0.25 (0.04)		1.10 (1.01–1.19)
rs6415084*	T (0.49)	0.22 (0.04)		1.07 (0.99–1.16)
rs7761293	A (0.47)	0.18 (0.04)		1.07 (0.99–1.15)
rs7449650	G (0.67)	0.16 (0.04)		1.04 (0.96–1.13)
rs1406888	T (0.53)	0.16 (0.04)		1.04 (0.96–1.12)
rs1358754	G (0.87)	0.12 (0.05)		1.04 (0.93–1.16)
rs1358753	C (0.86)	0.11 (0.05)		1.06 (0.95–1.19)
rs9364559	A (0.81)	0.04 (0.05)		1.06 (0.95–1.17)
rs9355296	G (0.84)	0.01 (0.05)		0.98 (0.88–1.09)
rs1084651	A (0.16)	0.01 (0.05)		1.04 (0.93–1.15)
rs1652507	T (0.84)	0.00 (0.05)		0.97 (0.88–1.08)
rs783149	C (0.84)	0.00 (0.05)		0.97 (0.88–1.08)

0.75 1.0   1.5   2.0   3.0

Figure 1. Associations of Single-Nucleotide Polymorphisms (SNPs) in *LPA* with the Lp(a) Lipoprotein Level and the Risk of Coronary Disease in the PROCARDIS Cohort.

The 27 SNPs in *LPA* are listed in order of their effect on the level of Lp(a) lipoprotein. Sixteen SNPs that were significantly associated with the Lp(a) lipoprotein level ($P<1.0\times10^{-6}$) are indicated by an asterisk. Seven SNPs that were also significantly associated with the level of Lp(a) lipoprotein by stepwise regression ($P<4.0\times10^{-9}$) are indicated with a dagger. The regression coefficients are for the log of the Lp(a) lipoprotein level, with the standard error (SE) providing a measure of the sampling variation. The boxes indicate odds ratios for coronary disease, and the horizontal lines indicate 95% confidence intervals (CI) for the allele associated with an increased Lp(a) lipoprotein level for each of the SNPs. The associations have been adjusted for the country of origin. The size of the squares is inversely proportional to the sampling variance, which is influenced by the frequency of the SNP. Thus, even though SNP rs4708871 has the fourth highest odds ratio in the list, the low frequency of this SNP means that the sampling error is large (as indicated by the large confidence interval), so the association with coronary disease was not significant.

Lp(a) lipoprotein was measured by means of one latex-enhanced immunoturbidimetric assay (Immuno)[16] in samples from case subjects that had been obtained in study clinics. In addition, Lp(a) lipoprotein was measured in a random subgroup of case subjects and control subjects with the use of a second latex-enhanced immunoturbidimetric assay (Randox Laboratories) on an ADVIA 1800 autoanalyzer (Siemens).

We estimated the size of the apolipoprotein(a) isoform by means of immunoblotting in samples from case subjects and control subjects who were selected according to *LPA* genotype. Sodium dodecyl sulfate–agarose gel electrophoresis was used to fractionate the reduced plasma proteins according to size before immunoblotting with apolipoprotein(a)-specific antibody.[17] A quantitative polymerase-chain-reaction (PCR) assay with genomic DNA as the template was used to determine the relative number of *LPA* kringle IV–type 2 repeats in a subgroup of subjects for whom isoform data were available.[18]

REPLICATION

We tested the replication of the associations in three independent populations that totaled 4846 case subjects with coronary disease and 4594 control subjects. These independent cohorts included

"trio families" (families either with a proband and two parents or with a proband, a parent, and at least one sibling) from the PROCARDIS cohort[19]; case subjects with nonfatal myocardial infarction and control subjects from the International Study of Infarct Survival (ISIS) in the United Kingdom[20]; and case subjects with nonfatal myocardial infarction and control subjects from two studies of persons living in the greater Stockholm area (the Stockholm Heart Epidemiology Program [SHEEP][21] and the Stockholm Coronary Artery Risk Factor [SCARF] study[22]) (for details, see the Methods section in the Supplementary Appendix).

We measured plasma levels of Lp(a) lipoprotein in samples obtained from subjects in the ISIS, SHEEP, and SCARF studies. The genotyping of two *LPA* SNPs (rs10455872 and rs3798220) was performed with the use of a TaqMan platform.

STATISTICAL ANALYSIS

Analysis of the association of candidate genes with coronary disease was carried out after the exclusion of SNPs with low call rates (<95%), very low frequency (minor allele frequency, <1%), or Hardy–Weinberg disequilibrium in controls ($P<1.0\times10^{-6}$). Hidden relatedness was sought with the use of identity-by-state methods, and the identity of any first-degree relatives was recorded for subsequent analysis.

We used logistic and linear regression models, which allowed for familial clustering, to perform the association analyses. The test statistics were inspected for overdispersion, and the genomic control measure was calculated.[23] Lp(a) lipoprotein levels were $\log_e$-transformed for regression analysis; study-specific standard-deviation units were used to allow for heterogeneity in measurements between studies (for additional details, see the Methods section of the Supplementary Appendix).

RESULTS

STUDY POPULATION

Samples from 3145 case subjects in the PROCARDIS cohort were genotyped with the use of the HumanCVD BeadChip. Among these samples were 2200 from subjects who had received a diagnosis of myocardial infarction (with confirmation of the diagnosis based on hospital-discharge or general-practice records for 91% of the subjects), 480 from subjects who had undergone coronary-artery revascularization (of whom 424 had received a diagno-

sis of angina), and 465 from subjects with angina only. The control group included 3352 subjects from the PROCARDIS cohort who had no personal or sibling history of coronary disease before the age of 66 years.

