Published in final edited form as:
Curr Opin Cardiol. 2010 May ; 25(3): 238–242. doi:10.1097/HCO.0b013e328338574d.

Genetics of Lipid Disorders

James Pirruccello[1,2,3] and **Sekar Kathiresan, MD**[1,2,4,5,*]

[1]Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts 02114, USA

[2]Broad Institute, Cambridge, Massachusetts 02142, USA

[3]Johns Hopkins University School of Medicine, Baltimore, Maryland 21287, USA

[4]Cardiovascular Research Center, Massachusetts General Hospital, Boston, Massachusetts 02114, USA

[5]Department of Medicine, Harvard Medical School, Boston, Massachusetts 02115, USA

Abstract

Purpose of review—In this review, we will highlight recent advances in identifying genes and gene regions responsible for the variation in serum lipid levels. We will also consider the next directions for research based on these advances.

Recent findings—Large-scale genome-wide association studies have successfully screened common variants across the genome for association with serum lipids and have generated novel hypotheses about the causes of serum lipid variation.

Summary—Deep sequencing of GWA signals promises to expand the catalog of variants responsible for serum lipid variation and with a full catalog of variants, we may develop a panel of polymorphisms with clinical utility. In parallel, functional exploration of the GWA signals should expand our knowledge of lipoprotein metabolism and generate targets for pharmacologic intervention.

Keywords

genetics; cholesterol; genome-wide association; lipids; triglycerides

Introduction

Disorders of lipoprotein metabolism lead to atherosclerosis and cardiovascular disease (CVD), including myocardial infarction and stroke. Sixteen million adults in the United States alone live with coronary artery disease, which kills 450,000 people annually [1]. Because of the well-established relationship between serum lipid levels and CVD, serum lipids have long been of clinical interest. In particular, LDL-C is recognized not only as a biomarker of CVD risk, but also as a causal participant in the disease process.

Serum lipid levels including LDL-C, high-density lipoprotein cholesterol (HDL-C) and triglycerides (TG) are highly heritable: studies consistently estimate that over 50% of the total inter-individual variation in serum lipid levels can be explained by the genetic variation

*Author to whom correspondence should be addressed. skathiresan@partners.org. Phone: (617) 643-6120.

The authors declare that they have no conflict of interest to report.

[2]. Therefore, there is substantial interest in discovering the genetic determinants of lipid levels, particularly genes that alter both serum lipids and CVD risk.

Mendelian lipoprotein disorders

Most of the well-characterized disorders of lipoprotein metabolism are monogenic, familial disorders with extreme phenotypes amenable to linkage analysis. For example, proprotein convertase subtilisin/kexin type 9 (*PCKS9*) has recently been implicated in a form of autosomal-dominant hypercholesterolemia (ADH). Its protein product normally associates with hepatic and extrahepatic LDL receptors and appears to target them for endosomal degradation [3]. Gain-of-function variants (such as D374Y) lead to a 10-fold reduction in LDL receptor levels, causing increased LDL-C levels and resulting in ADH [4,5].

Various Mendelian disorders of lipoprotein metabolism have now been described for each of the three lipoprotein traits, and up to now they have contributed the bulk of our knowledge about lipoprotein genetics [6]. Although those who harbor such traits experience profound phenotypes, the population-wide impact of the known Mendelian disorders is attenuated by their rarity. The monogenic lipid disorders discovered thus far can explain little of overall lipoprotein heritability. Indeed, the normal distribution of serum lipoprotein levels in the population suggests a polygenic model of inheritance: LDL-C, HDL-C, and TG levels are determined by the additive contribution of multiple loci.

Common variants

In contrast to the linkage studies that were well suited to identifying rare variants with profound effects, genome-wide association (GWA) studies have been designed to ascertain whether common genetic variants contribute to population-wide lipid variability. In such studies, hundreds of thousands of single-nucleotide polymorphisms (SNPs) are interrogated for association with serum lipid levels in large cohorts. GWA has become a viable technique for studying lipoprotein genetics on a large scale because serum lipids are accurately measured, lipid phenotypes are readily available in many cohorts, and genotyping hundreds of thousands of polymorphisms has become relatively inexpensive.

