Clinical and Population Studies

Common SNPs in HMGCR in Micronesians and Whites Associated With LDL-Cholesterol Levels Affect Alternative Splicing of Exon13

Ralph Burkhardt, Eimear E. Kenny, Jennifer K. Lowe, Andrew Birkeland, Rebecca Josowitz, Martha Noel, Jacqueline Salit, Julian B. Maller, Itsik Pe‘er, Mark J. Daly, David Altshuler, Markus Stoffel, Jeffrey M. Friedman, Jan L. Breslow

Background—Variation in LDL-cholesterol (LDL-C) among individuals is a complex genetic trait involving multiple genes and gene–environment interactions.

Methods and Results—In a genome-wide association study (GWAS) to identify genetic variants influencing LDL-C in an isolated population from Kosrae, we observed associations for SNPs in the gene encoding 3-hydroxy-3-methylglutaryl (HMG)-coenzyme A (CoA) reductase (HMGCR). Three of these SNPs (rs7703051, rs12654264, and rs3846663) met the statistical threshold of genome-wide significance when combined with data from the Diabetes Genetics Initiative GWAS. We followed up the association results and identified a functional SNP in intron13 (rs3846662), which was in linkage disequilibrium with the SNPs of genome-wide significance and affected alternative splicing of HMGCR mRNA.

In vitro studies in human lymphoblastoid cells demonstrated that homozygosity for the rs3846662 minor allele was associated with up to 2.2-fold lower expression of alternatively spliced HMGCR mRNA lacking exon13, and minigene transfection assays confirmed that allele status at rs3846662 directly modulated alternative splicing of HMGCR exon13 (42.9±3.9 versus 63.7±1.0% Δexon13/total HMGCR mRNA, P=0.02). Further, the alternative splice variant could not restore HMGCR activity when expressed in HMGCR deficient UT-2 cells.

Conclusion—We identified variants in HMGCR that are associated with LDL-C across populations and affect alternative splicing of HMGCR exon13. (Arterioscler Thromb Vasc Biol. 2008;28:2078-2084)

Key Words: HMG-CoA reductase ■ SNP ■ genome-wide association study ■ LDL-C ■ alternative splicing

Elevated levels of LDL-cholesterol (LDL-C) are a primary risk factor for atherosclerotic cardiovascular disease, the major cause of morbidity and mortality in industrialized countries today. Variation in LDL-C among individuals is a complex genetic trait, involving multiple genes and significant gene–environment interactions. Candidate gene and linkage studies have identified some of the genetic factors contributing to the population variance in plasma lipoprotein levels, but these factors only explain a small fraction of the heritability, suggesting that additional variants influencing lipid levels remain to be identified.

Our group has previously used candidate gene and linkage approaches to identify genetic loci affecting plasma lipid and lipoprotein levels in a cohort from the Island of Kosrae, Federated States of Micronesia, a genetic isolate with significant founder effects and a high prevalence of traits related to the metabolic syndrome.

Recently, genome-wide association studies (GWAS) have been shown to be successful in gene discovery for complex traits and offer a new approach to identify common genetic variants with modest effects. Using Affymetrix gene chip 500k arrays, we have performed a GWAS in ~2400 Kosraens for LDL-C and other metabolic traits (Lowe et al, submitted). The strongest association for LDL-C was found for 2 SNPs on chromosome 19q13 near APOE, a candidate gene with known common coding polymorphisms. The second best hits for LDL-C were SNPs that mapped to the HMG-CoA reductase (HMGCR) gene, the rate-limiting enzyme in cholesterol biosynthesis, and the target of LDL-C lowering statin drugs. During the preparation of this manuscript, 4 GWAS for LDL-C in white cohorts were published, and 1 of them included an association between the same SNPs in HMGCR and plasma LDL-C in whites.
Identifying a strong and replicable association signal in a GWAS is just the first step in elucidating the specific genetic variants involved in predisposing individuals to complex traits. In most cases, the underlying causal variant is not directly captured on the SNP array. Rather, one or more SNPs on the array are acting as a proxy for a functional noncoding-type SNP with which it is in linkage disequilibrium. Fine-mapping, resequencing, and hypothesis-driven approaches have been proposed to unearth the actual causal variants.

