Common and Rare Alleles in Apolipoprotein B Contribute to Plasma Levels of Low-Density Lipoprotein Cholesterol in the General Population

Marianne Benn, Maria C. A. Stene, Børge G. Nordestgaard, Gorm B. Jensen, Rolf Steffensen, and Anne Tybjærg-Hansen

Department of Clinical Biochemistry (M.B., M.C.A.S., A.T.-H.), Rigshospitalet, Copenhagen University Hospital; Department of Clinical Biochemistry (B.G.N.), Herlev University Hospital, The Copenhagen City Heart Study (B.G.N., G.B.J., A.T.-H.), Bispebjerg University Hospital, and Department of Medicine B (R.S.), Hillerød Hospital, all Faculty of Health Sciences, University of Copenhagen, DK-2100 Copenhagen, Denmark

Context: We have previously shown that rare mutations in the apolipoprotein B gene (*APOB*) may result in not only severe hypercholesterolemia and ischemic heart disease but also hypocholesterolemia. Despite this, common single-nucleotide polymorphisms (SNPs) in *APOB* have not convincingly been demonstrated to affect low-density lipoprotein (LDL) cholesterol levels.

Objective: We tested the hypothesis that nonsynonymous SNPs in three important functional domains of *APOB* and *APOB* tag SNPs predict levels of LDL cholesterol and apolipoprotein B and risk of ischemic heart disease.

Design: This was a prospective study with 25 yr 100% follow up, The Copenhagen City Heart Study.

Setting: The study was conducted in the Danish general population.

Participants: Participants included 9185 women and men aged 20–80+ yr.

Main Outcome Measures: Levels of LDL cholesterol and apolipoprotein B and risk of ischemic heart disease and myocardial infarction were measured. The hypothesis was formulated before genotyping.

Results: We genotyped 9185 individuals for *APOB* T71I (minor allele frequency: 0.33), Ivs4+171c>a (0.14), A591V (0.47), Ivs18+379a>c (0.30), Ivs18+1708g>t (0.45), T2488Tc>t (0.48), P2712L (0.21), R3611Q (0.09), E4154K (0.17), and N4311S (0.21). SNPs were associated with increases (T71I, Ivs181708g>t, T2488Tc>t, R3611) or decreases (Ivs4+171c>a, A591V, Ivs18+379a>c, P2712L, E4154, N4311S) in LDL cholesterol from −4.7 to +8.2% (−0.28 to 0.30 mmol/liter; $P \leq 0.002$), and corresponding effects on cholesterol and apolipoprotein B levels. However, as predicted from the magnitude of the observed LDL cholesterol effects, none of these SNPs predicted risk of ischemic heart disease prospectively in the general population, in a case-control study, or as haplotypes.

Conclusions: Multiple common and rare alleles in *APOB* contribute to plasma levels of LDL cholesterol in the general population, although the effects of common alleles and haplotypes are modest. **(*J Clin Endocrinol Metab* 93: 1038–1045, 2008)**

Twin studies suggest that about 50–60% of the variation in plasma levels of apolipoprotein B is genetically determined (1). Apolipoprotein B is crucial in the initial steps of chylomicron and very low-density lipoprotein (VLDL) formation, as well as in the binding and clearance of low-density lipoprotein (LDL) by the LDL receptor. Rare missense mutations in the apolipoprotein

0021-972X/08/$15.00/0

Printed in U.S.A.

doi: 10.1210/jc.2007-1365 Received June 19, 2007. Accepted December 17, 2007.

First Published Online December 26, 2007

Abbreviations: *APOB*, Apolipoprotein B gene; LDL, low-density lipoprotein; SNP, single-nucleotide polymorphism; VLDL, very low-density lipoprotein.

B gene (*APOB*) may result in not only severe hypercholesterolemia and increased risk of ischemic heart disease but also hypocholesterolemia (2–4). In contrast, previous studies examining the association of common nonsynonymous single-nucleotide polymorphisms (SNPs) in *APOB* with lipid and lipoprotein levels and with risk of ischemic heart disease have been conflicting (5, 6).

For this reason, we selected six nonsynonymous SNPs in *APOB*, located in important functional domains crucial for lipidation of the nascent apolipoprotein B (T71I, A591V) (7, 8), involved in structural changes of apolipoprotein B during the conversion of VLDL to LDL (P2712L) (9) or known or suspected of regulating binding to the LDL receptor (R3611Q, E4154K, N4311S) (10, 11). In addition, we selected four other SNPs [Ivs4 + 171c>a, Ivs18 + 379a>c, Ivs18 + 1708 g>t, and T2488Tc>t (12)] because they together with T71I, A591V, and E4154K are predicted by HapMap to tag the genetic variation in the entire coding and intronic regions of *APOB*, comprising approximately 43 kb of genomic DNA.

We genotyped 9185 individuals from the Danish general population followed up prospectively for 25 yr in the Copenhagen City Heart Study and tested the following hypotheses: T71I, Ivs4 + 171c>a, A591V, Ivs18 + 379a>c, Ivs18 + 1708 g>t, T2488Tc>t, P2712L, R3611Q, E4154K, and N4311S in *APOB* predict levels of LDL cholesterol and apolipoprotein B and risk of ischemic heart disease. Results for risk of ischemic heart disease were verified in an independent case-control study comprising 944 cases and 7664 controls.

Subjects and Methods

Subjects

General population sample

The Copenhagen City Heart Study is a prospective cardiovascular study of the Danish general population initiated in 1976–1978 with follow-up examinations in 1981–1983, 1991–1994, and 2001–2003 (13, 14). Individuals were selected based on the national Central Population Register code to reflect the adult Danish population aged 20–80+ yr. Blood samples for DNA extraction were available on 9259 participants; of these 9185 were genotyped for all ten SNPs in *APOB* (T71I, Ivs4 + 171c>a, A591V, Ivs18 + 379a>c, Ivs18 + 1708 g>t, T2488Tc>t, P2712L, R3611Q, E4154K, and N4311S).

Information on diagnosis of ischemic heart disease (World Health Organization; *International Classification of Diseases*, 8th edition: codes 410–414; 10th edition: I20-I25) was collected and verified until the beginning of 2004 by reviewing all hospital admissions and diagnoses entered in the national Danish Patient Registry and all causes of death entered in the national Danish Causes of Death Registry. Ischemic heart disease was myocardial infarction (codes 410 and I21) or characteristic symptoms of angina pectoris (codes 411 and I20) (15). A diagnosis of myocardial infarction required the presence of at least two of the following criteria: characteristic chest pain, elevated cardiac enzymes, and electrocardiographic changes indicative of myocardial infarction.