Each GWA study has shared similar approaches. Hundreds of thousands of SNPs are genotyped in a discovery cohort, with follow-up genotyping performed on a subset of those SNPs in a replication cohort. With linear regression, each allele is correlated to the serum level of each lipoprotein trait. Only genetic loci passing genome-wide significance ($P < 5 \times 10^{-8}$) are reported; the success of this stringent criterion in minimizing false associations is underscored by the fact that, in each successively larger GWA study, no previously-significant loci have dropped from significance.

The first large-scale lipids GWA study, published in 2007, involved 2,800 individuals genotyped at nearly 400,000 SNPs [7]. This study re-discovered associations known from prior genetic studies: that of HDL-C with CETP, and of LDL-C with APOE, which effectively served as positive controls that confirmed GWA as a valid approach to identifying lipid-associated loci. Progressively larger studies have been conducted; the most recent lipids GWA studies from 2009 involved up to 40,000 individuals of European descent and identified more than 30 chromosomal loci with common variants associated with lipid levels [8**,9**]. An in-progress study by the Global Lipids Genetics Consortium involving over 100,000 individuals will likely identify still more loci.

The lipids GWA studies have led to the observation that genes with rare Mendelian mutations of large effect size also harbor common variants of more modest effect sizes [8**, 9**] (Table 1). These findings prompt the reciprocal hypothesis: many lipid-associated loci

mapped using common variants will also harbor novel low-frequency (1% to 5% frequency) and rare (<1% frequency) variants with more pronounced effects.

Already, 30 common lipid-associated variants have explained 9.3%, 7.7%, and 7.4% of inter-individual variability for HDL-C, LDL-C, and TG, respectively [8**]. These variants, therefore, explain ~15–18% of heritability – a proportion that likely will increase with publication of the 100,000-individual lipids GWA study. Much of the remaining heritability may ultimately be attributable to low-frequency and rare variants.

Next steps after genetic mapping using common variants

GWA has provided a flood of common genetic variants that are strongly associated with lipoprotein levels and have begun to explain substantial fractions of heritability. The tasks now at hand are to narrow the scope of localization from the chromosomal region to the level of the functional element (gene or regulatory element) and to explain the molecular mechanisms that ultimately lead to the observed lipid phenotypes. DNA sequencing and functional biology will be instrumental to these ends.

Sequencing

Sequencing offers the promise of identifying the variants driving the GWA signals. Because of the underlying patterns of linkage disequilibrium, a GWA SNP may only narrow the search to a several-hundred-kilobase region. Targeted, large-scale sequencing in unrelated individuals should ascertain all of the variants in the region and, therefore, to reduce the scope of the functional search down to a focused assortment of dozens of specific variants, one or more of which is likely to have functional consequences.

In some cases, sequencing may even identify smoking-gun variants, such as early truncations, that strongly indicate causation. Studies of *PCSK9* are suggestive in this regard. The gene harbors a common variant (19% minor allele frequency) identified in a 40,000-person lipids GWA study [10]. *PCSK9* has also been shown to contain low-frequency nonsense variants that lead to a 15–28 percent reduction in LDL-C and a 47–88 percent decrease in lifetime risk of CVD [11]. Additionally, the rare gain of function variants, discussed above, implicate *PCSK9* in a form of ADH. If this pattern of common, low-frequency, and rare variants being present at a lipid locus is typical, then sequencing near GWA loci may not only identify the source of the GWA hit, but also a trove of causal low-frequency and rare variants. Nevertheless, even variants that appear likely to be causal (truncations, nonconservative substitutions) will require functional validation.

As the efficiency of sequencing increases and the cost falls, targeted sequencing will give way to whole-exome sequencing, and, in turn, to whole-genome sequencing – far exceeding the association boundaries of the regions initially highlighted by GWA studies. Such enhancement of our catalogue of lipid variants by sequencing may also prove fruitful for improved risk prediction and for the identification of a set of genes enriched for plausible pharmacological targets.

