Curr Atheroscler Rep (2011) 13:190–192
DOI 10.1007/s11883-011-0168-1

CLINICAL TRIAL REPORT

Genome-Wide Association Studies Complemented with Mechanistic Biological Studies Identify Sortilin 1 as a Novel Regulator of Cholesterol Trafficking

Ali J. Marian

Published online: 2 February 2011

Musunuru K, Strong A, Frank-Kamenetsky A, Lee NE, Ahfeldt T, Sachs KV, et al. From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. Nature 2010;466:714–9.

Rating: ••Of major importance.

Introduction: Genome-wide association studies (GWAS) have shown a strong association between a locus on chromosome 1p13 and both plasma low-density lipoprotein cholesterol (LDL-C) and myocardial infarction (MI) in humans.

Aims: The aim of this study was to identify the responsible molecular mechanisms.

Methods: The methods include population genetic studies to test association of the rs12740374 single nucleotide polymorphism (SNP) with plasma LDL-C levels; functional studies to detect the effects of the locus SNP on expression levels of the genes located at the 1p13 locus; assessment of the effects of the nucleotide change on a C/EBPα binding site and promoter activity; and in vivo genetic manipulation in mice to test effects of over-expression or suppression of expression of SORT1 on plasma LDL-C levels.

Results: The minor allele of the rs12740374 SNP was associated with lower plasma LDL-C levels. The allele created a C/EBP (CCAAT/enhancer binding protein) transcription factor binding site, which was associated with enhanced hepatic expression of the SORT1 gene. Over-expression and suppression of expression of Sort 1 in mouse liver led to lower and higher plasma LDL-C and very low-density lipoprotein (VLDL) levels, respectively.

The findings provide functional evidence for a novel regulatory pathway for lipoprotein metabolism and suggest that modulation of this pathway may alter risk for MI in humans.

Discussion: Through a series of studies in human cohorts, mice, and hepatocytes, the data provide strong evidence that a single noncoding DNA variant at the chromosome 1p13 locus influences LDL-C and MI risk. The responsible mechanism is liver-specific transcriptional regulation of the SORT1 gene by C/EBP transcription factors. The clinical importance of this novel pathway is defined by the 40% difference in MI risk between those homozygous for the major and minor alleles of the p13 locus. The effect is comparable to those of common variants of *LDLR* and *PCSK9* and larger than the effects of common variants in *HMGCR*. The 1p13 minor allele frequency is about 30% in Europeans and is also common in other ethnicities, including African Americans, Hispanics, Asian Indians, and Chinese. Therefore, this locus is an important global genetic determinant of MI risk. In addition, among lipid-regulating genes related to MI, SORT1 is unique for having been identified by GWAS mapping of common DNA variants, rather than by discovery of rare gene variants underlying Mendelian disorders.

A. J. Marian (✉)
Center for Cardiovascular Genetics,
The Brown Foundation Institute of Molecular Medicine,
The University of Texas Health Science Center at Houston,
6770 Bertner Street, DAC 900,
Houston, TX 77030, USA
e-mail: Ali.J.Marian@uth.tmc.edu

A. J. Marian
Texas Heart Institute at St. Luke's Episcopal Hospital,
Houston, TX 77030, USA

Comments

The findings illustrate the potential power of the GWAS, when complemented with mechanistic studies, in elucidation of functional effects of the DNA sequence variants and identification of new pathways for the pathogenesis of complex phenotypes [1]. The main strength of the present study is its multi-tier approach, which is exemplary for genetic association studies. Accordingly, the authors replicated the findings of the previous GWAS studies in two independent populations. Then they analyzed association of the rs646776 SNP located at this locus with the mRNA levels of *SARS*, *CELSR2*, *PSRC1*, *MYBPHL*, *SORT1*, *PSMA5*, and *SYPL2*, which are also located at the 1p13 locus. They found that the minor allele was associated with increased expression levels of *CELSR2*, *PSRC1,* and *SORT1* in human liver tissues. The strongest association was with SORT1 mRNA and protein. To identify the culprit SNP, the authors fine-mapped the region, defined the haplotypes, and analyzed the functional significance of the variants using a reporter assay, which led to identification of the rs12740374 SNP as a candidate SNP. The major allele of this SNP, when included in the haplotype, had the largest effect size on the transcriptional activity of the haplotype. Bioinformatics analysis showed that the minor allele created a C/EBP-α binding site by changing the wild-type sequence G**G**TGCTCAAT to G**T**TGCTCAAT on the *SORT1* promoter. Binding of the C/EBP-α to DNA sequence at this site increased the promoter activity and expression level of SORT1. To consolidate the findings, the authors over-expressed Sort1 in live mice using recombinant adeno-associated viruses and showed it reduced plasma LDL-C levels. Conversely, suppression of expression of Sort1 in the liver led to increased plasma LDL-C levels. Thus, the authors validated the effects of SORT1 on plasma LDL-C at multiple levels.

