Lnc-ing Common Polymorphisms to Statin Responsiveness at the MYLIP Locus

Sumeet A. Khetarpal, MS; John S. Millar, PhD

3′-Hydroxymethylglutaryl coenzyme A reductase inhibitors, commonly known as statins, are among the most widely prescribed drugs in the world. Indeed, in the United States at least one quarter of adults over the age of 40 years is prescribed a statin.1 Recent evidence suggests that wider prescription of statins to those with even moderate risk of cardiovascular disease may be greatly beneficial in primary risk prevention.2 Statins are highly efficacious in lowering low-density lipoprotein cholesterol (LDL-C) levels by inhibiting 3-hydroxymethylglutaryl coenzyme A reductase, the rate-limiting enzyme in the cholesterol biosynthetic pathway, resulting in increased transcription of hepatic LDL receptor levels.1 Yet despite their immense success, statin use is still limited, in part, by variable responsiveness with regard to LDL-C lowering and side effects, which include risk of developing insulin resistance, myopathy, cognitive impairment, and other morbidities.1 These concerns have prompted extensive studies focused on identifying the precise patients likely to benefit most from statins or as those at risk of developing adverse effects.

Pharmacogenomics studies have been used to identify patients who would benefit the most from statins or are at risk of side effects. These studies relate natural genetic variants and associated genes and pathways to efficacy and tolerability of drugs at the individual level. At a basic level, these studies compare genotype frequencies for polymorphisms at individual genes or across multiple genomic loci among individuals with differing responsiveness for a given drug and identify those single-nucleotide polymorphisms (SNPs) whose allele frequencies differ across the drug-responsive groups. Pharmacogenomics approaches have been applied widely to identify genes that influence the LDL-C-lowering efficacy of many statins as well as their variable adverse effect profiles.3 Such efforts have been performed at times to compare the SNP–drug interactions across different statins and in different ethnicities, and have often focused on either candidate genes thought to influence statin pharmacokinetics or those that influence the dynamics of the cholesterol biosynthetic pathway directly targeted by these medications.

Studies of both candidate genes and genome-wide association studies (GWAS) surveying SNPs across the genome for relationships to statin efficacy have identified several genes involved in statin uptake by the liver and efflux of statins through biliary excretion. These studies have implicated common SNPs in the key statin uptake transporter gene SLCO1B1 as well as several ATP-binding cassette transporter genes, such as ABCB1, ABCC2, and ABCG2, which are proposed to affect biliary excretion of statins, as modulators of LDL-C lowering in response to at least one of the individual statins. These and other studies have also identified SNPs at genomic loci putatively involved in cholesterol biosynthesis and hepatic lipid metabolism that may modulate the pharmacodynamics of statin efficacy. Among these loci are the HMGCR gene, whose protein product, 3-hydroxymethylglutaryl coenzyme A reductase, is the direct target of statins, and other genes regulating the LDL receptor pathway of LDL-C clearance, such as LDLR and PCSK9. Additional loci that have been implicated in statin responsiveness through GWAS studies include APOE, LPA, the SORT1/CERS2/PSRC1 locus, and DNAJC5B.3

Other studies have identified polymorphisms associated with the adverse effects of statins, namely statin-associated myopathy. GWAS approaches studying this adverse effect have consistently shown the SLCO1B1 gene again as a significant contributor.3,4 Additional candidate gene and GWAS have also implicated the coenzyme Q2 gene, COQ2, and the GATM gene, which encodes the rate-limiting enzyme in creatine synthesis, glycine aminotransferase. However, the latter 2 examples have not been confirmed in subsequent replication efforts.3

Statin pharmacogenomics studies have met with the central limitation of difficulty in ascertaining the causal SNP, causal gene, and mechanism of action linking the identified genetic signals to the observed differences in drug responsiveness. Indeed, this is a key hurdle shared by many in the post-GWAS era who seek to uncover causal genes and new biology from the wealth of novel loci borne from the last decade of SNP association studies. In the case of statin pharmacogenomics, although some of the identified loci associated with statin responsiveness underlie genes with plausible roles in regulating statin pharmacokinetics or downstream pathways related to hepatic lipid metabolism, empirical data demonstrating precisely which genes are regulated by a particular SNP and how they are related is largely lacking. Recent efforts to identify causal genes at GWAS-associated loci for various traits have suggested that the identified SNPs from these studies may regulate genes as far as 500 kb away, thus cautioning against the current convention of focusing on the genes
nearest to the associated SNPs. Moreover, as our understanding of noncoding genes and the complexity of cis- and trans-regulatory control of gene expression grows, it becomes ever more important to consider these mechanisms when trying to pinpoint how a polymorphism mediates its effect on a trait, such as statin responsiveness.