Patients with ischemic heart disease

A second cohort comprised 944 patients from the greater Copenhagen area referred for coronary angiography to Copenhagen University Hospital during the period 1991 through 1994. These patients had documented ischemic heart disease based on characteristic symptoms of stable angina pectoris (15), plus at least one of the following: severe stenosis on coronary angiography, a previous myocardial infarction, or a positive exercise electrocardiography test. The diagnosis of myocardial infarction was established with the same criteria as in the general population sample.

Study designs

Studies were approved by institutional review boards and Danish ethical committees [no. (KF)V.100.2039/91, Copenhagen and Frederiksberg committee, and no. KA93125, Copenhagen County committee] and conducted according to the Declaration of Helsinki. Informed consent was obtained from participants. More than 99% were white and of Danish descent.

Prospective study of risk of ischemic heart disease

We included 9185 participants from The Copenhagen City Heart Study. All end points were recorded in the follow-up period 1976–2004. The median follow-up time was 25 yr (186,985 person-years). Individuals diagnosed with ischemic heart disease before entry were excluded (n = 61). We observed the following incident events: ischemic heart disease 1460, and myocardial infarction 729 (Table 1).

Case-control study of risk of ischemic heart disease

We included 944 cases with ischemic heart disease and 7664 unmatched controls from The Copenhagen City Heart Study without ischemic heart disease.

TABLE 1. Characteristics of individuals in the prospective study

	Controls (n = 7664)	Participants with ischemic heart disease (n = 1460)	Participants with myocardial infarction (n = 729)
Age (yr)	56 ± 15	67 ± 10[a]	67 ± 10[a]
Total cholesterol (mmol/liter)	6.1 ± 1.3	6.5 ± 1.3[a]	6.6 ± 1.3[a]
LDL cholesterol (mmol/liter)	3.7 ± 1.1	4.1 ± 1.2[a]	4.2 ± 1.1[a]
Apolipoprotein B (mg/dl)	85 ± 23	95 ± 23[a]	97 ± 23[a]
HDL cholesterol (mmol/liter)	1.6 ± 0.5	1.4 ± 0.5[a]	1.4 ± 0.5[a]
Triglycerides (mmol/liter)	1.8 ± 1.5	2.2 ± 1.7[a]	2.3 ± 1.6[a]
Body mass index (kg/m^2)	25 ± 4.3	27 ± 4.5[a]	27 ± 4.3[a]
Hypertension (%)	50	74[a]	76[a]
Diabetes (%)	4	9[a]	10[a]
Smokers (%)	49	51	54[a]

Values are means ± SD or percentages. Incident cases with ischemic heart disease or myocardial infarction were compared with controls without disease by Mann-Whitney *U* test or Pearson χ^2 test.

[a] $P < 0.001$.

Laboratory analyses

SNP genotyping

Genotyping was by TaqMan chemistry using an ABI Prism 7900HT sequence detection system (Applied Biosystems Inc., Foster City, CA) for T71I (rs1367117), Ivs4 + 171c>a (rs531819), A591V (rs679899), Ivs18 + 379a>c (rs10199768), and Ivs18 + 1708 g>t (rs3791980) and by PCR followed by digestion with *Xba*I (T2488Tc>t; rs693), *Bfa*I (P2712L; rs676210), *Msp*I (R3611Q; rs1801701), *Eco*RI (E4154K; rs1042031), or *Eco*57I (N4311S; rs1042034), respectively. Primers, TaqMan probes, and PCR conditions are available from the authors.

Biochemical analyses

Colorimetric and turbidimetric assays were used to measure plasma levels of total cholesterol, apolipoprotein B, HDL cholesterol, and triglycerides. LDL cholesterol was calculated using the Friedewald equation (16), and non-HDL cholesterol was total cholesterol-HDL cholesterol.

Other covariates

The risk factors, diabetes mellitus, smoking, and hypertension were dichotomized and defined as ever-diabetics (self-reported disease, use of antidiabetic medication and/or a nonfasting plasma glucose > 11.0 mmol/liter), ever-smokers (ex-smoker or current smoker), ever-hypertensives (systolic blood pressure ≥ 140 mm Hg or diastolic blood pressure ≥ 90 mm Hg and/or use of antihypertensive medication). Body mass index was weight (kilograms) divided by height squared (square meters).

Statistical analysis

Data were analyzed using Stata/SE 9.2 (Stata Corp., College Station, TX). Two-sided probability values less than 0.05 were considered significant. Pairwise linkage disequilibrium was estimated using Haploview (http://www.broad.mit.edu/mpg/haploview/download/php). Mann-Whitney *U* test and Pearson's χ^2 test were used in two-group comparisons. The effect of SNP genotype on levels of cholesterol, LDL cholesterol, apolipoprotein B, and non-HDL cholesterol was determined by ANOVA and Student's *t* test.

In the prospective study, with the use of left truncation (delayed entry), Cox proportional hazards regression models with age as time scale estimated hazard ratios. Multifactorial adjustment was for age, gender, total cholesterol, LDL cholesterol, HDL cholesterol, triglycerides, body mass index, hypertension, diabetes, smoking, and menopausal status and use of hormonal replacement therapy for women. Bivariate tests of interaction between the covariates mentioned above and SNP genotypes on lipid, lipoprotein, and apolipoprotein levels and risk of ischemic heart disease and myocardial infarction were all nonsignificant. In the case-control study, logistic regression analysis was used to estimate odds ratios.

Estimated haplotypes containing the seven SNPs that according to HapMap tag the genetic variation in the coding and intronic regions of *APOB* (T71I, Ivs4 + 171c>a, A591V, Ivs18 + 379a>c, Ivs18 + 1708 g>t, T2488Tc>t, and E4154K) were inferred using the freely available WHAP software (17), as were levels of total cholesterol, LDL cholesterol, and apolipoprotein B and risk of ischemic heart disease and myocardial infarctions as a function of haplotypes.

Evolutionary conservation of nonsynonymous sequence variations in *APOB*

To compare evolutionary conservation with functional effects of the six nonsynonymous SNPs, we aligned the human *APOB* amino acid sequence with the orthologous sequences from other mammals, birds, fish, and sea urchin. For comparison, we included three *APOB* mutations (R3480P, R3500Q, R3531C) with known functional effects on LDL metabolism (2, 4). Next we examined the ability of three computer-based algorithms, PolyPhen (18), SIFT (19), and PANTHER (20), to predict the functional effects of the nonsynonymous SNPs and mutations in *APOB*. All three programs use sequence similarity to predict whether an amino acid substitution affects protein function, and PolyPhen in addition uses structural information.

