Loss-of-Function Mutations in APOC3, Triglycerides, and Coronary Disease

The TG and HDL Working Group of the Exome Sequencing Project, National Heart, Lung, and Blood Institute*

ABSTRACT

BACKGROUND
Plasma triglyceride levels are heritable and are correlated with the risk of coronary heart disease. Sequencing of the protein-coding regions of the human genome (the exome) has the potential to identify rare mutations that have a large effect on phenotype.

METHODS
We sequenced the protein-coding regions of 18,666 genes in each of 3734 participants of European or African ancestry in the Exome Sequencing Project. We conducted tests to determine whether rare mutations in coding sequence, individually or in aggregate within a gene, were associated with plasma triglyceride levels. For mutations associated with triglyceride levels, we subsequently evaluated their association with the risk of coronary heart disease in 110,970 persons.

RESULTS
An aggregate of rare mutations in the gene encoding apolipoprotein C3 (APOC3) was associated with lower plasma triglyceride levels. Among the four mutations that drove this result, three were loss-of-function mutations: a nonsense mutation (R19X) and two splice-site mutations (IVS2+1G→A and IVS3+1G→T). The fourth was a missense mutation (A43T). Approximately 1 in 150 persons in the study was a heterozygous carrier of at least one of these four mutations. Triglyceride levels in the carriers were 39% lower than levels in noncarriers (P<1×10−20), and circulating levels of APOC3 in carriers were 46% lower than levels in noncarriers (P=8×10−10). The risk of coronary heart disease among 498 carriers of any rare APOC3 mutation was 40% lower than the risk among 110,472 noncarriers (odds ratio, 0.60; 95% confidence interval, 0.47 to 0.75; P = 4×10−6).

CONCLUSIONS
Rare mutations that disrupt APOC3 function were associated with lower levels of plasma triglycerides and APOC3. Carriers of these mutations were found to have a reduced risk of coronary heart disease. (Funded by the National Heart, Lung, and Blood Institute and others.)
In observational studies, plasma triglyceride levels are associated with the risk of coronary heart disease. Heritability accounts for more than 50% of the individual variation in triglyceride levels. Genomewide association studies have identified common DNA sequence variants at more than 150 genetic loci that are related to plasma lipids and have suggested that plasma triglyceride-rich lipoproteins directly influence the risk of coronary heart disease. These findings lead to two unanswered questions: first, to what extent do rare DNA sequence variants, particularly those in protein-coding sequences, contribute to individual variation in plasma triglyceride levels and the risk of coronary heart disease at the population level, and second, are there specific genetic variants that might lower triglyceride levels and reduce the risk of coronary heart disease?

Recent advances in DNA sequencing technology allow comprehensive detection of rare DNA sequence variants. When sequencing is performed in large populations, a sufficient number of mutation carriers can be identified to evaluate the correlation of genotype with phenotype. In particular, it is advantageous to focus sequencing on exons, the elements of the genome that code for proteins (collectively called the exome), since mutations in protein-coding sequences (e.g., missense, nonsense, or splice-site mutations) are most readily interpreted.

To address the two questions posed above, we sequenced the exomes of 3734 persons in the United States, identified mutations, and tested the mutations for association with plasma triglyceride levels. We subsequently investigated whether the same mutations were related to the risk of clinical coronary heart disease.

METHODS

STUDY DESIGN AND OVERSIGHT

The study was conducted as part of the Exome Sequencing Project of the National Heart, Lung, and Blood Institute (NHLBI) (https://esp.gs.washington.edu/drupal). The study was designed by the Triglycerides and High-Density Lipoprotein (TG and HDL) Working Group of the Exome Sequencing Project (see the Appendix). The institutional review boards at the Broad Institute and all participating sites approved the NHLBI Exome Sequencing Project study protocols. The last two authors vouch for the accuracy and completeness of the data and all analyses.

