Whole-exome sequencing identifies rare and low-frequency coding variants associated with LDL cholesterol.

Abstract:

Elevated low-density lipoprotein cholesterol (LDL-C) is a treatable, heritable risk factor for cardiovascular disease. Genome-wide association studies (GWASs) have identified 157 variants associated with lipid levels but are not well suited to assess the impact of rare and low-frequency variants. To determine whether rare or low-frequency coding variants are associated with LDL-C, we exome sequenced 2,005 individuals, including 554 individuals selected for extreme LDL-C (>98\textsuperscript{th} or <2\textsuperscript{nd} percentile). Follow-up analyses included sequencing of 1,302 additional individuals and genotype-based analysis of 52,221 individuals. We observed significant evidence of association between LDL-C and the burden of rare or low-frequency variants in PNPLA5, encoding a phospholipase-domain-containing protein, and both known and previously unidentified variants in PCSK9, LDLR and APOB, three known lipid-related genes. The effect sizes for the burden of rare variants for each associated gene were substantially higher than those observed for individual SNPs identified from GWASs. We replicated the PNPLA5 signal in an independent large-scale sequencing study of 2,084 individuals. In conclusion, this large whole-exome-sequencing study for LDL-C identified a gene not known to be implicated in LDL-C and provides unique insight into the design and analysis of similar experiments.