Functional validation

GWA and deep resequencing are highly informative, but fall short of elucidating a functional link between DNA sequence variant and lipoprotein phenotype. To provide convincing evidence of causality, functional studies will be required. Many GWA-identified loci are in linkage disequilibrium blocks spanning one or more genes; such loci immediately suggest protein targets to interrogate, whether through animal models, molecular biology, or emerging techniques. More difficult to study will be the loci that harbor no known genes,

though many will ultimately be discovered to harbor regulatory elements important for modulating the expression of genes in lipoprotein metabolic pathways.

Ultimately, the fulfillment of a genetic Koch's postulates would provide the most convincing evidence of causality, requiring: (a) identification of the SNP that, presumptively, causes variation in a lipid trait; (b) development of a valid model, such as that from induced pluripotent stem cells (iPS) from individuals with and without the variant; (c) highly targeted manipulation of the DNA to replace the common allele with the variant; and (d) observation of a change in phenotype consistent with the proposed directionality and degree of variation (i.e., substituting the variant allele for the common allele and observing a phenotype consistent with that found in those natively harboring the variant allele). Such an approach would be particularly useful in confirming the causality of variants that appear to reside in intergenic "deserts."

Clinical applications

In parallel to the work being done on elucidating the biology underlying the GWA signals, efforts are underway to take advantage of the presently available data for clinical risk prediction and personalized pharmacology.

Genetic risk stratification

There are advantageous theoretical properties to a genetic screen for plasma lipoproteins. Genetic scores are immutable within an individual, which may reduce costs due to decreased repeat testing, and allow for very early detection of individuals at increased risk for CVD. Such scores may better represent the cumulative lifetime burden of exposure to lipids than point-estimates of fasting lipid levels [12].

Because CVD is the ultimate outcome of interest, not lipoprotein levels *per se*, genetic screens may be particularly helpful if not all causes of lipid variation confer equal CVD risk (for example, if certain causes of increased HDL-C are protective, whereas others are not). Such a possibility was brought to the fore by the failure of torcetrapib, the *CETP* inhibitor, to reduce cardiovascular mortality or intermediate endpoints, despite increasing HDL (though this result should not be overinterpreted, as torcetrapib has molecule-specific off-target effects that trigger mineralocorticoid excess) [13–15]. On the other hand, persons with null mutations in *APOC3* experience increased HDL-C, reduced TG, and a reduction in coronary calcification [16*]. This suggests that not all causes of elevated HDL-C carry the same consequences for CVD. Specific variants contributing to the observed serum lipoprotein levels may convey substantial information beyond that encoded by the scalar lipid levels themselves.

Nevertheless, genetic scores have not proven superior to lipoprotein levels and family history for CVD risk discrimination. For example, the use of one risk allele in the well-known 9p21 locus in white women did not improve discrimination or reclassification beyond that achievable using serum lipids, C-reactive protein (CRP), and family history [17*]. In contrast, we have demonstrated a lipid genetic panel using SNPs identified in the pre-GWA era; the panel did not improve risk discrimination but did improve risk reclassification [12]. As more variants are identified through GWA studies and sequencing, expanded panels will be developed that may achieve the goal of improved discrimination. At that point, the clinical utility of such panel should be examined in the context of a clinical trial.

Pharmacogenetics

Whereas genetic risk profiling is used to rule-in those who would ordinarily not be considered for therapy, pharmacogenetics may be used to exclude patients from what would otherwise be the default treatment. Here we will split the subject of lipid pharmacogenetics into two topics: efficacy-oriented pharmacogenetics and toxicity-oriented pharmacogenetics.

Trials have now demonstrated proof-of-principle for efficacy-oriented pharmacogenetic approaches to predicting statin response [18*,19*]. For example, the PROVE-IT TIMI 22 trial examined the effect of SNPs from genes in LDL-C metabolism pathways and statin pharmacokinetics pathways on the response to atorvastatin or pravastatin [18*]. Carriers with apolipoprotein E (*APOE*) isoform ε2 demonstrated greater reduction in LDL-C did than ε4 carriers, and were more likely to reach target LDL-C levels [18*]. In this study, *APOE* isoforms explained 3.8% of the residual variance in response to statin therapy. Nevertheless, statin therapy can already be titrated based on lipid levels at follow-up. Perhaps a pharmacogenetic panel could be built to address whether a failure to reduce LDL-C despite statin therapy was due to nonadherence or, instead, to genetically-encoded resistance, but the clinical utility of such a tool is unclear. Efficacy-oriented pharmacogenetic approaches are unlikely to be clinically compelling in the absence of a trial that demonstrates that an interaction between statin choice and genotype can lead to differential outcomes, even after controlling for lipoprotein levels.