In this article, we report associations of SNPs in HMGCR with LDL-C in Kosraens that, in combination with similar findings from studies in whites, indicate that the same genetic variants at HMGCR contribute to differences in LDL-C across populations. To follow-up the association signal, we have implemented a hypothesis-driven strategy and performed in vitro studies to identify a functional variant in intron13 of HMGCR that affects alternative splicing of exon13.

### Methods

An expanded methods section is provided in the supplemental Methods (please see http://atvb.ahajournals.org).

### Kosrae Study Subjects and GWAS

A genome-wide association study for LDL-C was carried out in 2346 people of the Island of Kosrae, Federated States of Micronesia using SNPs from the Affymetrix 500k platform. Details of this study will be described elsewhere (Lowe et al, submitted). All participants in the study provided written informed consent, and IRB approval was obtained from all participating institutions.

We used publicly available data from Saxena et al14 to validate our findings. The probability values of the Kosrae and DGI study were combined using Fisher’s method15 to quantify the overall evidence for association. We set 5 values.

### Cell Lines

Lymphoblastoid cell lines from 18 white CEU individuals of the HapMap collection were obtained from the Coriell Institute for Medical Research. These individuals were homozygous for the chromosome 5q13 containing the HMGCR gene (Table). Identities of the minigenes were confirmed by DNA sequencing.

### Quantitative Real-Time Polymerase Chain Reaction

RNA was extracted using TRIzol reagent (Invitrogen) and cDNA was synthesized using SuperScript III (Invitrogen). Quantitative RT-PCR was performed in an ABI PRISM 7500 Sequence Detector (Applied Biosystems). Specific primers and probes for full-length HMGCR, Δexon13 HMGCR, and -actin mRNA were selected to span exon junctions to avoid coamplification of genomic DNA. mRNA expression studies were carried out at a density of 250,000 cells/ml medium.

### Stable Expressing Cell Lines

The open reading frames of HMGCR full-length and Δexon13 mRNA were PCR amplified and cloned into the pcDNA3.1 expression vector (Invitrogen). UT-2 cells with stable expression of human full-length (UT-2+FL) or Δexon13 HMGCR (UT-2+ex13) were generated by G418 selection of FuGene6 transfected UT-2 cells.

### Other Assays

Cell protein was measured using the BioRad DC kit. HMG-CoA reductase activity was measured in detergent solubilized cell extracts as described,17 except that mevalonolactone was separated by ion exchange chromatography.18

### Results

**Common Variants in HMGCR Are Associated With Plasma LDL-C Across Populations**

We performed genome-wide association analysis for LDL-C using 2346 individuals from the Micronesian Island of Kosrae. The strongest association was found for SNPs on chromosome 19q13 in the APOE/C1/C4/C2 gene cluster (rs4420638, \( P = 1.89 \times 10^{-7} \)). The effect of the APOE polymorphisms on LDL-C is well established in many ethnicities, and our result thus implicates that the same genetic factors are important in the Kosraen population.

The second best locus for LDL-C mapped to a region on chromosome 5q13 containing the HMGCR gene (Table). This gene encodes HMG-CoA reductase, the rate-liming enzyme in cholesterol biosynthesis and thus represents an interesting candidate gene with high plausibility. The difference in LDL-C between homozygotes at this locus was 0.30 mmol/L (11.6 mg/dL), and the fraction of the population variance for LDL-C explained by this locus was 2.1%.

However, the results did not surpass genome-wide significance when the Bonferroni correction was applied for multiple testing (most associated SNP rs3846663: \( P = 1.28 \times 10^{-6} \)). To validate our findings in an independent cohort, we combined data from our analysis with public data for LDL-C of the Diabetes Genetics Initiative (DGI) GWAS. Because Kosraens (Micronesian) and DGI participants (European whites) are of different ancestry, we first investigated linkage disequilibrium (LD) patterns in a 1-Mb region around the locus. As shown in supplemental Figure I, pairwise LD (\( r^2 \)) was
symmetrical between HapMap samples of European ancestry and Micronesians from Kosrae.

Combining probability values across both studies, we validated multiple SNPs at the HMGCR locus, 3 of which surpassed a genome-wide significance of \( P < 5 \times 10^{-8} \) (rs7703051, rs12654264, and rs3846663; Table and supplemental Figure II). A regional association plot for the combined results showed a peak of association signal over a 47-kb region containing the HMGCR gene (supplemental Figure II). In both study populations, the minor allele frequencies were comparable and the minor alleles were associated with an increase in LDL-C.