ASSOCIATION WITH CORONARY DISEASE

After quality-control filtering, 34,399 SNPs were used to test associations with coronary disease in samples from case subjects and control subjects. The genomic control measure λ was 1.03, indicating that the statistical modeling assumptions were appropriate. We identified 33 SNPs with an unadjusted P value of less than 1.0×10^{-6} (or $P<0.05$ after adjustment for multiple testing), which mapped to three distinct chromosomal regions (6q26–27, 9p21, and 1p13) (Table 1 in the Supplementary Appendix). The *LPA* locus on chromosome 6q26–27 showed the strongest associations with coronary artery disease, with one SNP (rs10455872) having a P value of 3.4×10^{-15}. The 25 SNPs that mapped to 9p21 are in a region that previous genomewide association studies have shown to be associated with coronary disease[1-4] and type 2 diabetes.[24,25] Similarly, the six SNPs that were localized to chromosome 1p13 are in a region that in previous genomewide association studies has been shown to be associated with coronary disease[4] and levels of low-density lipoprotein cholesterol.[26,27] The complete set of association results is available in the European Genome–Phenome Archive (www.ebi.ac.uk/ega/) under accession number EGAS00000000055.

ASSOCIATION WITH LP(A) LIPOPROTEIN LEVEL

Lp(a) lipoprotein levels were measured in samples from 1822 case subjects. The median level was 33 mg per deciliter (interquartile range, 15 to 89). There was a significant association between the Lp(a) lipoprotein level and 16 of the 27 SNPs that were studied at the *LPA* locus ($P<1.0\times10^{-6}$) (Table 1). We confirmed previous evidence of a strong association for rs3798220 ($P=5.9\times10^{-51}$),[28,29] which explained about 8% of total (both genetic and individual-specific) variation in Lp(a) lipoprotein levels. We also found an exceptionally strong association for rs10455872 ($P=3.6\times10^{-166}$), which explained about 25% of the variation. A joint analysis showed that rs3798220 and rs10455872 each conferred significant main effects that together explained 36% of variation in Lp(a) lipoprotein levels. Subjects with one or more variant alleles at either SNP had an increased geometric mean Lp(a)

Table 1. Association of *LPA* SNPs with the Plasma Level of Lp(a) Lipoprotein in 1822 Case Subjects in the PROCARDIS Cohort.*

SNP	Absolute Location	Relative Location	Risk Allele (Frequency)	Association with Increased Lp(a) Lipoprotein Level		
				Regression Coefficient for Lp(a) Level	SE for Regression Coefficient	P Value
	base pairs					
rs3127596†	160,873,025	0	G (0.30)	0.30	0.04	1.5×10^{-14}
rs7449650	160,877,104	4,079	G (0.67)	0.16	0.04	9.1×10^{-5}
rs6919346†‡	160,880,349	7,324	C (0.83)	0.43	0.05	1.6×10^{-16}
rs3798220†‡	160,881,127	8,102	C (0.02)	1.27	0.08	5.9×10^{-51}
rs11751605†‡	160,883,220	10,195	C (0.16)	0.50	0.04	5.9×10^{-28}
rs4708871	160,889,086	16,061	T (0.93)	0.53	0.11	1.0×10^{-6}
rs10755578†	160,889,728	16,703	G (0.48)	0.27	0.04	3.4×10^{-13}
rs7761293	160,890,953	17,928	A (0.47)	0.18	0.04	2.3×10^{-6}
rs9364559	160,896,138	23,113	A (0.81)	0.04	0.05	3.9×10^{-1}
rs6415084†	160,900,320	27,295	T (0.49)	0.22	0.04	2.7×10^{-9}
rs9365171†	160,901,726	28,701	C (0.65)	0.25	0.04	4.0×10^{-11}
rs6923877†	160,916,787	43,762	A (0.67)	0.26	0.04	2.4×10^{-11}
rs3798221†‡	160,918,138	45,113	G (0.81)	0.28	0.05	2.0×10^{-9}
rs7765781†‡	160,927,486	54,461	G (0.67)	0.26	0.04	2.1×10^{-11}
rs7765803†	160,927,528	54,503	G (0.67)	0.26	0.04	2.4×10^{-11}
rs10455872†‡	160,930,108	57,083	G (0.07)	1.18	0.04	3.6×10^{-166}
rs9355296	160,937,983	64,958	G (0.84)	0.01	0.05	7.8×10^{-1}
rs13202636†	160,949,718	76,693	G (0.78)	0.33	0.04	2.0×10^{-14}
rs9355813†	160,951,162	78,137	T (0.64)	0.32	0.04	2.8×10^{-17}
rs10945682†‡	160,989,931	116,906	G (0.64)	0.32	0.04	1.8×10^{-17}
rs1652507	161,002,451	129,426	T (0.84)	0.00	0.05	9.4×10^{-1}
rs1321195†	161,004,146	131,121	G (0.86)	0.26	0.05	4.0×10^{-7}
rs783149	161,008,908	135,883	C (0.84)	0.00	0.05	9.7×10^{-1}
rs1084651	161,009,807	136,782	A (0.16)	0.01	0.05	8.4×10^{-1}
rs1358754	161,010,510	137,485	G (0.87)	0.12	0.05	2.8×10^{-2}
rs1358753	161,010,560	137,535	C (0.86)	0.11	0.05	3.8×10^{-2}
rs1406888	161,011,583	138,558	T (0.53)	0.16	0.04	2.1×10^{-5}

* The association of *LPA* single-nucleotide polymorphisms (SNPs) with an increased Lp(a) lipoprotein level is shown by a regression coefficient and standard error (SE) for each SNP. Lp(a) lipoprotein levels were first log-transformed to reduce skewness and then scaled as study-specific SD units to allow for heterogeneity among the studies.