Results

Location of the 10 SNPs relative to the amino acid sequence and structural and functional domains of apolipoprotein B are shown in Fig. 1. Minor allele frequencies were T71I: 0.33, Ivs4 + 171c>a: 0.14, A591V: 0.47, Ivs18 + 379a>c: 0.30, Ivs18 + 1708 g>t: 0.45, T2488Tc>t: 0.48, P2712L: 0.21, R3611Q: 0.09, E4154K: 0.17, and N4311S: 0.21 (Fig. 2). All genotype distributions were in Hardy-Weinberg equilibrium.

Linkage disequilibrium

Linkage disequilibrium as r^2 and D′ is shown for all 10 SNPs in Fig. 2. Generally, a high degree of linkage disequilibrium was present throughout the gene, especially between the five most C-terminal SNPs, indicating that these SNPs are on the same haplotypes. However, only the minor alleles of P2712L and N4311S were also highly correlated (r^2 = 1.0, all other r^2s < 0.70) and could therefore tag or serve as proxy for the other SNP.

Nevertheless, according to HapMap, T71I, Ivs4 + 171c>a, A591V, Ivs18 + 379a>c, Ivs18 + 1708 g>t, T2488Tc>t, and E4154K are seven tag SNPs covering the entire *APOB* gene of approximately 43 kb (http://www.hapmap.org/index.html.en): T71I and Ivs4 + 171c>a form a haploblock covering the N-terminal part of the gene (12,986 nucleotides) and Ivs18 + 1708 g>t, T2488Tc>t, E4154K another block covering the most C-terminal end of the gene (18,719 nucleotides), whereas A591V and Ivs18 + 379a>c cover two haploblocks of, respectively, 5,585 and 1,329 nucleotides in between. It follows that Ivs18 + 1708 g>t, T2488Tc>t, and E4154K according to HapMap

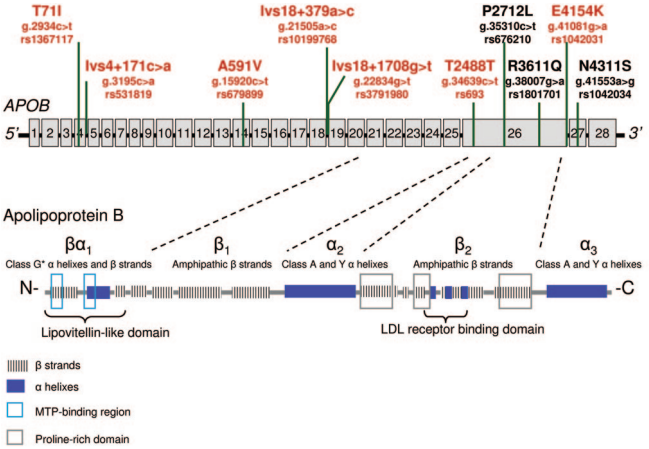

FIG. 1. Location of the 10 SNPs relative to the amino acid sequence and the structural and functional domains of apolipoprotein B. T71V (rs1367117) and A591V(rs679899) are located in domains crucial for lipidation of the nascent apolipoprotein B (7, 8); P2712(rs676210) is in a domain involved in structural changes of apolipoprotein B during the conversion of VLDL to LDL (9); and R3611Q (rs1801701), E4154K (rs1042031), and N4311S (rs1042034) are in domains known to or suspected of regulating binding to the LDL receptor (11). The seven SNPs predicted by HapMap (http://www.hapmap.org/index.html.en) to tag for the genetic variation in the entire *APOB* gene (coding regions and introns) are marked in red. MTP, Microsomal triglyceride transfer protein.

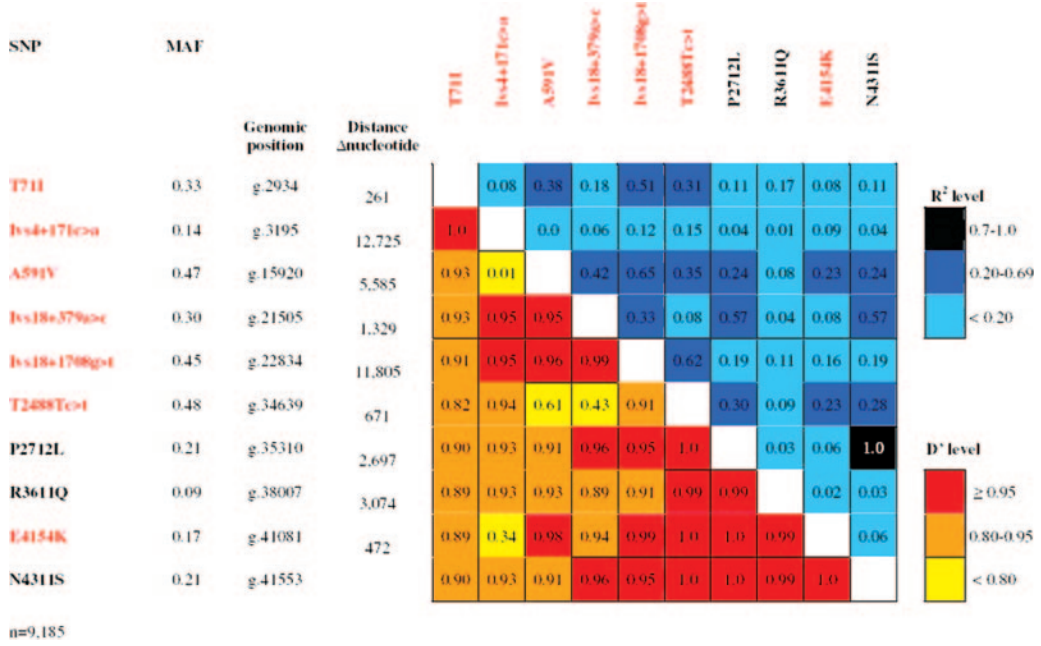

FIG. 2. Pairwise linkage disequilibrium between the 10 SNPs examined in the present study. The seven SNPs predicted by HapMap (http://www.hapmap.org/index.html.en) to tag for the genetic variation in the entire *APOB* gene (coding regions and introns) are marked in *red*. Disequilibrium statistics reported as exact values of D′, ranging from −1 to +1, below the diagonal, and of r^2 above the diagonal. All D′ values were positive, indicating that the rare alleles at each locus segregate together. The color code also indicates the degree of linkage disequilibrium between SNPs. MAF, Minor allele frequency.

should tag P2712L, R3611Q, and N4311S because they are in the same haploblock. In this study, although these six SNPs are often on the same haplotypes (all D′>0.90), none are highly correlated, with the exception of P2712L with N4311S (Fig. 2; $r^2 = 1.0$ for P2712L with N4311S; $r^2 = 0.30$ for T2488Tc>t with P2712L/N4311S, all other r^2s < 0.11). This indicates that Ivs18 + 1708 g>t, T2488Tc>t, and E4154K cannot tag or serve as a proxy for P2712L, R3611Q, or N4311S.