We also combined the probability values of the association analysis results for plasma total cholesterol from the Kosrae and DGI studies and found genome-wide significance for the same three SNPs at the HMGCR locus (supplemental Table I).

Association Between SNP rs3846662 and Alternative Splicing of HMGCR Exon13 in Human Lymphoblastoid Cell Lines

To follow-up the association results, we next aimed to discover functional variants at the HMGCR locus and study their molecular mode of action. We used existing resequencing data of the region containing the entire HMGCR gene from 23 whites—estimated to have >99% power to detect variants with a minor allele frequency of >5%—to identify candidate functional SNPs. Under our hypothesis-driven model we focused on SNPs that would have strong potential for changing HMGCR function or levels.

All associated SNPs of the Kosrae and DGI studies were noncoding variants and in LD (\( r^2 \approx 0.81 \) to 1), suggesting they represent the same association. The lack of LD (\( r^2 \leq 0.02 \)) between these SNPs and the only known nonsynonymous SNP in HMGCR, rs5908 (I638V) in exon15, suggested that the association was not attributable to this protein-coding mutation (Figure 1). Because the existence of a second HMGCR mRNA transcript resulting from alternative splicing had been reported in humans,20 we looked for SNPs in the vicinity of exon-intron borders. We detected that SNP rs3846662 was located 47 bp downstream of exon13 and in LD with the genotyped variants of genome-wide significance (\( r^2:0.82 \) to 0.93, Figure 1). We hypothesized that this intronic SNP may be functional and modulate the splicing efficiency of exon13.

To analyze whether rs3846662 was associated with HMGCR splicing efficiency we obtained lymphoblastoid cell lines (LCLs) from white CEU individuals of the HapMap collection, who were either homozygous for the major (rs3846662/AA) or minor allele (rs3846662/GG). Likewise, these individuals were also homozygous for the 3 proxy SNPs with genome-wide significance in the association meta-analysis (rs7703051, rs12654264, and rs3846663), where these SNPs denote a haplotype. LCLs (n=9 per genotype) were seeded in medium supplemented with 10% fetal calf serum, and mRNA was harvested at various time points (0 hour, 10 hours, 24 hours, and 48 hours) for expression analysis. We first analyzed whether total HMGCR mRNA expression differed between both groups. As shown in Figure 2A, we did not detect significant differences in total HMGCR mRNA levels between the 2 groups at any time point, indicating that total HMGCR mRNA expression is not influenced by allele status at these SNPs.

We went on to determine the amounts of full-length and alternatively spliced HMGCR (Δexon13) mRNA separately. Δexon13 HMGCR mRNA was detectable in all samples and showed significant variation along the time course (Figure 2B). We observed a distinct decrease in the percentage of Δexon13 HMGCR mRNA in both genotype groups over the first 10 hours. However, the decrease in percentage of...
Alternative splicing of HMGCR mRNA leads to an in-frame deletion of 53 amino acids in the catalytic domain of the protein. To investigate the effect of this deletion on enzyme activity, we stably expressed human full-length (UT-2 FL) and Δexon13 (UT-2 +ex13) HMGCR variants at comparable levels in UT-2 cells (Figure 5A), a CHO cell-line that lacks HMGCR activity and requires exogenous mevalonate for growth.16 UT-2 +FL cells displayed 51% HMGCR enzyme activity of wild-type CHO cells, whereas UT-2 +ex13 cells lacked enzyme activity and were indistinguishable from control UT-2 cells (Figure 5B). Further, UT-2 +FL cells grew in the absence of mevalonate, whereas UT-2 +ex13 and parental UT-2 cells died without mevalonate supplementation (Figure 5C), suggesting that the Δexon13 HMGCR variant is unable to restore enzyme activity in these cells.

**Discussion**

We identified variants in the HMGCR gene that were among our top hits for LDL-C in a GWAS in a population from the Island of Kosrae. We then conducted in vitro studies to follow-up the association signals and to identify a functional variant at the HMGCR locus. We present evidence that a common intronic SNP (rs384662) that is in linkage disequilibrium with the variants typed in the genome scan modulates alternative splicing of HMGCR mRNA. The resulting splicing
variant could not restore enzyme activity when expressed in HMGCR deficient UT-2 cells.