The findings are also in accord with the results of previous genetic studies that have reported an association between SNPs at the 1p13 locus and plasma cholesterol levels and atherosclerosis [2–4]. The minor allele of the rs599839 SNP in *SORT1* gene was associated with a decrease of serum LDL-C by 0.14 mmol/L and a 9% decreased risk of coronary atherosclerosis [3]. Despite the strength of the genetic studies, there is uncertainty at the mechanistic level. A recent study showed that Sort1 interacts with apolipoprotein B100 (apoB100) in the Golgi apparatus and facilitates formation and hepatic export of apoB100-containing lipoproteins [5]. Consequently, in contrast to the findings by Musunuru et al. [1], over-expression of Sort 1 stimulated hepatic release of lipoproteins and led to increased plasma LDL levels [5]. Thus, the genetic data provide robust evidence for the causal role of SORT1 in regulating plasma LDL-C. However, the responsible mechanism(s) remain to be established.

The above discoveries are remarkable, as the previous functional studies of the GWAS results typically have faced formidable challenges in identifying new mechanisms. The demanding task of identifying the responsible mechanism (s) of allelic association studies is not only evident for SORT1 but also for the 9p21 locus, which in multiple independent studies has been linked to atherosclerosis and its complications [6–8]. The locus regulates expression levels of the cell cycle regulators *Cdkn2a* and *Cdkn2b*, which are tumor suppressor genes and markers of cell senescence [9]. Deletion of the orthologous region of 9p21 locus in the mouse genome is associated with significant reductions in the expression levels of *Cdkn2a* and *Cdkn2b* [10]. In accord with this finding, smooth muscle cells isolated from the 9p21-null mice showed excessive proliferation and reduced cell senescence [10]. These findings are insightful but perhaps to some extent intuitive, as they implicate biological aging as the potential mechanism for atherosclerosis. In contrast, functional studies of the 1p31 locus have been impressive in identifying SORT1 as a novel pathway in cholesterol trafficking, despite the exact mechanisms being completely elucidated.

GWAS, by genotyping the relatively common alleles, have the inherent limitation of identifying alleles that are typically expected to exert minimal or negligible effects on the phenotype [11]. By and large, the common alleles appear to explain only a small fraction of the phenotype [12]. Hence, a significant component of the heritability of the complex phenotypes such as atherosclerosis remains unaccounted for, which is often referred to as "the missing heritability" or "the dark matter of heritability." The proponents of GWAS contend that studies with larger sample sizes and a higher SNP density could lead to identification of the "missing heritability" by detecting alleles with very small effect sizes. Conversely, there is an increasing interest in direct DNA sequencing as a means of identifying rare alleles that exert large effect sizes [13]. Advances in DNA sequencing technologies, such as the next-generation sequencing platforms, afford the opportunity to sequence the whole exome or even the whole genome of every individual and identify rare and novel variants. Considering that each exome has approximately 10,000 non-synonymous variants and probably about 200–300 novel ns-synonymous variants, deep sequencing is likely to dominate genetic studies of complex phenotype in the coming years [14–18]. Whole genome sequencing also has the advantage of identifying non-coding sequences, including the regulatory variants with significant effect sizes, as was in the case of SORT1. However, "the dark matter of heritability" is likely due to complex interactions between various genetic, genomic, and epigenetic factors [12]. Likewise, the phenotype is the consequence of non-linear and stochastic interactions between different genetic and non-genetic constituents, including environmental factors.

Discovery of any new genetic variant or new molecular pathway naturally raises the interest in exploiting the discovery for therapeutic interventions. Sortilin 1 is a biologically plausible therapeutic target as it is involved in the transport and processing of LDL-C in the liver cells. However, the contrasting results of two mechanistic studies do not provide guidance on how to pursue interventions [1, 5]. Interventions aimed to increase SORT1 level might be more challenging than targeting its suppression, but both are potentially achievable. The next level of challenge is the potential fortuitous effect of interventions to enhance or suppress expression of SORT1, as this protein is diffusely expressed and is a multi-ligand receptor with functions in lysosomal trafficking in various organs, including the brain [19]. Therefore, exploiting the new discovery of SORT1 pathway for therapeutic gain could prove difficult but not necessarily impossible.