Lipids, lipoproteins, and apolipoprotein B levels

Overall, all 10 SNPs were associated with either increases (T71I, Ivs18 + 1708 g>t, T2488Tc>t, R3611Q) or decreases (Ivs4 + 171c>a, A591V, Ivs18 + 379a>c, P2712L, E4154K, N4311S) in total cholesterol, LDL cholesterol, apolipoprotein B, and non-HDL cholesterol ($P = 0.03$ to $P < 0.001$ by ANOVA) (Fig. 3). T71I, Ivs18 + 1708 g>t, T2488Tc>t, and R3611Q were associated with increases in LDL cholesterol of 3.8, 2.8, 2.8, and 2.7% (0.14, 0.10, 0.10, and 0.11 mmol/liter) in heterozygotes *vs.* noncarriers and T71I, Ivs18 + 1708 g>t, and T2488Tc>t also with an increase in LDL cholesterol of 8.2, 6.6, and 6.9% (0.30, 0.24, and 0.25 mmol/liter) in homozygotes *vs.* noncarriers. Ivs4 + 171c>a, A591V, Ivs18 + 379a>c, P2712L, E4154K, and N4311S were associated with decreases in LDL cholesterol of 4.0, 3.4, 2.1, 3.6, 1.9, and 3.6% (0.15, 0.12, 0.08, 0.14, 0.07, and 0.14 mmol/liter) in heterozygotes *vs.* noncarriers, and Ivs4 + 171c>a, A591V, and E4154K were also associated with decreases in LDL cholesterol of 7.4, 4.7, and 4.9% (0.28, 0.18, and 0.18 mmol/liter) in homozygotes *vs.* noncarriers. In absolute values, the maximum increase in LDL cholesterol as a function of SNP genotype was 0.30 mmol/liter for T71I homozygotes *vs.* noncarriers, whereas the maximum decrease in LDL cholesterol was 0.28 mmol/liter for Ivs4 + 171c>a homozygotes *vs.* noncarriers. None of the 10 SNPs were associated with increases or decreases in plasma levels of triglycerides, VLDL cholesterol, HDL cholesterol, or apolipoprotein AI (data not shown).

Of the estimated haplotypes containing the seven *APOB* tag SNPs as defined by HapMap (T71I, Ivs4 + 171c>a, A591V, Ivs18 + 379a>c, Ivs18 + 1708 g>t, T2488Tc>t, and E4154K), nine haplotypes had a frequency above 1% in the general population (supplemental Table 1, published as supplemental data on The Endocrine Society's Journals Online Web site at http://jcem.endojournals.org). Overall, these haplotypes associated with variation in levels of total cholesterol, LDL cholesterol, and apolipoprotein B (all global $P < 0.05$). The most common haplotype in the population, IcAattE (T71I: *I*, Ivs4 + 171c>a: *c*, A591V: *A*, Ivs18 + 379a>c: *a*, Ivs18 + 1708 g>t: *t*, T2488Tc>t: *t*, and E4154K: *E*) associated with the highest levels of total cholesterol, LDL cholesterol, and apolipoprotein B (6.34 mmol/liter, 3.96 mmol/liter, and 90.0 mg/dl, respectively) and was also a combination of the seven single-site tag SNPs associated with the highest LDL cholesterol levels. Compared with this haplotype, the second most common haplotype (TcVcgcE) was associated with significant but modest reductions in total cholesterol, LDL cholesterol, and apolipoprotein B of, respectively, 0.15 mmol/liter, 0.16 mmol/liter, and 2.80 mg/dl. Two other less common haplotypes (TaAagcE and TaVagcK) were associated with significant decreases in total cholesterol of, respectively, 0.19 and 0.17 mmol/liter and decreases in LDL cholesterol of 0.19 and 0.18 mmol/liter. The latter three haplotypes share the T, g, and c alleles of, respectively, *T*71I, Ivs18 + 1708*g*>t, and T2488T*c*>t, which are all associated with the lowest LDL cholesterol levels for these SNPs. Thus, the span in cholesterol levels