A report by Mitchel et al in this issue of Circulation Cardiovascular Genetics highlights this complexity in an effort to identify the mechanism governing the association of a SNP, rs6924995, at the MYLIP gene locus to statin responsiveness in LDL-C lowering. The authors explored the mechanism governing the sub-GWAS significant association of a SNP identified in a previous GWAS for statin effects on LDL-C lowering by the Justification for the Use of Statins in Primary Prevention: An Intervention Trial Evaluating Rosuvastatin (JUPITER) investigators. This SNP was appealing to study for a role in statin responsiveness because it occurs $\approx 10$ kb downstream of MYLIP, a gene encoding IDOL (Inducible Degrader of LDL receptor), a known regulator of LDL receptor turnover. IDOL is a liver-X-receptor (LXR) responsive E3 ubiquitin ligase and both noncoding and coding polymorphisms in MYLIP have been associated with LDL-C levels in humans. One study in 156 Brazilian participants with heterozygous familial hypercholesterolemia even suggested that a coding polymorphism in MYLIP underlies responsiveness to statin therapy in these patients, offering additional support to the notion that MYLIP may underlie the original trending association of the downstream SNP with statin responsiveness in the JUPITER study. However, the SNP itself lies within a processed pseudogene, RP1-13D10.2. Given that some such pseudogenes are more recently being revealed as encoding long noncoding RNAs (lncRNAs), Mitchell et al sought to address the possibility that RP1-13D10.2 is indeed an lncRNA and that it may directly mediate an effect conferred by the SNP on LDL-C levels in response to statins.

Through a combination of studies in lymphoblastoid cell lines derived from patients with previously identified differences in statin responsiveness and other experiments in hepatocyte cell lines, the authors offer evidence suggesting that RP1-13D10.2 is indeed an lncRNA that regulates LDLR gene expression in a sterol-responsive and SNP genotype-dependent manner in vitro. First, the authors demonstrate that RP1-13D10.2 encodes a transcript detectable in lymphoblastoid cell lines from 36 statin high responders versus 39 low responders and show that differences in RP1-13D10.2 expression were observed between the 2 groups, whereas no detectable differences were observed for MYLIP gene expression. Having shown that RP1-13D10.2 encodes an lncRNA, the authors next asked if this transcript regulates expression of genes involved in cholesterol metabolism in both HuH7 and HepG2 hepatocytes through overexpressing RP1-13D10.2 in these cell lines. They found that RP1-13D10.2 increased LDLR gene expression levels but did not affect the expression of any other genes involved in cholesterol biosynthesis or LDL-C uptake that they tested, including MYLIP. Importantly, they compared the effects of the 2 alleles of the rs6924995 SNP on LDLR expression modulating by overexpressing RP1-13D10.2 harboring either allele and comparing the impact on LDLR expression and LDL uptake in vitro. This showed that the A allele of the SNP conferred moderately higher LDLR transcript levels and a trend toward increased LDL-C uptake in vitro. Finally, the authors show through experiments in HuH7 cells that RP1-13D10.2 expression levels seem to be sterol responsive, with sterol depletion through SREBF2 knockdown conferring reduced expression and LXR activation causing a modest increase in expression.