In conclusion, the study reviewed here provides multiple lines of evidence that establish SORT1 as a novel regulator of plasma LDL-C levels and hence, as a potential target for therapeutic intervention for hypercholesterolemia and atherosclerosis.

Disclosure

A.J. Marian has received grants from the National Institutes of Health and the Burroughs Wellcome Fund.

References

1. Musunuru K, Strong A, Frank-Kamenetsky M, Lee NE, Ahfeldt T, Sachs KV, et al. From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. Nature 466(7307):714–9.
2. Kleber ME, Renner W, Grammer TB, Linsel-Nitschke P, Boehm BO, Winkelmann BR, et al. Association of the single nucleotide polymorphism rs599839 in the vicinity of the sortilin 1 gene with LDL and triglyceride metabolism, coronary heart disease and myocardial infarction. The Ludwigshafen Risk and Cardiovascular Health Study. Atherosclerosis 209 (2):492–7.
3. Linsel-Nitschke P, Heeren J, Aherrahrou Z, Bruse P, Gieger C, Illig T, et al. Genetic variation at chromosome 1p13.3 affects sortilin mRNA expression, cellular LDL-uptake and serum LDL levels which translates to the risk of coronary artery disease. Atherosclerosis 208(1):183–9.
4. Kathiresan S, Voight BF, Purcell S, Musunuru K, Ardissino D, Mannucci PM, et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. Nat Genet. 2009;41(3):334–41.
5. Kjolby M, Andersen OM, Breiderhoff T, Fjorback AW, Pedersen KM, Madsen P, et al. Sort1, encoded by the cardiovascular risk locus 1p13.3, is a regulator of hepatic lipoprotein export. Cell Metab. 12(3):213-3.
6. McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, Cox DR, et al. A common allele on chromosome 9 associated with coronary heart disease. Science. 2007;316(5830):1488–91.
7. Helgadottir A, Thorleifsson G, Manolescu A, Gretarsdottir S, Blondal T, Jonasdottir A, et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. Science. 2007;316(5830):1491–3.
8. Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, et al. Genomewide association analysis of coronary artery disease. N Engl J Med. 2007;357(5):443–53.
9. Collado M, Blasco MA, Serrano M. Cellular senescence in cancer and aging. Cell. 2007;130(2):223–33.
10. Visel A, Zhu Y, May D, Afzal V, Gong E, Attanasio C, et al. Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. Nature 464(7287):409–12.
11. Marian AJ. Nature's genetic gradients and the clinical phenotype. Circ Cardiovasc Genet. 2009;2(6):537–9.
12. Eichler EE, Flint J, Gibson G, Kong A, Leal SM, Moore JH, et al. Missing heritability and strategies for finding the underlying causes of complex disease. Nat Rev Genet. 11(6):446–50.
13. Cirulli ET, Goldstein DB. Uncovering the roles of rare variants in common disease through whole-genome sequencing. Nat Rev Genet. 11(6):415–25.
14. Wang J, Wang W, Li R, Li Y, Tian G, Goodman L, et al. The diploid genome sequence of an Asian individual. Nature. 2008;456(7218):60–5.
15. Levy S, Sutton G, Ng PC, Feuk L, Halpern AL, Walenz BP, et al. The diploid genome sequence of an individual human. PLoSBiol. 2007;5(10):e254.
16. Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, et al. The complete genome of an individual by massively parallel DNA sequencing. Nature. 2008;452(7189):872–6.
17. Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW, Lee C, et al. Targeted capture and massively parallel sequencing of 12 human exomes. Nature. 2009;461(7261):272–6.
18. Ng PC, Levy S, Huang J, Stockwell TB, Walenz BP, Li K, et al. Genetic variation in an individual human exome. PLoS Genet. 2008;4(8):e1000160.
19. Wahe A, Kasmapour B, Schmaderer C, Liebl D, Sandhoff K, Nykjaer A, et al. Golgi-to-phagosome transport of acid sphingomyelinase and prosaposin is mediated by sortilin. J Cell Sci. 123 (Pt 14):2502–11.