† A P value of less than 1.0×10^{-6} was considered to indicate a significant association with the Lp(a) lipoprotein level (as found in 16 SNPs).

‡ A P value of less than 4.0×10^{-9} (as calculated by stepwise regression) was considered to indicate a significant association with the Lp(a) lipoprotein level (as found in seven SNPs, which together explained 40% of the total variation).

lipoprotein level (Table 2 in the Supplementary Appendix).

Stepwise regression identified seven SNPs, including rs3798220 and rs10455872, that each had a significant association with Lp(a) lipoprotein levels ($P<4.0\times10^{-9}$) and together explained 40% of the total variation (Table 1). We previously mapped quantitative trait loci for Lp(a) lipoprotein by variance-components linkage analysis and estimated that about 74% of the total variation in Lp(a) lipoprotein levels was specific to the *LPA* locus.[30] Consequently, it appears that the SNPs that were in-

cluded in our study captured more than half the quantitative genetic variation encoded by *LPA*.

ASSOCIATION BETWEEN Lp(a) LIPOPROTEIN LEVEL AND CORONARY DISEASE

Variants at the *LPA* locus that were associated with increased effects on Lp(a) lipoprotein levels tended to be associated with increased effects on the risk of coronary disease (Fig. 1). In particular, the rs3798220 and rs10455872 SNPs, which were most strongly associated with Lp(a) lipoprotein levels, were most strongly associated with the risk of coronary disease.

REPLICATION OF ASSOCIATION

We studied independent populations to test for replication of the associations of SNPs rs3798220 and rs10455872 with both circulating Lp(a) lipoprotein levels and coronary disease. With respect to Lp(a) lipoprotein, data were available for 500 case subjects and 627 control subjects in the ISIS study and for 1151 case subjects and 1506 control subjects in the SHEEP and SCARF studies. After standardization for differences in the assay methods, there was little evidence of heterogeneity of the effects of each allele on Lp(a) lipoprotein levels in the three populations (Fig. 2, and Table 2 in the Supplementary Appendix).

With respect to coronary disease, data were available for 1259 trio families in the PROCARDIS cohort, for 2068 case subjects and 1484 control subjects in the ISIS study, and for 1519 case subjects and 1851 control subjects in the SHEEP and SCARF studies. The two SNPs were associated with coronary risk in these replication cohorts, although the association was somewhat weaker than that observed in the PROCARDIS cohort (Fig. 2). A meta-analysis of all four studies, involving nearly 8000 case subjects with coronary disease, provided an estimated odds ratio for coronary disease of 1.47 (95% confidence interval [CI], 1.35 to 1.60) for the rs10455872 SNP and 1.68 (95% CI, 1.43 to 1.98) for the rs3798220 SNP.

LPA GENOTYPE SCORE AND CORONARY DISEASE

Maximum likelihood analysis predicted that the rs3798220 and rs10455872 SNPs formed three haplotypes with frequencies (0.885 for T→A, 0.093 for T→G, and 0.023 for C→A) that were very close to the frequencies expected for alleles in linkage equilibrium (with no subjects simultaneously homozygous for the variant alleles at both SNPs). Hence,

the information on coronary risk that was encoded by rs3798220 and rs10455872 was combined into a single *LPA* genotype score with levels 0, 1, and 2 (or more) that corresponded to the sum of variant alleles (C for rs3798220 and G for rs10455872) in either of the two SNPs.

The odds ratio for coronary disease in subjects in the PROCARDIS cohort was 1.73 (95% CI, 1.51 to 1.98) with one variant allele and 4.87 (95% CI, 2.80 to 8.48) with two or more variant alleles. The score showed a strong association with coronary disease under an allele-dose risk model ($P=2.4\times10^{-21}$), with an odds ratio of 1.81 (95% CI, 1.60 to 2.04) per variant allele. The association between these *LPA* variants and the risk of coronary disease in subjects in the PROCARDIS cohort correlated with the effects of these variants on Lp(a) lipoprotein levels (Fig. 3).

We also examined the effects of the *LPA* genotype score on Lp(a) lipoprotein levels and on the risk of coronary disease in the replication cohorts (Fig. 2 in the Supplementary Appendix). There was no evidence of departure from an allele-dose risk model (P=0.11) or of heterogeneity among countries of origin (P=0.97). A meta-analysis of all the studies showed odds ratios for coronary disease of 1.51 (95% CI, 1.38 to 1.66) for one *LPA* variant allele and of 2.57 (95% CI, 1.80 to 3.67) for two or more alleles.