Toxicity-oriented pharmacogenetic approaches may be of clinical use if the consequences of an adverse event are severe – e.g., cause harm *per se* or, indirectly, by inducing reluctance to try any drug in the same class, even one with reduced likelihood of causing a future adverse event. Clinical trials have now examined the association between genotype and statin-induced myopathy. In the SEARCH trial, a *SLCO1B1* allele (*SLCO1B1*5*) raised the risk for myopathy in an allelic dose-dependent fashion (OR = 4.5 for CT heterozygotes vs TT homozygotes, and OR = 16.9 for CC homozygotes vs TT homozygotes) [20**]. Roughly 60% of the population risk of myopathy from taking 80mg simvastatin daily was attributed to this allele; nevertheless, the absolute risk of simvastatin-induced myopathy was modest (0.6% in TT homozygotes, 3% in CT heterozygotes, and 18% in CC homozygotes). Another group found similar relative effect sizes of allelic dose on risk of myopathy in simvastatin users, but not in those taking pravastatin [21*]. Therefore, knowledge of such allelic variants may help guide initial statin therapy in order to reduce the risk of harm and to increase the likelihood of adherence.

Conclusion

GWA studies have successfully screened common variants across the genome for association with serum lipids and have generated novel hypotheses about the causes of lipid variation. Deep sequencing of GWA signals promises to expand the catalog of variants responsible for serum lipid variation and with a full catalog of variants, we may develop a panel of polymorphisms with clinical utility. In parallel, functional exploration of the GWA signals should expand our knowledge of lipoprotein metabolism and generate targets for pharmacologic intervention.

Acknowledgments

The authors would like to thank Kiran Musunuru for his useful comments.

J.P. is a research fellow supported by Sarnoff Cardiovascular Research Foundation; S.K. is supported by NIH RC2 HL101864, NIH RC1 HL099793, and NIH RC1 HL099634.

References

1. Rosamond W, Flegal K, Furie K, et al. Heart Disease and Stroke Statistics--2008 Update: A Report From the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Circulation. 2008 Jan 29; 117(4):e25–146. [PubMed: 18086926]
2. Kathiresan S, Manning A, Demissie S, et al. A genome-wide association study for blood lipid phenotypes in the Framingham Heart Study. BMC Medical Genetics. 2007; 8(Suppl 1):S17. [PubMed: 17903299]
3. Lagace TA, Curtis DE, Garuti R, et al. Secreted PCSK9 decreases the number of LDL receptors in hepatocytes and in livers of parabiotic mice. J Clin Invest. 2006 Nov; 116(11):2995–3005. [PubMed: 17080197]
4. Abifadel M, Varret M, Rabes J, et al. Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. Nat Genet. 2003 Jun; 34(2):154–156. [PubMed: 12730697]
5. Cohen JC, Boerwinkle E, Mosley TH, Hobbs HH. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. N Engl J Med. 2006 Mar 23; 354(12):1264–1272. [PubMed: 16554528]
6. Willard, H. Genomic and personalized medicine. 1. Amsterdam & Boston: Elsevier/Academic Press; 2009.
7. Saxena R, Voight BF, Lyssenko V, et al. Diabetes Genetics Initiative of Broad Institute of Harvard and MIT LU. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science. 2007 Jun 1; 316(5829):1331–1336. [PubMed: 17463246]