HMG-CoA reductase is a key enzyme in cholesterol homeostasis and catalyzes the rate-limiting step in cholesterol biosynthesis. In contrast to other well known determinants of cholesterol homeostasis, eg, LDL-receptor or Apolipoprotein E, associations between variants in HMGCR and LDL-C have only recently emerged in the context of GWAS. As in the Kosrae study, the initial results of the DGI GWAS in 2758 whites supported associations between SNPs in HMGCR and LDL-C but did not meet the statistical threshold of genome-wide significance (best associated SNP rs12654264: \( P = 4.09 \times 10^{-10} \)). In this study, genome-wide significance was clearly established for HMGCR SNP rs12654264 after validation in 3 additional white cohorts, resulting in a combined probability value of \( 1 \times 10^{-20} \) in a total of \( \approx 18,000 \) subjects. However, in a separate study, the associations between SNPs in HMGCR and LDL-C but did not meet the statistical threshold of genome-wide significance by themselves (best associated SNP rs12654264: \( P = 4.09 \times 10^{-10} \)). In this study, genome-wide significance was clearly established for HMGCR SNP rs12654264 after validation in 3 additional white cohorts, resulting in a combined probability value of \( 1 \times 10^{-20} \) in a total of \( \approx 18,000 \) subjects. However, in a separate study, the associations between SNPs in HMGCR and LDL-C that were observed in the DGI study were not strengthened by a meta analysis approach, consisting of the DGI and 2 other GWAS in whites (best associated SNP rs3846663: \( P = 2.79 \times 10^{-4} \)). This discrepancy might be attributable to some source of heterogeneity, eg, differences in sample ascertainment or the impact of nonadditive interactions with other genetic variants or unaccounted environmental exposures. Combining the association results from the Kosrae and DGI studies revealed 3 variants in LD \( (r^2 > 0.81) \) with genome-wide significance at the HMGCR locus, including the 2 SNPs mentioned above (rs12654264, rs3846663) and SNP rs7703051. Our data obtained in the Kosrae isolate thereby adds important evi-

dence about the generalizability of genetic associations at the HMGCR locus, demonstrating that these associations also extend to other ancestries. Interestingly, 2 pharmacogenetic studies investigating whether genetic variants in HMGCR influence response to statin therapy demonstrated that common SNP haplotypes in HMGCR contribute to variation in statin response. These haplotypes included the SNPs that were associated with plasma LDL-C in the Kosrae and DGI studies, and it is possible that the same underlying mechanisms contribute to variation in LDL-C levels and variation in statin response.

A major aspect of our study was to follow-up the findings from the GWAS and to identify the putative functional variant at the HMGCR locus. To address this question we used human lymphoblastoid cells from the HapMap CEU collection, which have previously been established as a suitable model to study the regulation of cholesterol biosynthesis in normal subjects and subjects with genetic abnormalities in lipid metabolism. Our efforts were facilitated by a near complete inventory (99%) of all common (>5% minor allele frequency) regional sequence variations, resulting from resequencing of the complete HMGCR locus in 23 whites. Because the only known common coding SNP in HMGCR (rs5908, I638V) is not in LD with any of the genotyped SNPs, we consider it to be unlikely that this variant is responsible for the association signal. Likewise, because we did not detect
significant differences in total HMGCR mRNA expression, we consider it to be unlikely that the causal SNP is located in a regulatory element affecting HMGCR transcription. On the other hand, we provide mutually supportive evidence that a common intronic variant (rs3846662) in LD with the genotyped variants is functional and alters the efficiency of HMGCR exon13 alternative splicing: We could demonstrate that (1) expression levels of alternatively spliced Δexon13 HMGCR mRNA were significantly lower in lymphoblastoid cells from homozygotes for the rs3846662 minor allele, and (2) allele status at rs3846662 directly modulated alternative splicing of HMGCR mRNA in minigene constructs. Further, alternative splicing of HMGCR appeared to be regulated and was present in vivo, as we could detect Δexon13 HMGCR mRNA in all 11 human tissues that we studied.