INDEPENDENCE OF KNOWN RISK FACTORS

In contrast with previous studies that reported a positive association between the Lp(a) lipoprotein level and coronary disease only in patients with elevated levels of low-density lipoprotein (LDL) cholesterol,[31] there was no significant heterogeneity of the odds ratio for coronary disease among subgroups of patients with various LDL cholesterol levels (Fig. 3 in the Supplementary Appendix). There was also no heterogeneity in the effects of the *LPA* genotype score on the risk of coronary disease according to the level of high-density lipoprotein (HDL) cholesterol, age, sex, presence or absence of a history of myocardial infarction, body-mass index, presence or absence of diabetes mellitus or hypertension, and smoking status. The two SNPs alone or combined in the genotype score showed little consistent evidence of an association with blood levels of lipids or inflammatory markers in a subgroup of control subjects in the PROCARDIS cohort who were not taking a statin drug (Table 3 in the Supplementary Appendix).
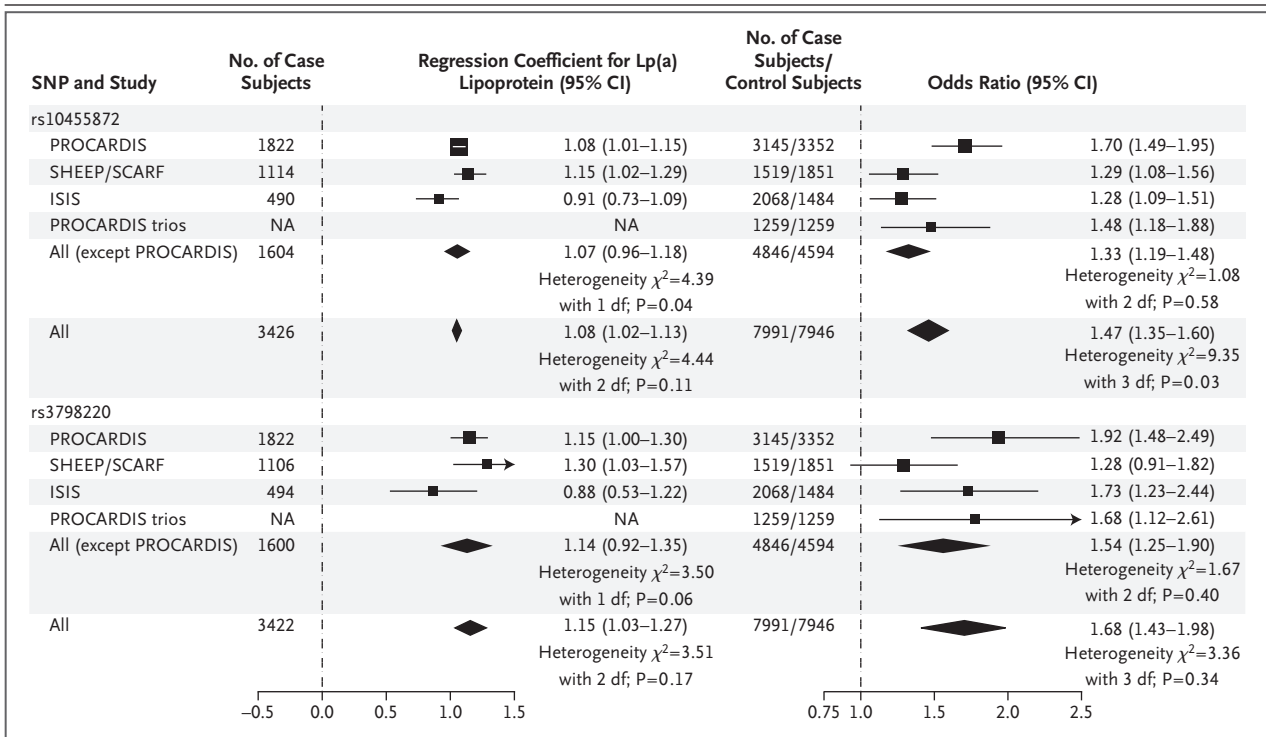
Figure 2. Replication of the Association of Single-Nucleotide Polymorphisms (SNPs) in *LPA* with Lp(a) Lipoprotein Levels and the Risk of Coronary Disease in Independent Populations, as Compared with the PROCARDIS Cohort.

The results of replication testing are shown for the associations between two variants that were strongly linked with an increased level of Lp(a) lipoprotein in the PROCARDIS cohort — rs10455872 and rs3798220 — in three independent populations: case subjects with non-fatal myocardial infarction and control subjects from two studies of persons living in the greater Stockholm area, the Stockholm Heart Epidemiology Program (SHEEP) and the Stockholm Coronary Artery Risk Factor (SCARF) study; case subjects with nonfatal myocardial infarction and control subjects from the International Study of Infarct Survival (ISIS); and PROCARDIS trio families (families either with a proband and two parents or with a proband, a parent, and at least one sibling). The values are regression coefficients and 95% confidence intervals (CIs) for Lp(a) lipoprotein levels. Levels were first log-transformed to reduce skewness and then scaled as study-specific SD units to allow for heterogeneity among the studies. The odds ratios (squares, with the size inversely proportional to the sampling variance) are for the associations of the two SNPs with the risk of coronary disease in the various studies. The combined odds ratios and 95% confidence intervals are indicated by the diamonds. The associations with Lp(a) lipoprotein levels have been adjusted for country, age, and sex. For the heterogeneity data, df denotes degree of freedom. NA denotes missing data for trio families.