**8. Kathiresan S, Willer CJ, Peloso GM, et al. Common variants at 30 loci contribute to polygenic dyslipidemia. Nat Genet. 2009 Jan; 41(1):56–65. Using GWA data from 19,840 individuals and targeted replication in up to 20,623 individuals, compelling statistical evidence is provided for 30 loci associated with LDL cholesterol, HDL cholesterol, and/or triglycerides. [PubMed: 19060906]

9. Aulchenko YS, Ripatti S, Lindqvist I, et al. Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. Nat Genet. 2009 Jan; 41(1):47–55. Using GWA data from 22,562 individuals, compelling statistical evidence is provided for 22 loci associated with LDL cholesterol, HDL cholesterol, and/or triglycerides. Almost all of these loci overlap with the work of Kathiresan et al. [8]. [PubMed: 19060911]

10. Kotowski IK, Pertsemlidis A, Luke A, et al. A Spectrum of PCSK9 Alleles Contributes to Plasma Levels of Low-Density Lipoprotein Cholesterol. The American Journal of Human Genetics. 2006 Mar; 78(3):410–422.
11. Cohen JC, Boerwinkle E, Mosley TH, Hobbs HH. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. N Engl J Med. 2006 Mar 23; 354(12):1264–1272. [PubMed: 16554528]
12. Kathiresan S, Melander O, Anevski D, et al. Polymorphisms associated with cholesterol and risk of cardiovascular events. N Engl J Med. 2008 Mar 20; 358(12):1240–1249. [PubMed: 18354102]
13. Rader DJ. Illuminating HDL – Is it still a viable therapeutic target? N Engl J Med. 2007 Nov 22; 357(21):2180–2183. [PubMed: 17984168]
14. Kastelein JJ, van Leuven SI, Burgess L, et al. Effect of torcetrapib on carotid atherosclerosis in familial hypercholesterolemia. N Engl J Med. 2007 Apr 19; 356(16):1620–1630. [PubMed: 17387131]
15. Nissen SE, Tardif J, Nicholls SJ, et al. Effect of torcetrapib on the progression of coronary atherosclerosis. N Engl J Med. 2007 Mar 29; 356(13):1304–1316. [PubMed: 17387129]

*16. Pollin TI, Damcott CM, Shen H, et al. A null mutation in human APOC3 confers a favorable plasma lipid profile and apparent cardioprotection. Science. 2008 Dec 12; 322(5908):1702–1705. This study demonstrated that a non-sense mutation in apoC-III was associated with an improved lipoprotein profile and decreased coronary artery calcification. [PubMed: 19074352]

*17. Paynter NP, Chasman DI, Buring JE, et al. Cardiovascular disease risk prediction with and without knowledge of genetic variation at chromosome 9p21.3. Ann Intern Med. 2009 Jan 20; 150(2):65–72. This study showed that the use of genetic information for cardiovascular risk

prediction did not offer any advantage over the use of serum lipids, CRP, and family history. [PubMed: 19153409]

*18. Mega JL, Morrow DA, Brown A, et al. Identification of genetic variants associated with response to statin therapy. Arterioscler Thromb Vasc Biol. 2009 Sep 1; 29(9):1310–1315. This genetic analysis from the PROVE-IT TIMI 22 trial demonstrated that specific isoforms of APOE were associated with a differential response to pravastatin or atorvastatin. [PubMed: 19667110]

19. Thompson JF, Hyde CL, Wood LS, et al. Comprehensive whole-genome and candidate gene analysis for response to statin therapy in the Treating to New Targets (TNT) cohort. Circ Cardiovasc Genet. 2009 Apr 1; 2(2):173–181. A GWA of nearly 300,000 SNPs in nearly 2,000 individuals found no genome-wide-significant associations between SNP and statin response. In agreement with the work of Mega et al [18], certain APOE variants were significantly associated with statin response, as well as PCSK9 and HMGCR variants. [PubMed: 20031582]

**20. The SEARCH Collaborative Group. SLCO1B1 variants and statin-induced myopathy – a genomewide study. N Engl J Med. 2008 Aug 21; 359(8):789–799. This study found that 60% of simvastatin-induced myopathy could be attributed to a variant in *SLCO1B1*, which encodes the organic anion transporter OATP1B1 that is responsible for the hepatic uptake of most statins. [PubMed: 18650507]