DISCOVERY COHORTS

The Exome Sequencing Project study population includes enrollees in seven population-based cohorts (Atherosclerosis Risk in Communities, Coronary Artery Risk Development in Young Adults, the Cardiovascular Health Study, the Framingham Heart Study, the Jackson Heart Study, the Multiethnic Study of Atherosclerosis, and the Women’s Health Initiative) and participants in a study of early-onset myocardial infarction (Myocardial Infarction Genetics Consortium) (see Table S1 in the Supplementary Appendix, available with the full text of this article at NEJM.org).

All participants provided consent for participation in genetic studies. Total cholesterol, HDL cholesterol, and triglyceride levels were measured in fasting venous blood samples, according to standard procedures. Low-density lipoprotein (LDL) cholesterol values were calculated according to the Friedewald formula, with the assignment of a missing value to persons with triglyceride levels higher than 400 mg per deciliter (4.5 mmol per liter).

EXOME SEQUENCING

Exome sequencing was performed at either the University of Washington or the Broad Institute. DNA sequence variants identified at either institution were then submitted to the University of Michigan, where they were subjected to quality control and filtering to exclude variants that were likely to be false positive findings. Protocols for exome sequencing and variant analysis are described in the Supplementary Appendix.

PRIMARY ANALYSES

Sequence Variants and Triglycerides

We used linear regression to test the association of plasma triglyceride levels with identified DNA sequence variants in race-specific models and in models including multiple races. The outcome variable was the natural logarithm of the triglyceride level. The primary independent variable was the number of copies of the variant allele (coded as 0, 1, or 2). Covariates included age, sex, two principal components of ancestry, an indicator variable for race (in the model in which races were combined), and indicator variables for the sequencing ascertainment scheme (see page 45 of the Supplementary Appendix for further in-
formation on the covariate of age). The single-variant association analysis was limited to variants with a frequency above 0.1%, corresponding to variants with approximately seven copies of the minor allele among the 3734 participants.

Since many of the identified variants occurred too infrequently to be analyzed individually, we aggregated rare variants within each gene; here, the gene (rather than an individual variant) became the unit of analysis. Within each gene, we included all missense, nonsense, and splice-site variants below a frequency threshold of 1% and considered whether plasma triglyceride levels differed between carriers of these mutations and noncarriers.\(^9\) Owing to considerations regarding statistical power, we limited the analysis to genes for which five or more persons carried minor alleles.

**Replication of Association between APOC3 Genotypes and Plasma Lipids**

The gene most strongly associated with plasma triglyceride levels in the discovery sample was the gene encoding apolipoprotein C3 (APOC3). To replicate this finding, we genotyped four APOC3 mutations (R19X, IVS2+1G→A, A43T, and IVS3+1G→T), using the HumanExome BeadChip (Illumina) in 41,671 participants of African or European ancestry in seven replication studies (Table S2 in the Supplementary Appendix). This cohort was separate from the discovery cohort. We tested the association of these four mutations with plasma lipid levels. The methods used for genotyping and association testing are described in the Supplementary Appendix.

**APOC3 Genotypes, APOC3 Levels, and Coronary Heart Disease**

We next tested the association of APOC3 mutations with coronary heart disease in participants of European, African, and Hispanic ancestry in 15 studies (for details, see Table S3 in the Supplementary Appendix). To investigate the association between APOC3 level and the risk of coronary heart disease, we used blood samples drawn from 3238 fasting participants in the Framingham Heart Study offspring cohort (examination cycle 5, 1991 to 1995).\(^20\) Plasma APOC3 levels were measured with the use of a commercially available immunochemical assay (Wako Diagnostics). All participants underwent continual surveillance for incident coronary heart disease events until December 31, 2010. As described in the Supplementary Appendix, we used proportional-hazards regression models to examine the relationship of plasma APOC3 levels (logarithmically transformed on a natural log scale) to the risk of incident coronary heart disease.

**SECONDARY ANALYSES**

In patients with angiographic coronary heart disease, we evaluated the association of plasma APOC3 levels with total and cardiovascular mortality. We also tested the association of APOC3 mutations with the presence of hepatic fat on computed tomography (CT). Details of the methods are presented in the Supplementary Appendix.