ASSOCIATION WITH APOLIPOPROTEIN(a) SIZE

The rare alleles of both the rs10455872 and rs3798220 SNPs were each correlated with a smaller apolipoprotein(a) isoform (as measured by Western blotting) and a lower copy number (as measured with a quantitative PCR assay) (Fig. 4). There was very little variation in the number of kringle IV–type 2 repeats for chromosomes carrying the variant rs10455872 allele, which suggested that this SNP tags a clade of short isoform alleles with 17 to 20 repeats (Fig. 4A). There was a similar pattern for chromosomes carrying the variant rs3798220 allele, which suggested that this SNP tags a clade of isoform alleles with 19 to 21 repeats.

CORONARY DISEASE ASSOCIATION ADJUSTED FOR Lp(a) LIPOPROTEIN LEVEL

We also measured levels of Lp(a) lipoprotein in a random subgroup of 1578 case subjects and 1726 control subjects in the PROCARDIS cohort, using a Siemens autoanalyzer. In the case subjects, the correlation between the Lp(a) lipoprotein level and the result of the Immuno immunoturbidimetric assay was 0.94. The median Lp(a) lipoprotein level was 15.1 mg per deciliter (interquartile range, 7.2 to 42.9) in case subjects and 10.3 mg per deciliter (interquartile range, 5.2 to 24.7) in control subjects (P<0.001). In a meta-analysis of 3137 subjects with coronary disease for whom Lp(a) lipo-

protein levels were available, the *LPA* genotype score was strongly associated with the risk of coronary disease (odds ratio, 1.52; 95% CI, 1.36 to 1.72) (Fig. 4 in the Supplementary Appendix). After adjustment for the Lp(a) lipoprotein level, the association between the *LPA* genotype score and the risk of coronary disease was abolished (odds ratio, 1.02; 95% CI, 0.88 to 1.18).

DISCUSSION

In this study, we identified risk loci for coronary disease by using a novel gene chip consisting of 48,742 SNPs for 2100 candidate genes that were selected for their potential relevance to coronary disease. With this gene chip, we confirmed the previous identification of three chromosomal regions that were correlated with the risk of coronary disease: 6q26–27, 9p21, and 1p13. Since the 6q26–27 region includes the *LPA* gene, we then used comprehensive SNP typing to characterize the spectrum of variation at the *LPA* locus and showed the independent relevance of several variants at the *LPA* locus for both the Lp(a) lipoprotein level and the risk of coronary disease.

Two common variants at rs10455872 and rs3798220 together explained 36% of the total variation in the Lp(a) lipoprotein level and were independently associated with an increased risk of coronary disease. We also found that the effects of the *LPA* variants on the risk of coronary disease correlated with the effects on the Lp(a) lipoprotein level. The linear dose–response relationship of the *LPA* variants with both the Lp(a) lipoprotein level and the risk of coronary disease provided compelling support for a causal role of an elevated plasma level of Lp(a) lipoprotein in the risk of coronary disease.

The rs3798220 SNP encodes a nonsynonymous variant in *LPA*. The frequency of this high-risk variant is about 2%. This SNP has previously been reported to have a strong association with the Lp(a) lipoprotein level,[28,29] a moderate association with the LDL cholesterol level,[26] and a tentative association with the risk of coronary disease.[28,29] Thus, our study, which involved nearly 8000 case subjects with coronary disease, provides a more definitive estimate of the strength of the association between rs3798220 and the risk of coronary disease. Furthermore, a haplotype analysis of Brit-
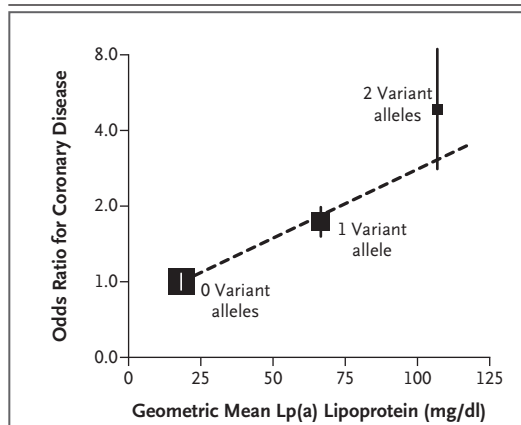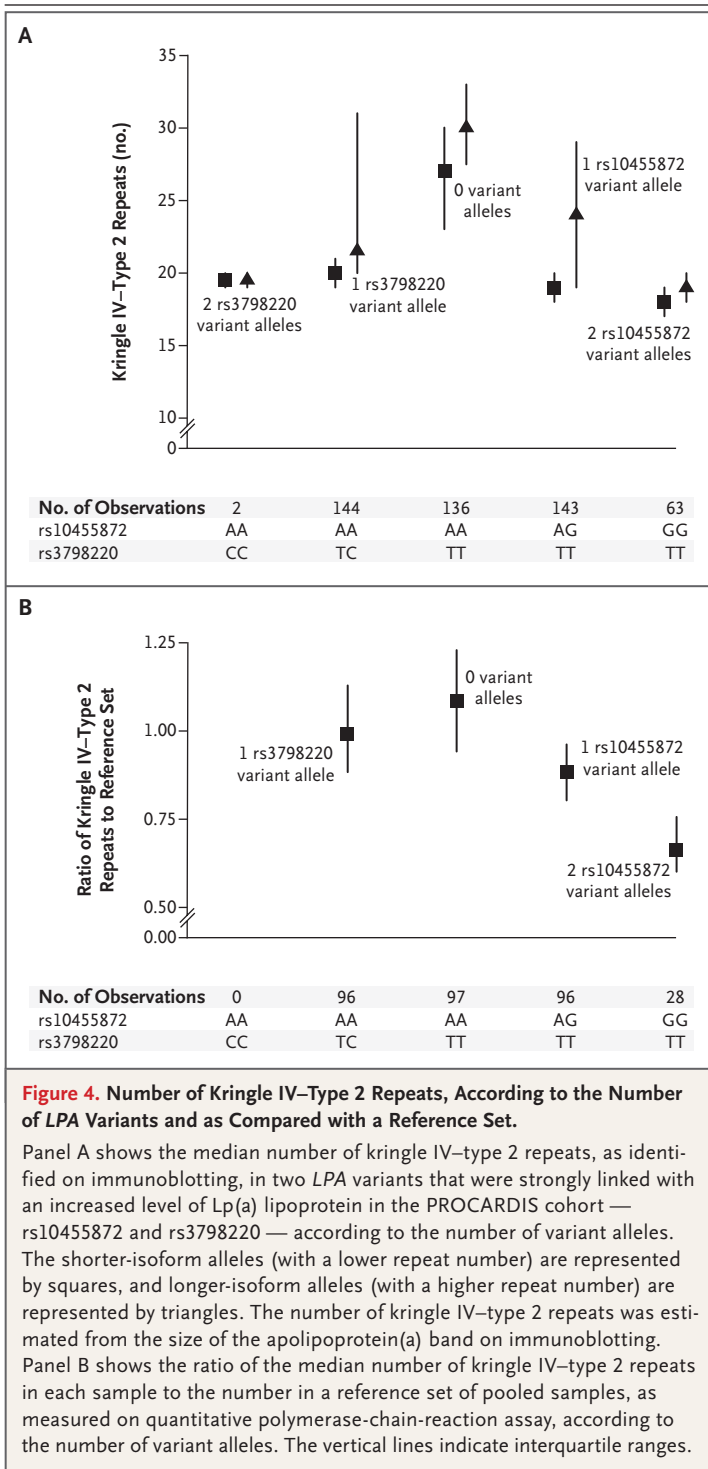