HMGCR mRNA lacking exon13 was described in a survey of alternative pre-mRNA splicing by Johnson et al., however its function and the underlying mechanisms remain unknown. The regulation of gene splicing in mammalians involves both cis- and trans-factors, which are composed of auxiliary element sequences in the pre mRNA, known as splicing enhancers and silencers and cellular splicing factors which include several protein families. The most likely explanation for the observed differences in HMGCR mRNA splicing between major and minor allele homozygotes at rs3846662 is that this SNP is located in a binding motif for a splice auxiliary protein and allele status changes the binding affinity of this protein.

Homozygosity for the major allele at rs3846662 increased the proportion of HMGCR mRNA lacking exon13. Skipping of exon 13 (159 bp) does not change the reading frame, and the resulting protein lacks 53 amino acids in the catalytic domain. When we stably expressed both HMGCR variants in CHO cells deficient of endogenous HMGCR activity, the Δexon13 variant appeared to be nonfunctional and was not able to restore cell growth in the absence of mevalonate. At present, we can only speculate about the exact underlying mechanisms of this observation. Exon13 encodes parts of the catalytic domain, and it contains the highly conserved sequence element ENVIGX3I/LP which is thought to mediate dimerization of the enzyme’s monomers. Thus, deletion of exon13 could potentially impact the stability of the enzyme, because experiments in which monomeric soluble proteins were fused to the HMGCR membrane domains illustrated that the protein was degraded faster when it was smaller than tetrameric. In addition, exon13 contains the E559 residue which is located at the front of the active site and was proposed to directly participate in the reduction of HMG-CoA by serving as a proton donor to mevaldehyde. Therefore, alternative splicing of HMGCR appears to result in altered enzymatic activity and could also lead to more rapid degradation of the protein. A decrease in HMGCR activity would lead to lower cellular cholesterol synthesis and subsequently a counter-regulatory increase of cholesterol uptake from the plasma via the LDL-receptor pathway to maintain intracellular cholesterol homeostasis. In accordance with this hypothesis the allele at rs3846662 that was causing higher levels of Δexon13 HMGCR mRNA in our in vitro studies was sharing a haplotype with the alleles that were associated with lower LDL-C in the genome-wide association studies.

HMG-CoA reductase activity is subject to multivarient control on transcriptional and posttranscriptional levels, and alternative splicing may be an additional regulatory mechanism. Modulation of alternatively spliced HMGCR mRNA levels could be of pharmacological interest with regard to response to statin therapy or as target for antisense-mediated exon skipping. Recently, an antisense oligonucleotide (AON)-mediated skipping approach related to lowering plasma cholesterol levels was applied by Khoo et al. In their study, AON-mediated exon27 skipping of the Apolipoprotein B transcript specifically lowered the amount of functional ApoB100 protein while maintaining ApoB48 levels.

Therefore, identification of specific factors that regulate HMGCR alternative splicing and elucidating the underlying mechanism may lead to a better understanding of its impact on regulating cellular cholesterol homeostasis and plasma cholesterol levels.

Addendum
While this manuscript was under review another study reported that alternative splicing of HMG-CoA Reductase exon13 is associated with plasma LDL-C response to simvastatin (Medina MW, et al. Circulation. 2008;118:355–362), further supporting the functional significance of HMGCR alternative splicing.

Sources of Funding
Ralph Burkhardt is a fellow of the Deutsche Forschungsgemeinschaft (DFG Bu2263/1-1).

Disclosures
None.

References


Common SNPs in HMGCR in Micronesians and Whites Associated With LDL-Cholesterol Levels Affect Alternative Splicing of Exon13

Ralph Burkhardt, Eimear E. Kenny, Jennifer K. Lowe, Andrew Birkeland, Rebecca Josowitz, Martha Noel, Jacqueline Salit, Julian B. Maller, Itsik Pe'er, Mark J. Daly, David Altshuler, Markus Stoffel, Jeffrey M. Friedman and Jan L. Breslow

Arterioscler Thromb Vasc Biol. 2008;28:2078-2084; originally published online September 18, 2008;
doi: 10.1161/ATVBAHA.108.172288

Arteriosclerosis, Thrombosis, and Vascular Biology is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2008 American Heart Association, Inc. All rights reserved.
Print ISSN: 1079-5642. Online ISSN: 1524-4636

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://atvb.ahajournals.org/content/28/11/2078