**STATISTICAL ANALYSIS**

For the single-variant analysis, we considered a P value of less than 5×10\(^{-7}\) to be significant because this threshold accounts for the 95,342 variants tested. For the gene-level analysis, we considered a P value of less than 3×10\(^{-6}\) to be significant to account for the testing of 18,666 genes. In the analyses of associations with coronary heart disease, we tested four variants and thus considered a P value of less than 0.0125 to be significant. Genetic association analyses were performed with use of PLINK (http://pngu.mgh.harvard.edu/~purcell/plink),\(^21\) SCORE-Seq (http://dlin.web.unc.edu/software/SCORE-Seq),\(^22\) SAS software, version 9.1, or R software, version R=2.14 (www.r-project.org).

**RESULTS**

**BASELINE CHARACTERISTICS OF THE STUDY PARTICIPANTS**

Among the 6823 persons for whom sequencing was performed through the NHLBI Exome Sequencing Project, 2861 were excluded because data on fasting triglyceride levels were not available, 149 were excluded because of relatedness, and 79 were excluded because of sample rejection during the quality-control process. These exclusions left 3734 persons eligible for the present investigation (the baseline characteristics of these participants are shown in Table S1 in the Supplementary Appendix). There were 2082 participants of European ancestry and 1652 of African ancestry.

**EXOME SEQUENCING**

We obtained sequence data for 256,143 exons from 18,666 genes. On average, targeted bases were read 89 times and 94% of targeted bases were covered...
with at least 20 sequencing reads. Each participant had an average of 46 mutations causing the gain or loss of a stop codon, 5712 had missense mutations, and 19 had mutations that altered canonical splice sites.

**ASSOCIATION OF SINGLE VARIANTS WITH TRIGLYCERIDES**

We tested the association of plasma triglyceride levels with 95,342 nonsense, missense, or splice-site mutations with an allele frequency greater than 0.1% (Fig. S1 and Tables S4, S5, and S6 in the Supplementary Appendix). In analyses of the full sample, the variant most strongly associated with triglyceride levels was the previously described S19W variant in APOA5,23 which had an allele frequency of 6% among participants of European ancestry (P=3×10⁻⁷).

**GENE-LEVEL ASSOCIATIONS WITH TRIGLYCERIDES AND REPLICATION**

We next assessed whether rare mutations (i.e., missense, nonsense, or splice-site variants with a frequency of <1%) tested collectively within each gene were associated with plasma triglyceride levels. APOC3 was the gene most strongly associated with plasma triglyceride levels when analyses included only participants of European ancestry (P=7×10⁻⁶) and when analyses included participants of European ancestry and participants of African ancestry (P=1×10⁻⁵) (Table S7 in the Supplementary Appendix). An aggregate of rare alleles at APOC3 was associated with lower plasma triglyceride levels (Table 1 and Fig. 1). As compared with noncarriers, carriers of any rare APOC3 mutation had plasma triglyceride levels that were 39% lower (P=6×10⁻⁹), HDL cholesterol levels that were 22% higher (P=3×10⁻⁶), and LDL cholesterol levels that were 16% lower (P=0.05) (Table 1). These rare mutations collectively explained 0.9% of the variance in triglyceride levels.

The association between plasma lipid levels and APOC3 was primarily attributable to four different sites within the gene. Three of the four mutations are predicted to severely disrupt the function of APOC3 — that is, lead to loss of function.24 Loss-of-function variants included

| Table 1. Rare Mutations in APOC3 and Plasma Levels of Triglycerides.* |
|-------------------|----------------|----------------|
| Chromosome Position | Mutation | Mutation Type |
| | No. of Carriers | Mean Triglyceride Level | Mean HDL Cholesterol Level | Mean LDL Cholesterol Level |
| Chr:11:116701284 | IVS1–2G→A | Splice site | 1 | 66.0 | 65.0 | 140.0 |
| Chr:11:116701326 | A10T | Missense | 2 | 71.5 | 64.0 | 138.5 |
| Chr:11:116701353 | R19X | Nonsense | 2 | 45.5 | 56.8 | 152.5 |
| Chr:11:116701354 | IVS2+1G→A | Splice site | 10 | 71.0 | 59.6 | 111.7 |
| Chr:11:116701560 | A43T | Missense | 12 | 79.0 | 64.5 | 110.4 |
| Chr:11:116701613 | IVS3+1G→T | Splice site | 5 | 82.8 | 67.0 | 154.2 |
| Chr:11:116703493 | D65 N | Missense | 1 | 415.0 | 32.0 | NA |
| Total APOC3 carriers | 33 | 84.5±86.0 | 61.9±16.6 | 122.8±54.8 |
| Total APOC3 noncarriers | 3701 | 137.5±90.1 | 50.7±15.6 | 146.2±61.8 |
| Carriers vs. noncarriers (%) | −38.5 | 22.1 | −16.0 |
| P value‡ | 6×10⁻⁶‡ | 4×10⁻⁶ | 0.05 |