*21. Voora D, Shah SH, Spasojevic I, et al. The SLCO1B1*5 genetic variant is associated with statin-induced side effects. Journal of the American College of Cardiology. 2009 Oct 20; 54(17):1609–1616. This study found that the effect of *SLCO1B1* variants did not signify the same risk of adverse events for all statins. [PubMed: 19833260]

Table 1

Loci associated with common SNPs in a 40,000-individual GWA study (adapted from [8**])

Trait	Chr.	SNP	P value	Locus	Associated Mendelian Disorders
LDL	1p13	rs12740374	2×10^{-42}	*CELSR2, PSRC1, SORT1*	
LDL	2p24	rs515135	5×10^{-29}	*APOB*	Autosomal dominant hypercholesterolemia
LDL	19q13	rs4420638	4×10^{-27}	*APOE-APOC1- APOC4- APOC2*	Type IB hyperlipoproteinemia (*APOC2*); Type III hyperlipoproteinemia (*APOE*)
LDL	19p13	rs6511720	2×10^{-26}	*LDLR*	Autosomal dominant hypercholesterolemia
LDL	2p21	rs6544713	2×10^{-20}	*ABCG8*	Sitosterolemia
LDL	5q13	rs3846663	8×10^{-12}	*HMGCR*	
LDL	5q23	rs1501908	1×10^{-11}	*TIMD4-HAVCR1*	
LDL	20q12	rs6102059	4×10^{-9}	*MAFB*	
LDL	12q24	rs2650000	2×10^{-8}	*HNF1A*	
LDL	19p13	rs10401969	2×10^{-8}	*NCAN, CILP2, PBX4*	
LDL	1p32	rs11206510	4×10^{-8}	*PCSK9*	Autosomal dominant hypercholesterolemia
HDL	16q13	rs173539	4×10^{-75}	*CETP*	Cholesteryl ester transfer protein deficiency
HDL	8p21	rs12678919	2×10^{-34}	*LPL*	
HDL	15q22	rs10468017	8×10^{-23}	*LIPC*	
HDL	18q21	rs4939883	7×10^{-15}	*LIPG*	
HDL	16q22	rs2271293	9×10^{-13}	*LCAT*	LCAT deficiency (fish-eye disease)
HDL	11q23	rs964184	1×10^{-12}	*APOA1-APOC3- APOA4- APOA5*	Primary hypoalphalipoproteinemia (*APOA1*)
HDL	11q12	rs174547	2×10^{-12}	*FADS1-FADS2- FADS3*	
HDL	12q24	rs2338104	1×10^{-10}	*MMAB, MVK*	
HDL	9p22	rs471364	3×10^{-10}	*TTC39B*	
HDL	20q13	rs1800961	8×10^{-10}	*HNF4A*	
HDL	9q31	rs1883025	1×10^{-9}	*ABCA1*	Tangier disease
HDL	20q13	rs7679	4×10^{-9}	*PLTP*	
HDL	19p13	rs2967605	1×10^{-8}	*ANGPTL4*	
HDL	1q42	rs4846914	4×10^{-8}	*GALNT2*	
TG	11q23	rs964184	4×10^{-62}	*APOA1-APOC3-APOA4- APOA5*	
TG	8p21	rs12678919	2×10^{-41}	*LPL*	Type IA hyperlipoproteinemia

Trait	Chr.	SNP	P value	Locus	Associated Mendelian Disorders
TG	2p23	rs1260326	2×10^{-31}	*GCKR*	
TG	8q24	rs2954029	3×10^{-19}	*TRIB1*	
TG	7q11	rs714052	3×10^{-15}	*MLXIPL*	
TG	11q12	rs174547	2×10^{-14}	*FADS1-FADS2-FADS3*	
TG	2p24	rs7557067	9×10^{-12}	*APOB*	
TG	19p13	rs17216525	4×10^{-11}	*NCAN, CILP2, PBX4*	
TG	20q13	rs7679	7×10^{-11}	*PLTP*	
TG	8p23	rs7819412	3×10^{-8}	*XKR6-AMAC1L2*	