Figure 3. Association of the *LPA* Genotype Score with the Lp(a) Lipoprotein Level and the Risk of Coronary Disease in the PROCARDIS Cohort.

The odds ratios (squares, with the size inversely proportional to the sampling variation) are for the association of the *LPA* genotype score (no variant alleles, one variant allele, or two variant alleles) with the risk of coronary disease, as measured with the use of "floating absolute risks" which summarize the sampling variation for the three genotype scores without the selection of an arbitrary baseline genotype score. The vertical lines indicate 95% confidence intervals.

ish control genotype data (available at www.wtccc .org.uk) showed that rs3798220 had a strong correlation ($r^2 = 0.86$) with a four-SNP haplotype (CCTC; approximate frequency, 2%), which would explain and refine part of the recently reported association between the *SLC22A3–LPAL2–LPA* gene cluster and coronary disease.[7]

The rs10455872 SNP maps to intron 25 in the *LPA* gene. The allele frequency of the high-risk variant is about 7%. Previous studies of Lp(a) lipoprotein levels and the risk of coronary disease did not measure rs10455872 (which was poorly tagged in genomewide SNP arrays) and thus did not identify the strong associations with this SNP.

The SNPs that were included in our study appeared to capture more than half the genetic variation encoded in the *LPA* locus in Lp(a) lipoprotein levels, a higher rate than that previously reported for a pentanucleotide variant in the promoter[32,33] or for other SNPs identified by genomewide association studies.[34] Variation in isoform size is believed to influence plasma levels through a direct mechanism involving protein secretory

Figure 4. Number of Kringle IV–Type 2 Repeats, According to the Number of *LPA* Variants and as Compared with a Reference Set.

Panel A shows the median number of kringle IV–type 2 repeats, as identified on immunoblotting, in two *LPA* variants that were strongly linked with an increased level of Lp(a) lipoprotein in the PROCARDIS cohort — rs10455872 and rs3798220 — according to the number of variant alleles. The shorter-isoform alleles (with a lower repeat number) are represented by squares, and longer-isoform alleles (with a higher repeat number) are represented by triangles. The number of kringle IV–type 2 repeats was estimated from the size of the apolipoprotein(a) band on immunoblotting. Panel B shows the ratio of the median number of kringle IV–type 2 repeats in each sample to the number in a reference set of pooled samples, as measured on quantitative polymerase-chain-reaction assay, according to the number of variant alleles. The vertical lines indicate interquartile ranges.

processes.[35,36] Both the rs10455872 and rs3798220 variants were inversely correlated with kringle IV–type 2 repeats and appear to tag clades of short

isoform alleles. Thus, the SNP associations are strongly dependent on kringle-repeat polymorphism.

The mechanism by which an increased level of Lp(a) lipoprotein increases the risk of coronary disease is less well understood; it may involve LDL lipoprotein cholesterol,[26] the inhibition of conversion of plasminogen to plasmin,[37] the inhibition of the expression of tissue factor,[38] or the carriage of proinflammatory oxidized phospholipids.[39] We found no significant association between either rs10455872 or rs3798220 and the plasma level of apolipoprotein B, fibrinogen, or C-reactive protein.