* Plus–minus values are means ±SD. The APOC3 mutations were identified by means of exome sequencing. To convert the values for cholesterol to millimoles per liter, multiply by 0.02586. To convert the values for triglycerides to millimoles per liter, multiply by 0.01129. HDL denotes high-density lipoprotein, LDL low-density lipoprotein, and NA not available.

† P values were derived from a linear regression model, with adjustments for age, sex, ancestry, and principal components of ancestry.

‡ The P value for the triglyceride phenotype is based on triglyceride levels that were logarithmically transformed on a natural log scale.
A nonsense substitution (R19X) and two DNA sequence variants that disrupt a splice site (IVS2+1G→A and IVS3+1G→T). The missense A43T mutation has been shown to decrease APOC3 function in vitro.25,26 When the participants of European ancestry were combined with the participants of African ancestry, approximately 1 in 150 persons carried any one of these four loss-of-function mutations (Table S8 in the Supplementary Appendix).27

To replicate the results derived from exome sequencing, we genotyped one or more of these four mutations in 34,432 participants of European ancestry and 7239 participants of African ancestry, independently of the discovery sequencing study. When compared with participants who did not carry any of the four mutations, those who were heterozygous for any of the four mutations had lower plasma triglyceride levels among participants of European ancestry (42% lower, P<1×10−20), among participants of African ancestry (32% lower, P=1×10−20), and among participants in the two groups combined (39% lower, P=4×10−20) (Table 2, and Table S9 in the Supplementary Appendix). APOC3 loss-of-function carrier status was associated with reduced levels of plasma triglycerides whether or not the S19W variant of APOA5 was present (Table S10 in the Supplementary Appendix). APOC3 loss-of-function mutation carriers also had significantly higher HDL cholesterol levels, with a trend toward lower LDL cholesterol levels (Table 2, and Table S9 in the Supplementary Appendix).

### APOC3 LOSS-OF-FUNCTION MUTATIONS, CORONARY HEART DISEASE, AND HEPATIC FAT

Among 110,970 study participants (34,002 patients with coronary heart disease and 76,968 controls) from 14 studies, 498 were heterozygous for at least one of the four APOC3 loss-of-function mutations. The risk of coronary heart disease for carriers of an APOC3 loss-of-function mutation was 40% lower than the risk for noncarriers (odds ratio, 0.60; 95% confidence interval [CI], 0.47 to 0.75; P=4×10−6) (Fig. 2, and Tables S3 and S11 in the Supplementary Appendix). There was no evidence of heterogeneity across the 15 studies (Breslow–Day test for heterogeneity, P=0.16).

We did not observe homozygosity or compound heterozygosity for APOC3 loss-of-function variants in any participant. However, with a combined allele frequency of 1:300 and genotyping of 110,970 participants, the expected mean (±SD) number of persons with homozygosity or compound heterozygosity for APOC3 loss-of-function variants was only 1±1 (Table S12 in the Supplementary Appendix).