Data Supplement (unedited) at:
http://atvb.ahajournals.org/content/suppl/2008/09/19/ATVBAHA.108.172288.DC1

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Arteriosclerosis, Thrombosis, and Vascular Biology can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at: http://www.lww.com/reprints

Subscriptions: Information about subscribing to Arteriosclerosis, Thrombosis, and Vascular Biology is online at: http://atvb.ahajournals.org//subscriptions/
Supplemental Table I: Primary associations and meta-analysis results for total cholesterol

<table>
<thead>
<tr>
<th>SNP</th>
<th>Chr</th>
<th>Position</th>
<th>Locus</th>
<th>A1/A2</th>
<th>MAF</th>
<th>p</th>
<th>beta</th>
<th>A1/A2</th>
<th>MAF</th>
<th>p</th>
<th>beta</th>
<th>p-comb</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs11957260</td>
<td>5</td>
<td>74644488</td>
<td>hmgcr</td>
<td>T/C</td>
<td>0.41</td>
<td>6.70x10^-5</td>
<td>0.16</td>
<td>T/C</td>
<td>0.43</td>
<td>1.62x10^-3</td>
<td>0.09</td>
<td>1.08x10^-7</td>
</tr>
<tr>
<td>rs7703051</td>
<td>5</td>
<td>74661243</td>
<td>hmgcr</td>
<td>A/C</td>
<td>0.41</td>
<td>6.15x10^-5</td>
<td>0.16</td>
<td>A/C</td>
<td>0.41</td>
<td>1.43x10^-4</td>
<td>0.11</td>
<td>8.78x10^-9</td>
</tr>
<tr>
<td>rs12654264</td>
<td>5</td>
<td>74684359</td>
<td>hmgcr</td>
<td>T/A</td>
<td>0.42</td>
<td>9.96x10^-5</td>
<td>0.16</td>
<td>T/A</td>
<td>0.41</td>
<td>1.82x10^-4</td>
<td>0.11</td>
<td>1.81x10^-8</td>
</tr>
<tr>
<td>rs3846663</td>
<td>5</td>
<td>74691482</td>
<td>hmgcr</td>
<td>T/C</td>
<td>0.39</td>
<td>5.46x10^-6</td>
<td>0.19</td>
<td>T/C</td>
<td>0.41</td>
<td>1.89x10^-4</td>
<td>0.11</td>
<td>1.03x10^-9</td>
</tr>
</tbody>
</table>

**SNP**: dbSNP ID, **Chr**: chromosome, **Position**: NCBI build 36, **Locus**: closest gene, **A1/A2**: allele 1(major) and allele2 (minor), **MAF**: minor allele frequency, **p**: p-value, **β**: beta-coefficient represents the proportion of 1 s.d. change in standardized LDL-C residual (mean=0, s.d =1) per copy of the allele modelled (additive), **p-comb**: combined p-value of Kosrae and DGI study according to Fisher’s method.
**Figure I**: Pairwise linkage disequilibrium ($r^2$) in HapMap CEPH individuals (top) and Kosraens (bottom) on chromosome 5q13 at 74-75 Mb.
Figure II: Regional association plot for chromosome 5q13 (74.3-75 Mb). Results of the Kosrae study in red, results of the DGI study in blue and combined results are shown in black. Genes in the region are displayed on top of the graph. Pairwise LD ($r^2$) for Kosraens is displayed at the bottom of the figure.
Figure III: HMGCR mRNA expression in lymphoblastoid cells of major and minor allele homozygotes at rs3846662 cultured in RPMI + 10% lipoprotein deficient serum. Panel (A) displays total HMGCR mRNA expression normalized to β-Actin as housekeeping gene over 48h. Panel (B) displays the amount of alternatively spliced HMGCR mRNA (Δex13 mRNA) as percentage of total HMGCR mRNA expression.
Supplemental Methods:

Kosrae Study Subjects and GWAS

A genome-wide association study of 13 quantitative traits was carried out in a cohort of the Island of Kosrae, Federated States of Micronesia. Details of this study will be described elsewhere (Lowe et al, manuscript in preparation). In brief, genotyping of 2906 individuals was performed on the Affymetrix 500k GeneChip. Of the 2906 participants, 2346 fell into groups of sibships and unrelateds which were used for an association analysis performed using the QFAM algorithm in PLINK 1. Nominal scores were permuted to control for relatedness and any remaining overdispersion was adjusted for by genomic control correction 2. All participants in the study provided written informed consent and IRB approval was obtained from all participating institutions.