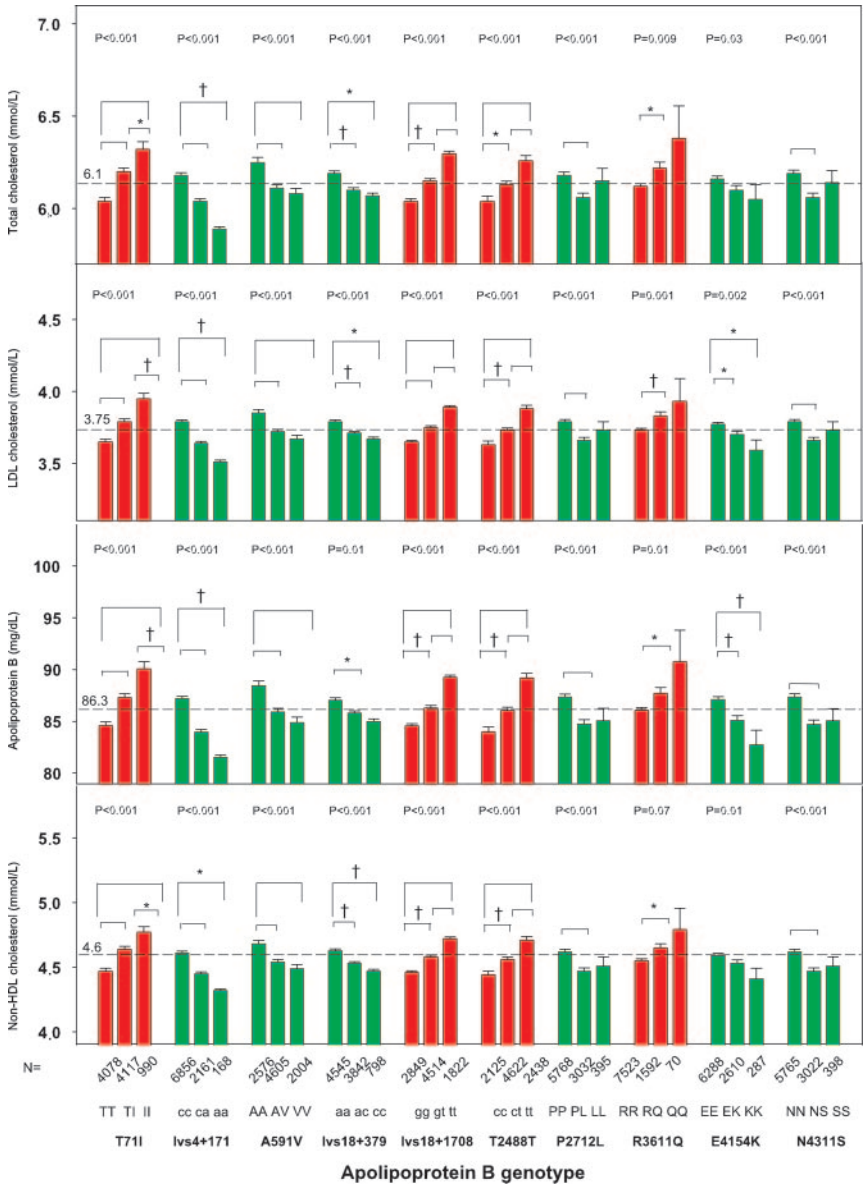

FIG. 3. Plasma levels of total cholesterol, LDL cholesterol, apolipoprotein B, and non-HDL cholesterol as a function of SNP genotypes in the general population, The Copenhagen City Heart Study. Values are means ± SE. Number of individuals within each genotype is given below the bars. *P* values above bars by ANOVA. *Post hoc* tests by Student's *t* test: *, $P < 0.05$; †, $P < 0.01$; ‡, $P < 0.001$. Bonferroni corrected significance level, $P < 0.008$. *Dashed lines* are population mean values.

from the most common haplotype associated with the highest levels of total and LDL cholesterol (IcAattE) to that with the lowest significant levels (TaAagcE) was a modest 0.19 mmol/liter for both total and LDL cholesterol.

Risk of ischemic heart disease and myocardial infarction

Characteristics of individuals in the general population by outcome are shown in Table 1. Despite the effects on plasma levels of total cholesterol, LDL cholesterol, and apolipoprotein B (Fig. 3 and supplemental Table 1), neither the 10 SNPs alone nor the estimated nine most common haplotypes containing the seven *APOB* tag SNPs from HapMap predicted risk of ischemic heart disease or myocardial infarction in the general population (Fig. 4 and supplemental Table 2). Furthermore, results on risk of ischemic heart disease were verified for all 10 SNPs in a case-control study including 944 cases with ischemic heart disease and 7664 controls (supplemental Table 3).

The prospective study had 80% power to detect a hazard ratio of 1.13 or above and 90% power to detect a hazard ratio of 1.15 or above for risk of ischemic heart disease in heterozygotes *vs.* noncarriers for the least frequent SNP (R3611Q: heterozygote frequency 0.17) (supplemental Fig. 1).

Evolutionary conservation of nonsynonymous sequence variations in APOB

Of the three *APOB* mutations (R3480P, R3500Q, R3531C) with known functional effects on LDL metabolism (2, 4), only two were highly conserved from humans to chicken (R3500Q, R3531C) and zebra fish (R3500Q) (Fig. 5). Of the six nonsynonymous SNPs in the present study, two (A591V, P2712L) were at amino acid residues conserved from humans to zebra fish and sea urchin, and one (T71I) was conserved in mammals. Two substituted amino acids conserved in primates and some mammals (E4154K, N4311S), whereas the last SNP (R3611Q) changed an amino acid conserved in primates only.

Five of the six nonsynonymous SNPs were predicted to be benign by all three or the two available algorithms, whereas only P2712L was predicted to be possibly or probably deleterious by SIFT and PolyPhen, respectively, despite the fact that all six nonsynonymous SNPs affected lipid phenotype. Of the three mutations known to have varying degrees of functional effects on LDL metabolism, R3500Q, which reduces the fractional catabolic rate of LDL by 33%, compared with noncarriers, and is associated with severe hypercholesterolemia and an increased risk of ischemic heart disease in the general population (2–4), was predicted to be benign by PolyPhen and possibly deleterious by SIFT and PANTHER. However, R3531C, which is associated with a marginal reduction in fractional catabolic rate of LDL of 12% and is not associated with hypercholesterolemia or risk of ischemic heart disease in the general population (2, 4), was predicted to be deleterious by both PANTHER and PolyPhen. Finally, R3480P, which is associated with a reduction in fractional catabolic rate of LDL of 26% but with an even larger reduction in the production rate of LDL from VLDL and with hypocholesterolemia in the general population (4), was predicted to be possibly deleterious by both PANTHER and PolyPhen.