A total of 3051 participants had both an APOC3 loss-of-function allele and a phenotype of hepatic steatosis on CT. The presence of hepatic fat on CT did not differ significantly between carriers of
Table 2. Association of Four APOC3 Mutations with Plasma Lipid Levels in Replication Studies.*

<table>
<thead>
<tr>
<th>Variable Affected by Mutation</th>
<th>Carriers of Any of Four APOC3 Mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>European Ancestry (N = 34,432)</td>
</tr>
<tr>
<td>Triglycerides (% difference)</td>
<td>−42</td>
</tr>
<tr>
<td>P value</td>
<td>&lt;1×10⁻²⁰</td>
</tr>
<tr>
<td>LDL cholesterol (mg/dl)</td>
<td>−9.3±3.4</td>
</tr>
<tr>
<td>P value</td>
<td>0.006</td>
</tr>
<tr>
<td>HDL cholesterol (mg/dl)</td>
<td>11.5±1.3</td>
</tr>
<tr>
<td>P value</td>
<td>&lt;1×10⁻²⁰</td>
</tr>
</tbody>
</table>

* Plus–minus values indicate effect size ±SE. Effect size represents the difference in mean lipid levels between carriers and noncarriers after adjustment for covariates. The four APOC3 mutations tested were R19X, splice-site mutation IVS2+1G→A, A43T, and splice-site mutation IVS3+1G→T. P values are for the comparison with noncarriers. P values were derived from a linear regression model, with adjustments for age, sex, ancestry, and principal components of ancestry. To convert the values for cholesterol to millimoles per liter, multiply by 0.01129. To convert the values for triglycerides to millimoles per liter, multiply by 0.02586. To convert the values for cholesterol to millimoles per liter, multiply by 0.01129.

Discussion

By sequencing the protein-coding regions of the genome in 3734 persons, we identified several rare coding-sequence variants of APOC3 that were associated with a large effect on plasma triglyceride levels. Approximately 1 in 150 persons carried any one of four protein-altering or splice-site variants. As compared with noncarriers, carriers of any of these four APOC3 mutations had plasma triglyceride levels that were 39% lower, HDL cholesterol levels that were 22% higher, LDL cholesterol levels that were 16% lower, and circulating APOC3 levels that were 46% lower, and they had a risk of coronary heart disease that was reduced by 40%.

These results show that loss of APOC3 function confers protection against clinical coronary heart disease. In a previous study, Pollin and colleagues reported that approximately 5% of Amish persons in Lancaster County, Pennsylvania, carried the APOC3 R19X null allele and that these carriers had a favorable lipid profile and a reduced amount of coronary artery calcium, a surrogate marker for atherosclerosis. The A43T missense mutation in APOC3 (referred to as A23T in the first coronary heart disease event. In models adjusted for age and sex, each decrease of 1 mg per deciliter in plasma levels of APOC3 was associated with a 4% decrease in the risk of incident coronary heart disease (hazard ratio, 0.96; 95% CI, 0.94 to 0.98; P<0.001) (Table S15 in the Supplementary Appendix). Participants with values in the lowest third of the distribution of plasma APOC3 levels had a reduced risk of incident coronary heart disease, as compared with participants with values in the highest third (odds ratio, 0.65; 95% CI, 0.48 to 0.87; P = 0.003) (Fig. 3, and Table S16 in the Supplementary Appendix). However, this association was not significant after adjustment for additional cardiovascular risk factors (Tables S15 and S16 in the Supplementary Appendix). In the Verona Heart Study cohort, which consisted of 794 patients with angiographic evidence of coronary artery disease, the lowest third of the distribution of APOC3 levels was associated with a reduced risk of death from cardiovascular disease, as compared with the highest third, in models that were minimally adjusted and fully adjusted for risk factors (Tables S17 through S20 in the Supplementary Appendix).
earlier nomenclature) was initially identified in two Mayan Indians with low plasma levels of triglycerides and APOC3. Functional characterization in vitro showed deficient lipid binding and attenuated secretion of very-low-density lipoprotein from hepatocytes. Recently, in three families, 12 carriers of the A43T mutation or the two APOC3 splice-site mutations were identified, and as compared with family members who were not carriers, the carriers had lower levels of plasma triglycerides and higher levels of HDL cholesterol. In our study, which included 110,970 participants, carriers of APOC3 loss-of-function mutations had a reduced risk of clinical coronary heart disease.