We used data from Saxena et al. 3 which is publicly available for download to validate our findings. The p-values of the Kosrae and DGI study analyses were combined using Fisher's method 4 to quantify the overall evidence for association. We set $5 \times 10^{-8}$ as threshold for genome-wide statistical significance of the combined p-values.

Cell lines

Lymphoblastoid cell lines from 18 Caucasian CEU individuals of the HapMap collection were obtained from the Coriell Institute for Medical Research. These individuals were homozygous for the major (n=9) or minor allele (n=9) at rs3846662 and the three proxy SNPs from the Affymetrix 500k array (rs3846663, rs7703051 and rs12654264). Cells were cultured at 37°C in a 5% CO₂ environment using RPMI1640 cell culture medium supplemented with 10% fetal calf serum, penicillin (100 U/ml) and streptomycin (100 µg/ml). For HMGCR mRNA expression studies cells were seeded at a density of 250 000 cells/ml medium.
HEK293 cells were grown DMEM supplemented with 10% fetal calf serum, penicillin (100 U/ml) and streptomycin (100 µg/ml).

UT-2 cells are mutant Chinese hamster ovary (CHO) cells that lack HMG-CoA reductase and require mevalonate for growth. Stock cultures of UT-2 cells were grown in medium A (F12:MEM (1:1), penicillin (100 U/ml), streptomycin (100 µg/ml), 0.2 mM mevalonate, 10% fetal calf serum). Stock cultures of UT-2+FL, UT-2+ex13 and CHO-K1 cells were grown under the same conditions as UT-2 cells.

Growth experiments were carried out as described by Mosley et al. In brief, on day 0, 5x10³ cells were seeded in 24-well plates in 1ml of medium A per well. On days 1, 3 and 5 the cells were fed medium A either with or without 0.2mM mevalonate. On day 7 the plates were fixed and stained in a 1% crystal violet/ 20% methanol solution.

For experiments in which HMGCR activity was determined, CHO-K1, UT-2, UT-2+FL and UT-2+ex13 cells were seeded in 60mm dishes in 4 ml medium A and grown over 3 days.

**Quantitative Real-Time PCR**

RNA was extracted using TRIzol reagent (Invitrogen). cDNA was synthesized from 2µg of RNA using SuperScript III (Invitrogen) and random hexamer primers. Quantitative RT-PCR was performed in an ABI PRISM 7700 Sequence Detector (PE Applied Biosystems) to determine the amount of full-length and Δexon13 HMGCR transcripts. Specific TaqMan primers and probes for full-length HMGCR, Δexon13 HMGCR and β-actin mRNA were selected to span exon junctions to avoid co-amplification of genomic DNA. To determine absolute mRNA copy numbers, standard curves were generated for each gene using plasmid dilution series containing the target sequences. The HMGCR PCR reactions were prepared in a final volume of 25 µL of a reaction mixture containing 5 µL cDNA (10 ng/µL), 2.5 mM MgCl₂, 2.5 µL 10 x AmpliTaq buffer A, 200 µM dNTP (each), 0.625 U AmpliTaq Gold (PE Applied Biosystems), 10% DMSO, 200 nM oligonucleotide probe, and 300 nM of each oligonucleotide primer. HMGCR Forward primer (5'
GCACGTGGAAGACGCACAA-3') and TaqMan Probe (5’FAM-
TGGGCCACGAGCTCATCCCATCTG-3’TAMRA) were identical in HMGCR full-length and
Δexon13 mRNA PCRs. The reverse primer was sequence specific for either the full-length (5’-
GAGGCTGCAGAGCAATAGGTCTT-3’) or Δexon13 (5’-
CAGGGATTATAATTACTCCTTGCTTG-3’) HMGCR transcript. The cycling conditions for both
HMGCR TaqMan PCRs were 95°C for 10 minutes and 40 two-step cycles of 95°C for 15
seconds and 63°C for 1 minute.