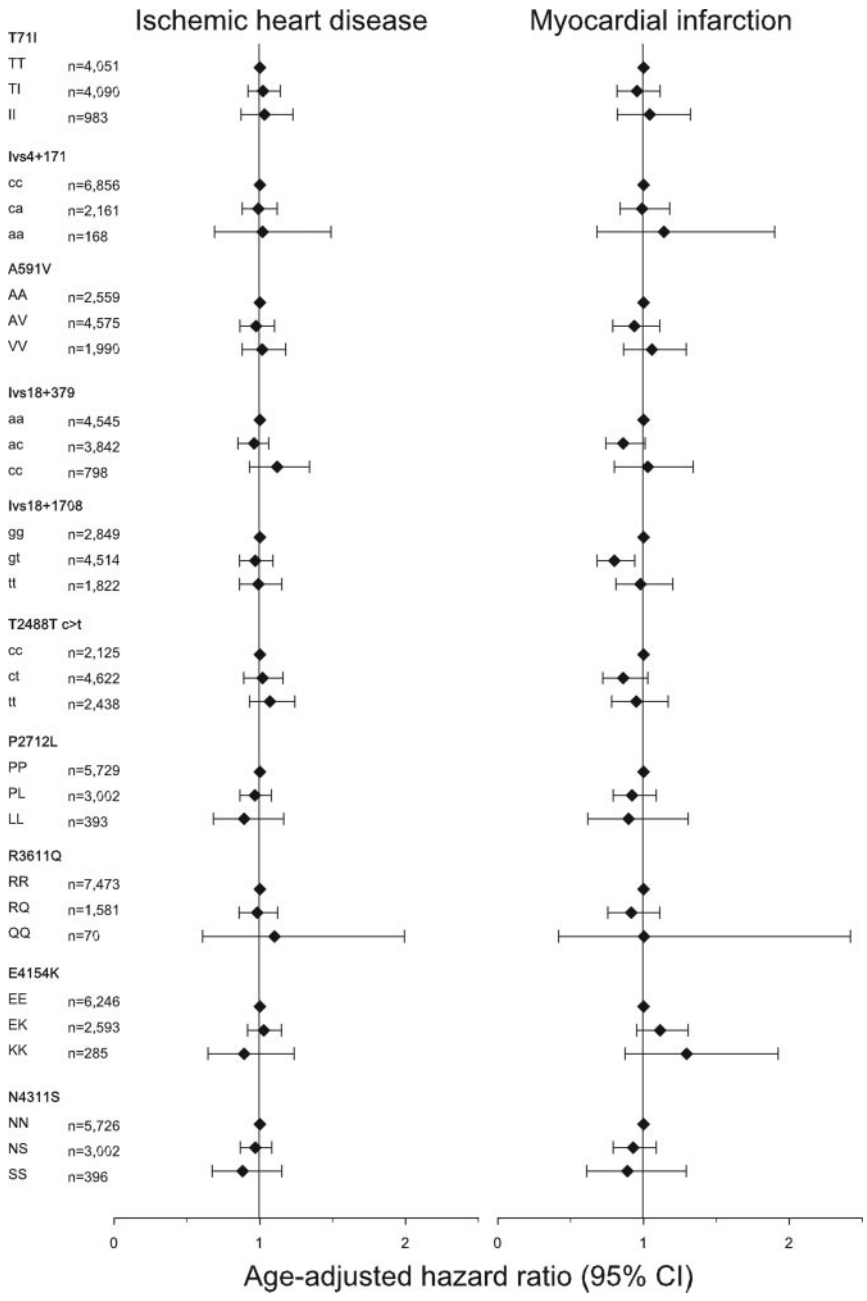

FIG. 4. Risk of ischemic heart disease and myocardial infarction by *APOB* T71I, Ivs4 + 171c>a, A591V, Ivs18 + 379a>c, Ivs18 + 1708 g>t, T2488Tc>t, P2712L, R3611Q, E4154K, and N4311S genotypes in the general population, The Copenhagen City Heart Study. CI, Confidence interval.

Discussion

The principal finding of this study was that six nonsynonymous SNPs in important functional domains of apolipoprotein B, a synonymous SNP, and three noncoding SNPs contribute to plasma levels of LDL cholesterol and apolipoprotein B in the general population. Four of the six nonsynonymous SNPs were associated with lower plasma levels of LDL cholesterol. These sequence variants most likely lower LDL cholesterol levels by either interfering with lipidation of nascent apolipoprotein B (A591V) by reducing the production of LDL from VLDL as we have previously reported for R3480P (4) or accelerating LDL clearance by the LDL receptor (P2712L/N4311S, E4154K). Two SNPs (T71I, R3611K) were associated with increases in LDL cholesterol and are likely to do so by mechanisms that are the opposite of those discussed above. Two SNPs in introns were associated with lower plasma levels of LDL cholesterol (Ivs4 + 171c>a and Ivs18 + 379a>c) and one intron SNP and a synonymous SNP (Ivs18 + 1708 g>t and T2488Tc>t) with an increase in LDL cholesterol. These SNPs are not predicted to be functional (in potential regulatory regions or splice site variants), although little is in fact known about the intronic regions of *APOB*. Alternatively, these SNPs could be in linkage disequilibrium with other functional SNPs. Using HapMap, we identified six SNPs in strong linkage disequilibrium and highly correlated with five (Ivs4 + 171c>a, A591V, Ivs18 + 379a>c, Ivs18 + 1708 g>t, and E4154K) of the seven tag SNPs in *APOB*. However, all six were SNPs in introns, and none were predicted to be functional.

If we include three previously described mutations in apolipoprotein B (R3480P, R3500Q, R3531C) (2–4), the spectrum of apolipoprotein B alleles associated with variation in LDL cholesterol spans a wide range of allele frequencies (from 0.0004 for R3480P to 0.48 for T2488Tc>t) and a magnitude of phenotypic effects (from 29% LDL cholesterol reduction for R3480P to 80% increase for R3500Q). Taken together, our results together with those from previous studies of mutations in *APOB* suggest that multiple common and rare alleles in *APOB* contribute to plasma levels of LDL cholesterol in the general population, although the effects of the common alleles and haplotypes are modest(2–4, 12).

This is the first prospective study to report associations between these 10 SNPs in *APOB* and levels of total cholesterol, LDL cholesterol, apolipoprotein B, and non-HDL cholesterol as well as the ability of these SNPs to predict risk of ischemic heart disease and myocardial infarction with 25 yr follow-up in a large general population cohort. Several case-control studies have reported conflicting associations between *APOB* SNPs and lipid and lipoprotein levels, but some studies reported increased levels of cholesterol and/or apolipoprotein B for T71I and reduced levels of cholesterol and apolipoprotein B for P2712L/N4311S as in the present study (21–23). Only one study previously reported on R3611Q and only three studies on E4154K, but all failed to find an association with variation in lipid and lipoprotein levels (24). However, the power of previous studies to convincingly