Several mechanisms may link APOC3 to atherogenesis. Resident on the surface of triglyceride-rich lipoproteins, APOC3 inhibits the lipolytic activity of lipoprotein lipase, thereby increasing the plasma level of atherogenic triglyceride-rich lipoproteins, including very-low-density lipoproteins and chylomicrons. APOC3 also delays the clearance of atherogenic, cholesterol-rich remnants of the postprandial metabolism of triglyceride-rich lipoproteins.

The question of whether triglyceride-rich lipoproteins in plasma have a causal effect on coronary heart disease remains unanswered. Several treatments that lower plasma levels of triglycerides, including fibrates and fish oils,
have failed to consistently reduce the risk of coronary heart disease in randomized, controlled trials. Our findings with respect to the human APOC3 loss-of-function alleles suggest that a lifelong decrease in APOC3 function is one means of reducing plasma levels of triglyceride-rich lipoproteins that may be atheroprotective. Recently, an antisense oligonucleotide therapeutic agent that decreased the production of APOC3 was shown to lower plasma levels of triglycerides and APOC3 in mice, nonhuman primates, and healthy human volunteers. Our data provide confidence that such a therapeutic strategy might be expected to reduce the risk of clinical coronary heart disease.

In more general terms, the results of our study highlight the potential usefulness of naturally occurring loss-of-function mutations in guiding the selection of therapeutic targets. For example, loss-of-function mutations in the gene encoding proprotein convertase subtilisin/kexin type 9 (PCSK9) were shown to reduce LDL cholesterol levels and lower the risk of coronary heart disease, a finding that prompted the development of monoclonal antibodies directed against PCSK9.

Several limitations of our study deserve mention. First, we are unable to pinpoint the specific mechanism by which loss of APOC3 function might be atheroprotective. Carriers of APOC3 loss-of-function mutations have a range of associated biomarkers, including lower levels of triglycerides, higher levels of HDL cholesterol, lower levels of LDL cholesterol, and lower levels of APOC3. Additional work will be required to understand the primary mechanism linking APOC3 to coronary heart disease. Second, earlier reports are conflicting as to whether common DNA sequence variants at APOC3 affect hepatic steatosis. We found no significant difference in CT findings with respect to hepatic fat between study participants who carried APOC3 loss-of-function mutations and those who did not. However, larger samples may be required to confidently answer this question.

In conclusion, we identified rare DNA sequence variants in APOC3 that were associated with lifelong reductions in plasma levels of triglycerides and APOC3. These variants conferred protection against coronary heart disease.

The views expressed in this article are solely those of the authors and do not necessarily represent the official views of the National Heart, Lung, and Blood Institute (NHLBI) or the National Institutes of Health (NIH).

Supported by the Grand Opportunity (GO) Exome Sequencing Project (ESP), through which grants were provided by the NHLBI (RC2 HL-103010 [HeartGo], RC2 HL-102923 [LungGo], RC2 HL-102924 [Women’s Health Initiative ESP], RC2 HL-102925 [BroadGo], and RC2 HL-102926 [SeattleGo]); the Russell and Diana Hawkins Family Foundation Discovery Fellowship, an award from the British Heart Foundation (BHF) (to Dr. Samani, who holds a chair supported by the BHF); a Banting Fellowship from the Canadian Institutes of Health Research (to Dr. Do), a career-development award from the NIH and NHLBI (K08-HL114642, to Dr. Stitziel), a Research Scholar Award from the Massachusetts General Hospital (MGH) (to Dr. Kathiresan), the Howard M. Goodman Fellowship from MGH, the Donovan Family Foundation, a grant from the NHLBI (T32HL007208, to Dr. Peloso), an investigator-initiated research grant from Merck (to Dr. Kathiresan), a grant from Fondation Leducq (to Dr. Kathiresan), a grant from the Swedish Heart–Lung Foundation (2020389, to the Swedish First Myocardial Infarction in AC County 3 [FIA3] Study), and a grant from Regionen Emilia-Romagna (RFPS-2007-3-644882, to the Italian Atherosclerosis, Thrombosis, and Vascular Biology Study).

We thank the site investigators, field staff, and study participants for their contributions to this study. A complete list of acknowledgements is provided in the Supplementary Appendix.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.
REFERENCES


Copyright © 2014 Massachusetts Medical Society.