Beta Actin was used as housekeeping gene to standardize HMGCR mRNA expression. The
human beta Actin PCR reactions were prepared in a final volume of 25 µL of a reaction mixture
containing 5 µL of cDNA (10 ng/µL), 5 mM MgCl₂, 2.5 µL 10 x AmpliTaq buffer A, 200 µM dNTP
(each), 0.625 U AmpliTaq Gold (PE Applied Biosystems), 200 nM oligonucleotide probe (5’FAM-
ATCAAGATCATTGCTCTCTGAGCGCA-3’TAMRA), and 300 nM of each oligonucleotide
primer (BA_fwd: 5’-CCTGGCACCCAGCACAAT-3’, BA_rev: 5’-
GCCGATCCACACGGAGTACTT-3’). The cycling conditions for the HMGCR TaqMan PCR were
95°C for 10 minutes and 40 two-step cycles of 95°C for 15 seconds and 60°C for 1 minute.
Analysis of the data was performed with the ABI PRISM sequence detection software.

**Minigene Constructs**

Minigenes were used to assess the influence of SNP rs3846662 on alternative splicing of
HMGCR exon13. HMGCR minigenes were generated for both alleles from the genomic DNA of
rs3846662/AA and rs3846662/GG individuals. HMGCR genomic DNA fragments containing
exons12-14, as well as the internal introns and parts of the flanking introns, were PCR amplified
with Primers HMGCR-Xho1-F: 5’-GAGACTCGAGGTGTATTAGTGCAAGCCTTGC-3’ and
HMGCR-BamH1-R: 5’- GAGAGGATCCCTGTATTGTGCCTCTTTAGCAGG-3’ (1406 bp product).
Xho1 and BamH1 digested PCR products were cloned into the exon trapping vector pSPL3.
Identities of the minigenes were confirmed by DNA sequencing. To confirm the causal
polymorphism of the observed effects in the minigene systems, we converted SNP rs3846662 in
the major form construct (allele A) to the minor form (allele G) by site directed mutagenesis
(QuikChange Lightning Kit, Stratagene).

Cell Transfection
HEK293 cells were grown to 80% confluency and transfected (four replicates each) with
HMGCR pSPL3 minigenes (A, G and A→G) and empty pSPL3 vector using FuGene6 reagent
(Roche Applied Bioscience). After 24h, RNA was isolated and reverse transcribed using the SA2
Primer (5'-ATCTCAGTGGTATTTGTGAGC-3'), corresponding to a transcribed exonic sequence
in the pSPL3 vector and thus allowing analysis of only vector-specific HMGCR transcripts.
HMGCR splicing pattern was analyzed by RealTime PCR as described above. RNA from cells
transfected with empty pSPL3 vector served as negative control.

Stable expression of full-length and Δexon13 HMGCR variants in UT-2 cells:
The complete full-length (FL) and Δexon13 (Δex13) HMGCR open reading frames were PCR
amplified from lymphocyte cDNA using Platinum Pfx proofreading Taq (Invitrogen) and primers
HMG_ORF_F: 5’-CACCATGTTGTCAAGACTTTTTCG-3’ and HMG_ORF_R: 5’-
TCAGGCTGTCTTCTTGGTGC-3’). The resulting PCR products were separated by gel-
electrophoresis, gel-extracted (Qiagen gel-extraction kit) and cloned into pcDNA3.1 expression
vectors according to the manufacturer’s protocol (Invitrogen). Identities of the clones were
verified by DNA sequencing. UT-2 cells were transfected with either FL-pcDNA3.1 or Δex13-
pcDNA3.1 vectors using FuGene6. The pcDNA3.1 vector contains the neomycin resistance
gene to allow selection of stable cell lines using G418. Three days after transfection cells were
grown in medium A (see above) and in the presence of 750 µg/ml G418 to select for stable
expressing cells. After 4 weeks individual colonies were picked, expanded and screened for
expression of human full-length (FL) and Δexon13 (Δex13) HMGCR, respectively. The resulting
cells were designated UT-2+FL (expressing human full-length HMGCR) and UT-2+ex13 (expressing human ∆exon 13 HMGCR), respectively. mRNA expression levels were quantified by real-time PCR and comparable between the two selected clones used in this study.

**Other assays**

Cell protein was measured using the BioRad DC kit. HMG-CoA reductase activity was measured in detergent solubilized cell extracts as described\(^6\) except that mevalonolactone was separated by ion exchange chromatography\(^7\).
References:

1. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559-575.