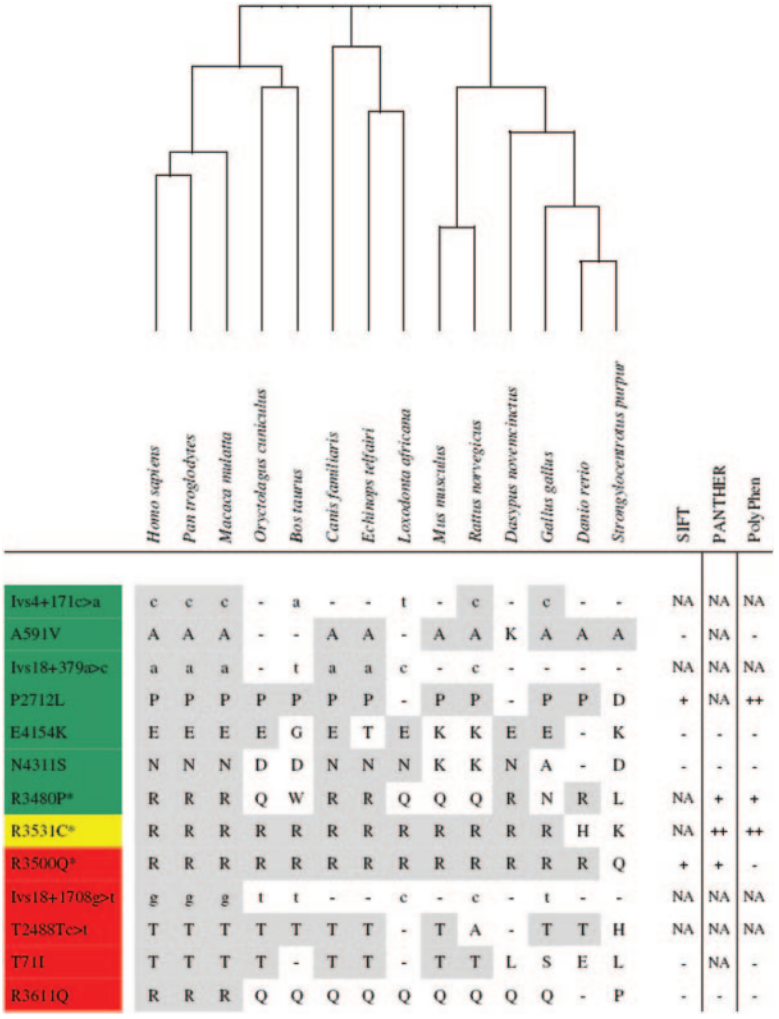

FIG. 5. Evolutionary sequence conservation and predicted functional effects of common nonsynonymous SNPs and rare mutations (*) in *APOB*. *Top*, Cladrogram depicting the relationship between the various apolipoprotein B protein sequences aligned. *Bottom*, Variants associated with decreased LDL cholesterol levels in the general population (*green*), variants without effect on LDL cholesterol in the general population (*yellow*), and variants associated with increased LDL cholesterol levels in the general population (*red*) (2, 4, 12). *Shaded boxes* indicate evolutionary sequence conservation between aligned sequences. The rare mutations R3480P and R3500Q are associated with, respectively, a substantial reduction of 29% and a substantial increase of 80% in plasma levels of LDL cholesterol in the general population. The R3531C mutation is not associated with effects on plasma LDL cholesterol levels in the general population but with a slight reduction in the fractional catabolic rate of LDL cholesterol of 12% (2, 4, 12). The sequences are shown for: *Homo sapiens* (human), *Pan troglodytes* (chimpanzee), *Macaca mulatta* (rhesus monkey), *Oryctolagus cuniculus* (rabbit), *Bos taurus* (cow), *Canis familiaris* (dog), *Echinops telfairi* (hedgehog), *Loxodonta Africana* (African elephant), *Mus musculus* (mouse), *Rattus norvegicus* (rat), *Dasypus novemcinctus* (nine banded armadillo), *Gallus gallus* (chicken), *Danio rerio* (zebra fish), and *Strongylocentrotus purpur* (sea urchin). The predicted effect of each amino acid variant on protein function is shown to the right. SIFT version 1(19): −, tolerated; + deleterious (low confidence prediction); 2+, deleterious. PANTHER version 6(20): −, unlikely functional effect; +, possible deleterious functional effect; 2+, high probability of deleterious functional effect. PolyPhen (18): −, benign; +, possibly damaging; 2+, probably damaging. NA, Not modeled by the algorithm.

show an effect on lipid or lipoprotein levels has generally been very limited due to the size of the studies.

The effects of these SNPs on risk of ischemic heart disease have been even more contradictory, as exemplified by two recent metaanalyses on risk of ischemic heart disease conferred by E4154K: one of these reported an increased risk of ischemic heart disease and an overall odds ratio of 1.32 (1.14–1.54; n = 3870) (5), whereas the other reported no increase in risk and an overall odds ratio of 1.15 (0.78–1.70; n = 2014) (6). Both of these metaanalyses were considerably smaller than the present study alone. In this study, we had 80% power to exclude a hazard ratio at or above 1.10 for E4154K heterozygotes. E4154K heterozygosity (EK) and homozygosity (KK) was associated with, respectively, a 0.07 and 0.18 mmol/liter decrease in LDL cholesterol levels and did not predict risk of ischemic heart disease.

In The Copenhagen City Heart Study, a 0.5 mmol/liter increase in LDL cholesterol predicts a hazard ratio of 1.06 (95% confidence interval 1.03–1.09) for ischemic heart disease and a 0.5 mmol/liter decrease in LDL cholesterol predicts a hazard ratio of 0.94 (0.92–0.97). Thus, the effect on risk of ischemic heart disease predicted by the largest observed effects on LDL cholesterol for the SNPs and haplotypes studied (0.30 mmol/liter increase for T71I) would correspond to a hazard ratio close to 1, in agreement with the lack of effect of both SNPs and haplotpes on risk of ischemic heart disease found in this study.

The extent of evolutionary sequence conservation did not reliably predict the impact of either SNPs or mutations on protein function. Only one computer-based prediction algorithm predicted one of the six SNPs associated with plasma LDL cholesterol level to have a deleterious effect with high probability. Moreover, one of the three functional mutations (R3500Q) was predicted to be benign by one algorithm, and this was the mutation with by far the largest functional effect and the largest effect on phenotype. In contrast, the mutation with the smallest functional effect (R3531C) was predicted by the two algorithms available to have a deleterious effect with high probability. Thus, *in silico* prediction methods were poor predictors of functional variation in apolipoprotein B. This suggests that there are limitations to the usefulness of these methods, as demonstrated previously for genetic variation in *PCSK9* (25). Taken together, this suggests that neither evolutionary sequence variation nor computer-based prediction of functional effects can reliably predict effects of nonsynonymous mutations or SNPs in *APOB* on plasma levels of lipids, lipoproteins, or apolipoprotein B. For the nonsynonymous SNPs, this assumes that no unknown nonsynonymous or regulatory SNPs in linkage disequilibrium with these six SNPs can account for the lipid phenotypes. We cannot exclude that such SNPs might exist, although at present none have been described in HapMap or in other SNP databases.

Our data support a role for common variants in *APOB* in

determining plasma levels of LDL cholesterol. Not surprisingly, the effects of these common variants were much smaller than those of the previously described rare mutations in *APOB*. The associations with plasma LDL cholesterol levels for all six nonsynonymous SNPs would not have been revealed using the current tag SNPs available in HapMap because the linkage analysis showed that the correlations (r^2s) between some of these tag SNPs and the nonsynonymous SNPs in the proposed haploblock were modest. This suggests that much denser SNP analysis of *APOB* and other genes is required for HapMap to be useful in genetic association studies.

In conclusion, the present results together with those from previous studies of mutations in *APOB* suggest that multiple common and rare alleles in *APOB* contribute significantly to plasma levels of LDL cholesterol in the general population, although the effects of the common alleles and haplotypes are modest.

Acknowledgments

The authors thank Mette Refstrup and Hanne Damm for expert technical assistance. We are indebted to the staff and participants of The Copenhagen City Heart Study for their important contributions.

Address all correspondence and requests for reprints to: Anne Tybjærg-Hansen, M.D., D.M.Sc., Chief Physician, Associate Professor, Department of Clinical Biochemistry KB3011, Section for Molecular Genetics, Rigshospitalet, Copenhagen University Hospital, Blegdamsvej 9, DK-2100 Copenhagen Ø, Denmark. E-mail: at-h@rh.regionh.dk.

This work was supported by grants from The Danish Heart Foundation, The Danish Medical Research Council, the Research Fund at Rigshospitalet, Copenhagen University Hospital, Chief Physician Johan Boserup and Lise Boserup's Fund, Ingeborg and Leo Dannin's Grant, and Henry Hansen and Wife's Grant.

Disclosure Statement: The authors have nothing to disclose.