The study by Mitchel et al provides an example of the complexity of mechanisms by which noncoding variants may regulate gene expression and associate with multifactorial traits, such as drug responsiveness and plasma lipids. Here, a SNP was identified in relatively close proximity ($\approx 10$ kb) to a gene with plausible and demonstrated roles in influencing the associated trait. And yet, the authors provide compelling evidence that the newly identified lncRNA RP1-13D10.2 at least, in part, is responsible for the association and further suggest that the established LDL-C gene at the locus, MYLIP, is not regulated in a SNP allele–dependent manner. Although not definitive, the study is suggestive and opens up the possibility of LDL receptor response to statin therapy being regulated, in part, by noncoding RNAs. This work follows several other examples of noncoding RNAs such as miRNAs that regulate both cholesterol efflux and LDL-C uptake in vivo, such as miR-33, miR-148a, and others, as well as fewer examples of lncRNAs that have been shown to regulate lipoprotein metabolism.

Although the study is promising, there are some questions that will need to be addressed in future efforts. One critical area is the validation and replication of the original association of this SNP. The locus studied here for statin responsiveness was originally identified in an analysis of 6989 patients from the JUPITER study, which treated patients with relatively low-baseline LDL-C levels ($<130$ mg/dL). In that study, this locus was not genome-wide significant as a statin-responsiveness locus but trended toward significance ($<1\times10^{-4}$). Although one might have predicted this trend to yield a significant association in a more highly powered study, a subsequent GWAS meta-analysis of $\approx 18500$ participants for statin responsiveness did not report any significant association for this locus while confirming several other previously reported associations. The authors suggest that the discrepant findings may be because of differences in study design, with the JUPITER study involving patients with low-baseline LDL-C levels but increased plasma C-reactive protein levels indicative of elevated inflammation and the more recent study by Postmus et al comprising a range of baseline LDL-C levels as part of a meta-analysis. However, this was not explored further and is of importance in further establishing this locus and putatively novel lncRNA as a mediator of this trait.

Furthermore, the potential that the SNP may also influence MYLIP expression in addition to that of RP1-13D10.2 must remain considered. The studies by Mitchel et al suggest that MYLIP is unlikely to be regulated by the rs6924995 SNP because of the lack of finding of a difference in MYLIP gene expression in the lymphoblastoid cell lines from high versus low responders to statins among the cell lines tested. However,

Khetarpal and Millar
given that statins primarily act to regulate LDL-C levels in the liver in humans, a human hepatocyte cell line would be a more appropriate model for evaluating the impact of the SNP on the regulation of expression of genes that may be involved in determining the statin response. This is even more important given the finding that the expression of RP1-13D10.2 seems to be responsive to sterol depletion and LXR activation, as these perturbations often impart large changes to metabolic transcriptional programs in the liver. An ideal experiment would be to attempt to measure allelic-specific expression\(^2\) for the expression of RP1-13D10.2, MYLIP, and any other genes at the MYLIP locus in human hepatocytes from participants heterozygous for the rs6924995 SNP by RNA sequencing and correlation of the transcript levels for any of the genes on chromosome 6 derived from the A allele versus the G allele for this SNP. Such a measure would allow the direct interrogation of the impact of the SNP genotype of all the genes at the locus and determine whether it is RP1-13D10.2 alone or additional genes such as MYLIP that are influenced by the SNP endogenously.

Another point that will need to be clarified is whether the associated SNP encodes a protein of low abundance or is truly a lncRNA. The authors here posit that RP1-13D10.2 is indeed a lncRNA based on indirect support from 3 in silico tools that suggested that the RP1-13D10.2 sequence was unlikely to encode a protein and the absence of any effect of cyclohexamide-mediated inhibition of protein synthesis on the \(\text{LDLR}\) upregulation conferred by RP1-13D10.2 overexpression. These attempts are suggestive but not conclusive evidence that RP1-13D10.2 indeed encodes a lncRNA and that the observed findings are directly because of the noncoding transcript itself. Additional efforts to evaluate the functions of RP1-13D10.2 will undoubtedly help decipher its putative role as a functional noncoding RNA. For example, studies to determine the interaction of RP1-13D10.2 with the genome, such as through fluorescence in situ hybridization, CRISPR-based approaches to couple the lncRNA to guide RNAs localizing to target genomic regions, and locked nucleic acid oligonucleotides to disrupt the interaction of a lncRNA with putative target DNA, could all aid in pinpointing a role for the noncoding transcript itself in regulating gene expression.\(^14\) Such approaches could start by evaluating the direct interaction of the RP1-13D10.2 transcript with the \(\text{LDLR}\) gene itself as means to confirm that the proposed relationship reported here is directly related to \(\text{LDLR}\) transcriptional modulation. The nature of the regulation of \(\text{LDLR}\) by RP1-13D10.2 is especially important to consider further given that the experiments here involved massive (10\(^{10}\) fold) overexpression of the lncRNA in the hepatocyte lines. Additional efforts to study the possible genome-wide functions of RP1-13D10.2, such as the identification of protein interactors of the lncRNA by RNA immunoprecipitation, unbiased transcriptomics to reveal additional genes regulated by RP1-13D10.2, and higher throughput methods for discovering RNA-bound DNA and proteins, such as ChIRP-Seq,\(^15\) would subsequently confirm the function of RP1-13D10.2 as a lncRNA and may also provide compelling additional evidence about its potentially myriad functions.