Our results extend the findings of a recent Danish study of kringle IV–type 2 repeats and the risk of coronary disease.[40] In that study, the kringle IV–type 2 genetic variation, as estimated with a quantitative PCR assay (which averaged the two alleles), explained 22% of the variation in the Lp(a) lipoprotein level; in 599 case subjects with coronary disease, the association between kringle repeats and disease was partially attenuated by adjustment for the Lp(a) lipoprotein level. By contrast, our two-SNP *LPA* genotype score explained 36% of the variation in the Lp(a) lipoprotein level, and the association between *LPA* and the risk of coronary disease was abolished after adjustment for the Lp(a) lipoprotein level in a meta-analysis of 3137 subjects with coronary disease — findings that are consistent with a causal role of an increased Lp(a) lipoprotein level in coronary disease.

In conclusion, we have identified two common SNPs in *LPA* that correlate with both the Lp(a) lipoprotein level and the risk of coronary disease. These SNPs explain 36% of the variation in the Lp(a) lipoprotein level. One in six persons carries a variant *LPA* allele and thus has a risk of coronary disease that is increased by a factor of 1.5.

APPENDIX

The authors' affiliations are as follows: the Clinical Trial Service Unit and Epidemiological Studies Unit (R. Clarke, J.C.H., S.P., D.B., R. Collins) and the Department of Cardiovascular Medicine and the Wellcome Trust Centre for Human Genetics (J.F.P., T.K., A.G., H.W., M.F.), University of Oxford, Oxford; and the Division of Biochemical Sciences, Faculty of Health and Medical Sciences, University of Surrey, Surrey (F.R.G.) — all in the United Kingdom; Centre National de Genotypage, Institut Genomique, Commissariat à l'Énergie Atomique, Evry, France (S.C.H., M.L.); the Department of Cardiovascular Research, Istituto di Ricerche Farmacologiche Mario Negri, Milan (S.B., M.G.F.); Leibniz-Institut für Arterioscleroseforschung an der Universität Münster, Münster, Germany (S.R., U.S.); and the Atherosclerosis Research Unit, Department of Medicine (A.S., A.M., A.H.), and the Division of Cardiovascular Epidemiology, Institute of Environmental Medicine (B.G., K.L., U.F.) — both at Karolinska Institutet, Stockholm.

REFERENCES

1. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007; 447:661-78.

2. McPherson R, Pertsemlidis A, Kavaslar N, et al. A common allele on chromosome 9 associated with coronary heart disease. Science 2007;316:1488-91.

3. Helgadottir A, Thorleifsson G, Manolescu A, et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. Science 2007;316: 1491-3.

4. Samani NJ, Erdmann J, Hall AS, et al. Genomewide association analysis of coronary artery disease. N Engl J Med 2007; 357:443-53.

5. Watkins H, Farrall M. Genetic susceptibility to coronary artery disease: from promise to progress. Nat Rev Genet 2006;7: 163-73.

6. Hamsten A, Eriksson P. Identifying the susceptibility genes for coronary artery disease: from hyperbole through doubt to cautious optimism. J Intern Med 2008;263:538-52.

7. Trégouët DA, König IR, Erdmann J, et al. Genome-wide haplotype association study identifies the SLC22A3-LPAL2-LPA gene cluster as a risk locus for coronary artery disease. Nat Genet 2009;41:283-5.

8. Berg K, Dahlén G, Børresen AL. Lp(a) phenotypes, other lipoprotein parameters, and a family history of coronary heart disease in middle-aged males. Clin Genet 1979;16:347-52.

9. The Emerging Risk Factors Collaboration. Lipoprotein(a) concentration and the risk of coronary heart disease, stroke and nonvascular mortality. JAMA 2009;302: 412-23.

10. Kronenberg F, Kronenberg MF, Kiechl S, et al. Role of lipoprotein(a) and apolipoprotein(a) phenotype in atherogenesis: prospective results from the Bruneck study. Circulation 1999;100:1154-60.

11. Smolders B, Lemmens R, Thijs V. Lipoprotein(a) and stroke: a meta-analysis of observational studies. Stroke 2007;38: 1959-66.

12. Berglund L, Ramakrishnan R. Lipoprotein(a): an elusive cardiovascular risk factor. Arterioscler Thromb Vasc Biol 2004;24:2219-26.

13. Keating BJ, Tischfield S, Murray SS, et al. Concept, design and implementation of a cardiovascular gene-centric 50 k SNP array for large-scale genomic association studies. PLoS One 2008;3(10):e3583.

14. Farrall M, Green FR, Peden JF, et al. Genome-wide mapping of susceptibility to coronary artery disease identifies a novel replicated locus on chromosome 17. PLoS Genet 2006;2(5):e72.

15. Broadbent HM, Peden JF, Lorkowski S, et al. Susceptibility to coronary artery disease and diabetes is encoded by distinct, tightly linked SNPs in the ANRIL locus on chromosome 9p. Hum Mol Genet 2008;17:806-14.

16. Siekmeier R, Marz W, Scharnagel H, et al. Determination of lipoprotein(a): comparison of a novel latex enhanced immunoturbidimetric assay and an immunoradiometric assay. Lab Medizin 1996;20: 294-8. (In German.)

17. Marcovina SM, Hobbs HH, Albers JJ. Relation between number of apolipoprotein(a) kringle 4 repeats and mobility of isoforms in agarose gel: basis for a standardized isoform nomenclature. Clin Chem 1996;42:436-9.

18. Lanktree M, Rajakumar C, Brunt JH, Koschinsky ML, Connelly PW, Hegele RA. Determination of lipoprotein(a) kringle repeat number from genomic DNA: copy number variation genotyping using qPCR. J Lipid Res 2009;50:768-72.