References

1. **Beekman M, Heijmans BT, Martin NG, Pedersen NL, Whitfield JB, DeFaire U, van Baal GC, Snieder H, Vogler GP, Slagboom PE, Boomsma DI** 2002 Heritabilities of apolipoprotein and lipid levels in three countries. Twin Res 5:87–97
2. **Tybjærg-Hansen A, Steffensen R, Meinertz H, Schnohr P, Nordestgaard BG** 1998 Association of mutations in the apolipoprotein B gene with hypercholesterolemia and the risk of ischemic heart disease. N Engl J Med 338:1577–1584
3. **Tybjærg-Hansen A, Jensen HK, Benn M, Steffensen R, Jensen G, Nordestgaard BG** 2005 Phenotype of heterozygotes for low-density lipoprotein receptor mutations identified in different background populations. Arterioscler Thromb Vasc Biol 25:211–215
4. **Benn M, Nordestgaard BG, Jensen JS, Nilausen K, Meinertz H, Tybjærg-Hansen A** 2005 Mutation in apolipoprotein B associated with hypobetalipoproteinemia despite decreased binding to the low density lipoprotein receptor. J Biol Chem 280:21052–21060
5. **Chiodini BD, Barlera S, Franzosi MG, Beceiro VL, Introna M, Tognoni G** 2003 APO B gene polymorphisms and coronary artery disease: a meta-analysis. Atherosclerosis 167:355–366
6. **Boekholdt SM, Peters RJG, Fountoulaki K, Kastelein JJP, Sijbrands EJG** 2003 Molecular variation at the apolipoprotein B gene locus in relation to lipids and cardiovascular disease: a systematic meta-analysis. Hum Genet 113:417–425
7. **Mann CJ, Anderson TA, Read J, Chester SA, Harrison GB, Kochl S, Ritchie PJ, Bradbury P, Hussain FS, Amey J, Vanloo B, Rosseneu M, Infante R, Hancock JM, Levitt DG, Banaszak LJ, Scott J, Shoulders CC** 1999 The structure of vitellogenin provides a molecular model for the assembly and secretion of atherogenic lipoproteins. J Mol Biol 285:391–408
8. **Bradbury P, Mann CJ, Kochl S, Anderson TA, Chester SA, Hancock JM, Ritchie PJ, Amey J, Harrison GB, Levitt DG, Banaszak LJ, Scott J, Shoulders CC** 1999 A common binding site on the microsomal triglyceride transfer protein for apolipoprotein B and protein disulfide isomerase. J Biol Chem 274:3159–3164
9. **Segrest JP, Jones MK, De Loof H, Dashti N** 2001 Structure of apolipoprotein B-100 in low density lipoproteins. J Lipid Res 42:1346–1367
10. **Chatterton JE, Phillips ML, Curtiss LK, Milne R, Fruchart JC, Schumaker VN** 1995 Immunoelectron microscopy of low-density lipoproteins yields a ribbon and bow model for the conformation of apolipoprotein-B on the lipoprotein surface. J Lipid Res 36:2027–2037
11. **Borén J, Lee I, Zhu W, Arnold K, Taylor S, Innerarity TL** 1998 Identification of the low density lipoprotein receptor-binding site in apolipoprotein B100 and the modulation of its binding activity by the carboxyl terminus in familial defective apo-B100. J Clin Invest 101:1084–1093
12. **Benn M, Nordestgaard BG, Jensen JS, Grande P, Sillesen H, Tybjærg-Hansen A** 2005 Polymorphism in APOB associated with increased low-density lipoprotein levels in both genders in the general population. J Clin Endocrinol Metab 90:5797–5803
13. **Schnohr P, Jensen G, Lange P, Scharling H, Appleyard M** 2001 The Copenhagen City heart study—introduction. Eur Heart J 3(Suppl H):H1–H83
14. **Benn M, Nordestgaard BG, Jensen GB, Tybjærg-Hansen A** 2007 Improving prediction of ischemic cardiovascular disease in the general population using apolipoprotein B. The Copenhagen City Heart Study. Arterioscler Thromb Vasc Biol 27:661–670
15. **Julian DG, Bertrand ME, Hjalmarson A, Fox K, Simoons ML, Ceremuzynski L, Maseri A, Meinertz T, Meyer J, Pyorala K, Rehnqvist N, Tavazzi L, Toutouzas P, Treasure T** 1997 Management of stable angina pectoris—recommendations of the Task Force of the European Society of Cardiology. Eur Heart J 18:394–413
16. **Friedewald WT, Levy RI, Fredrickson DS** 1972 Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin Chem 18:499–502
17. **Purcell S, Daly MJ, Sham PC** 2007 WHAP: haplotype-based association analysis. Bioinformatics 23:255–256
18. **Sunyaev S, Ramensky V, Koch I, Lathe III W, Kondrashov AS, Bork P** 2001 Prediction of deleterious human alleles. Hum Mol Genet 10:591–597
19. **Ng PC, Henikoff S** 2003 SIFT: predicting amino acid changes that affect protein function. Nucleic Acids Res 31:3812–3814
20. **Thomas PD, Kejariwal A, Campbell MJ, Mi H, Diemer K, Guo N, Ladunga I, Ulitsky-Lazareva B, Muruganujan A, Rabkin S, Vandergriff JA, Doremieux O** 2003 PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. Nucleic Acids Res 31:334–341
21. **Keavney B, Palmer A, Parish S, Clark S, Youngman L, Danesh J, McKenzie C, Delepine M, Lathrop M, Peto R, Collins R** 2004 Lipid-related genes and myocardial infarction in 4685 cases and 3460 controls: discrepancies between genotype, blood lipid concentrations, and coronary disease risk. Int J Epidemiol 33:1002–1013
22. **Dunning AM, Renges HH, Xu CF, Peacock R, Brasseur R, Laxer G, Tikkanen MJ, Butler R, Saha N, Hamsten A** 1992 Two amino acid substitutions in apolipoprotein B are in complete allelic association with the antigen group (x/y) polymorphism: evidence for little recombination in the 3′ end of the human gene. Am J Hum Genet 50:208–221
23. **Bentzen J, Jørgensen T, Fenger M** 2002 The effect of six polymorphisms in the apolipoprotein B gene on parameters of lipid metabolism in a Danish population. Clin Genet 61:126–134
24. **Talmud PJ, Barni N, Kessling AM, Carlsson P, Darnfors C, Bjursell G, Galton D, Wynn V, Kirk H, Hayden MR** 1987 Apolipoprotein B gene variants are involved in the determination of serum cholesterol levels: a study in normo- and hyperlipidaemic individuals. Atherosclerosis 67:81–89
25. **Kotowski IK, Pertsemlidis A, Luke A, Cooper RS, Vega GL, Cohen JC, Hobbs HH** 2006 A spectrum of PCSK9 alleles contributes to plasma levels of low-density lipoprotein cholesterol. Am J Hum Genet 78:410–422