If RP1-13D10.2 is ultimately determined to be a lncRNA that acts as a guide molecule to modulate LDL receptor expression this could have several clinical implications. Polymorphisms at this locus could be used in combination with those for other known genetic determinants to predict the overall LDL-C response to statin therapy. Because transcription of this lncRNA seems to be modulated by both SREBP2 and LXR, it is presumably sensitive to conditions in which intracellular cholesterol levels are altered and may also be useful in predicting the LDL response to lifestyle interventions as well, although that remains to be determined. Indeed, the direct consequences of RP1-13D10.2 must be further explored such as through knockdown or gene targeting in animal models and study of the impact on LDL-C metabolism and statin responsiveness in vivo to determine whether RP1-13D10.2 is a physiologically relevant contributor to the observed association of the SNP.

Another exciting possibility would be the development of a therapeutic based on this lncRNA. This study showed that increased RP1-13D10.2 expression was associated with a greater LDL-C reduction in response to statins in vitro. Therefore, it seems reasonable that a homolog to this lncRNA could conceivably be used to promote LDL receptor expression independent of statin treatment. When given as a therapeutic, this lncRNA would be expected to increase LDL receptor expression by promoting transcription of the \(\text{LDLR}\) independent of SREBP2 or LXR activation. This would provide a way of promoting LDL receptor transcription without inhibiting cholesterol synthesis and thus could be used in patients who are unable to tolerate statins because of side effects associated with reduced cholesterol synthesis.

Thus, although much work remains to be done to link the newly identified lncRNA RP1-13D10.2 to not only statin responsiveness but LDL-C biology, the current work by Mitchel et al\(^6\) provides provocative evidence for yet another noncoding RNA in mediating hepatic cholesterol homeostasis. Furthermore, the study here shows another example of the complexity of tackling GWAS loci for complex traits and honing in on the causal genes and mechanisms governing the genetic associations. As tools for both annotation of lncRNAs increase alongside methods for studying their functions in regulating gene expression, additional examples will undoubtedly surface about the relationship of common SNPs to noncoding RNAs related to disease traits that will hopefully better our understanding of the regulatory mechanisms but also pinpoint new biomarkers and putative drug targets.

**Sources of Funding**

S. Khetarpal is supported by NIH grant F30 HL124967.

**Disclosures**

None.

**References**


Key Words: Editorials ◼ cardiovascular disease ◼ genotype ◼ hydroxymethylglutaryl-CoA reductase inhibitors ◼ LDL receptor ◼ low-density lipoprotein cholesterol
Lnc-ing Common Polymorphisms to Statin Responsiveness at the MYLIP Locus
Šumeet A. Khetarpal and John S. Millar

doi: 10.1161/CIRCGENETICS.116.001458
Circulation: Cardiovascular Genetics is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2016 American Heart Association, Inc. All rights reserved.
Print ISSN: 1942-325X. Online ISSN: 1942-3268

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circgenetics.ahajournals.org/content/9/3/206

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation: Cardiovascular Genetics 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 Circulation: Cardiovascular Genetics is online at:
http://circgenetics.ahajournals.org/subscriptions/