19. PROCARDIS Consortium. A trio family showing association of the lymphotoxin-alpha N26 (804A) allele with coronary artery disease. Eur J Hum Genet 2004; 12:770-4.

20. Clarke R, Xu P, Bennett D, et al. Lymphotoxin-alpha gene and risk of myocardial infarction in 6,928 cases and 2,712 controls in the ISIS case-control study. PLoS Genet 2006;2(7):e107.

21. Reuterwall C, Hallqvist J, Ahlbom A, et al. Higher relative, but lower absolute risks of myocardial infarction in women than in men: analysis of some major risk factors in the SHEEP study. J Intern Med 1999;246:161-74.

22. Samnegård A, Silveira A, Lundman P, et al. Serum matrix metalloproteinase-3 concentration is influenced by MMP-3 -1612 5A/6A promoter genotype and associated with myocardial infarction. J Intern Med 2005;258:411-9.

23. Devlin B, Roeder K. Genomic control for association studies. Biometrics 1999; 55:997-1004.

24. Scott LJ, Mohlke KL, Bonnycastle LL, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 2007;316: 1341-5.

25. Zeggini E, Weedon MN, Lindgren CM, et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science 2007; 316:1336-41.

26. Kathiresan S, Willer CJ, Peloso GM, et al. Common variants at 30 loci contribute to polygenic dyslipidemia. Nat Genet 2009; 41:56-65.

27. Willer CJ, Sanna S, Jackson AU, et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. Nat Genet 2008;40:161-9.

28. Luke MM, Kane JP, Liu DM, et al. A polymorphism in the protease-like domain of apolipoprotein(a) is associated with severe coronary artery disease. Arterioscler Thromb Vasc Biol 2007;27:2030-6.

29. Chasman DI, Shiffman D, Zee RY, et al. Polymorphism in the apolipoprotein(a) gene, plasma lipoprotein(a), cardiovascular disease, and low-dose aspirin therapy. Atherosclerosis 2009;203:371-6.

30. Barlera S, Specchia C, Farrall M, et al. Multiple QTL influence the serum Lp(a) concentration: a genome-wide linkage screen in the PROCARDIS study. Eur J Hum Genet 2007;15:221-7.

31. Suk Danik J, Rifai N, Buring JE, Ridker PM. Lipoprotein(a), measured with an assay independent of apolipoprotein(a) isoform size, and risk of future cardiovascular events among initially healthy women. JAMA 2006;296:1363-70.

32. Røsby O, Berg K. LPA gene: interaction between the apolipoprotein(a) size ('kringle IV' repeat) polymorphism and a pentanucleotide repeat polymorphism influences Lp(a) lipoprotein level. J Intern Med 2000;247:139-52.

33. Trommsdorff M, Köchl S, Lingenhel A, et al. A pentanucleotide repeat polymorphism in the 5′ control region of the apolipoprotein(a) gene is associated with lipoprotein(a) plasma concentrations in Caucasians. J Clin Invest 1995;96:150-7.

34. Ober C, Nord AS, Thompson EE, et al. Genome-wide association study of plasma lipoprotein (a) levels identifies multiple genes on chromosome 6q. J Lipid Res 2009; 50:798-806.

35. White AL, Hixson JE, Rainwater DL, Lanford RE. Molecular basis for "null" lipoprotein(a) phenotypes and the influence of apolipoprotein(a) size on plasma lipoprotein(a) level in the baboon. J Biol Chem 1994;269:9060-6.

36. Wang J, Boedeker H, Hobbs H, White AL. Determinants of human apolipoprotein[a] secretion from mouse hepatocyte cultures. J Lipid Res 2001;42:60-9.

37. Grainger DJ, Kemp PR, Liu AC, Lawn RM, Metcalfe JC. Activation of transforming growth factor-beta is inhibited in transgenic apolipoprotein(a) mice. Nature 1994;370:460-2.

38. Caplice NM, Panetta C, Peterson TE, et al. Lipoprotein(a) binds and inactivates tissue factor pathway inhibitor: a novel link between lipoproteins and thrombosis. Blood 2001;98:2980-7.

39. Tsimikas S, Brilakis ES, Miller ER, et al. Oxidized phospholipids, Lp(a) lipoprotein, and coronary artery disease. N Engl J Med 2005;353:46-57.

40. Kamstrup PR, Tybjaerg-Hansen A, Steffensen R, Nordestgaard BG. Genetically elevated lipoprotein(a) and increased risk of myocardial infarction. JAMA 2009; 301:2331-9.

Copyright © 2009 Massachusetts Medical Society.

FULL TEXT OF ALL *JOURNAL* ARTICLES ON THE WORLD WIDE WEB

Access to the complete contents of the *Journal* on the Internet is free to all subscribers. To use this Web site, subscribers should go to the *Journal*'s home page (**NEJM.org**) and register by entering their names and subscriber numbers as they appear on their mailing labels. After this one-time registration, subscribers can use their passwords to log on for electronic access to the entire *Journal* from any computer that is connected to the Internet. Features include a library of all issues since January 1993 and abstracts since January 1975, a full-text search capacity, and a personal archive for saving articles and search results of interest. All articles can be printed in a format that is virtually identical to that of the typeset pages. Beginning 6 months after publication, the full text of all Original Articles and Special Articles is available free to nonsubscribers.