Published in final edited form as:

Curr Opin Lipidol. 2016 April ; 27(2): 99–104. doi:10.1097/MOL.0000000000000277.

What can we learn about lipoprotein metabolism and coronary heart disease from studying rare variants?

Janina M. Jeff¹, Gina M. Peloso², and Ron Do^{1,3,4,5}

¹Charles F. Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY

²Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA

³The Center for Statistical Genetics, Icahn School of Medicine at Mount Sinai, New York, NY

⁴The Zena and Michael A. Weiner Cardiovascular Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA

⁵Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY

Abstract

Purpose of review—Rare variant association studies (RVAS) target the class of genetic variation with frequencies less than 1%. Recently, investigators have used exome sequencing in RVAS to identify rare alleles responsible for Mendelian diseases but have experienced greater difficulty discovering such alleles for complex diseases. In this review, we describe what we have learned about lipoprotein metabolism and coronary heart disease (CHD) through the conduct of RVAS.

Recent findings—Rare protein-altering genetic variation can provide important insights that are not as easily attainable from common variant association studies. First, RVAS can facilitate gene discovery by identifying novel rare protein-altering variants in specific genes that are associated with disease. Second, rare variant associations can provide supportive evidence for putative drug targets for novel therapies. Finally, rare variants can uncover new pathways and reveal new biologic mechanisms.

Summary—The field of human genetics has already made tremendous progress in understanding lipoprotein metabolism and the etiology of CHD in the context of rare variants. As next generation sequencing becomes more cost-effective, RVAS with larger sample sizes will be conducted. This will lead to more novel rare variant discoveries, and the translation of genomic data into biological knowledge and clinical insights for cardiovascular disease.

Corresponding author: Ron Do, Ph.D., The Charles Bronfman Institute for Personalized Medicine, Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Pl., Box 1003, New York, NY 10029, Tel: 212-241-6206, ron.do@mssm.edu.

Conflicts of interest
None.

Keywords

Rare variant association studies; exome sequencing; lipoprotein metabolism; coronary heart disease

INTRODUCTION

The frequency of a variant can be classified into four categories: fixed, common, low frequency, and rare. Fixed variants, also called monomorphic, do not vary among individuals in a population, whereas, common single nucleotide variants (SNVs) [frequency $\geq 5\%$] represents the other end of the allele frequency spectrum that contains the most frequent number of alleles in a SNV [1]. Because of the high frequency of common variants, these SNVs are typically captured using standard genotyping arrays and potential disease loci can be more easily detectable in populations of moderate size [2]. Furthermore, low frequency variants (frequency between 0.5–5%) have been included on most genotyping arrays [1] and imputation of low frequency variants is now possible in most human populations [3]. Unlike low-frequency variants, rare variants (frequency $< 0.5\%$) are seen in one or a few individuals and are optimally detected using high-throughput sequencing approaches.

Protein-altering, or non-synonymous variants, are those SNVs predicted to change the amino acid sequence of a protein. A special class of rare protein-altering variants are those that result in a complete loss of function (LoF). Complete LoF mutations are SNVs predicted to create a premature stop codon (nonsense), disrupt a splice-site and indels predicted to alter the reading frame of a protein (frameshift). These variants are likely to be under strong negative selection and have a strong deleterious impact on protein function [4].

Common variant association studies (CVAS; sometimes referred to as genome-wide association studies (GWAS)), typically target the class of genetic variation with frequencies greater than 1%. These studies rely on genotyping arrays, and have been successful in identifying common genetic variation associated with complex traits and diseases, including lipid metabolism and coronary heart disease (CHD) [5–9]. The first wave of CVAS discovered an association with CHD at the 9p21 locus [5, 6, 8, 9]. Subsequently, larger studies have now identified over >160 loci associated with blood lipids [9–12] and 58 loci associated with CHD [13].

Loci discovered from CVAS can span several megabases and involve dozens of genes (due to linkage disequilibrium). Fine-mapping [11] and functional validation studies [14] have helped refine association signals. However, such studies are time-consuming and in a number of cases, it still remains difficult to localize specific variants and/or genes underlying disease associations.

Rare variant association studies (RVAS), on the other hand, target the class of genetic variation that is generally missed by CVAS; that is, these studies can capture SNVs with frequencies less than 1%. These studies typically rely on sequencing technologies. Recently, several large-scale RVAS studies have been performed using next generation sequencing coupled with targeted capture of the protein-coding regions (“exome sequencing”). To date,

investigators have used exome sequencing to identify rare alleles responsible for Mendelian diseases [15], but have experienced greater difficulty discovering such alleles for complex diseases [16].

RVAS has several advantages over CVAS. In Table 1, we highlight the advantages and disadvantages of RVAS compared to CVAS. Rare protein-altering genetic variation can provide impactful biological and clinical insights that common variants studied in CVAS cannot. First, RVAS can facilitate gene discovery by identifying novel rare protein-altering variants associated with disease. Second, rare variant associations can provide supportive evidence for putative drug targets for novel therapies. Finally, rare variants can uncover new biology involved in disease pathways. In the present review, we discuss each of these impacts that RVAS has had in informing lipoprotein metabolism and CHD.

EXOME SEQUENCING AS A TOOL TO DISCOVER NOVEL GENES

Exome sequencing to discover inherited causes of Mendelian disorders

By studying rare protein-altering variants segregating in families with putative Mendelian dyslipidemias, one can implicate causal variants and genes for disease [15]. An example is a study by Musunuru et al. that examined the exome sequences of two siblings with very low levels of plasma low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C) and triglycerides [17]. Two siblings had both copies of the angiopoietin-like 3 (*ANGPTL3*) gene knocked out (i.e. compound heterozygotes) by two different nonsense mutations, p.S17X and p.E129X, where both mutations affected LDL-C and triglycerides in a manner that was dependent on the number of mutations carried. The *ANGPTL3* protein is secreted by the liver, involved in the regulation of extracellular lipases, and is associated with lipid levels in mice [18]. More recently, a larger study of nine families and 32 unrelated individuals has confirmed the p.S17X variant association with Familial Combined Hypolipidemia (FCH) [19]. FCH causes a global reduction of plasma lipoproteins. Specifically, FCH is characterized by a reduction in apolipoprotein (Apo) B- [very low-density lipoprotein (VLDL) and low-density lipoprotein (LDL)] and apoA-1-containing [high-density lipoprotein (HDL)] lipoproteins [19]. Homozygotes for the p.S17X variant had a reduction in plasma lipid and no circulating *ANGPTL3* [19]. A significant carrier effect is observed for p.S17X variant and *ANGPTL3* levels, with carriers having a 42% reduction in *ANGPTL3* levels compared to non-carriers [19].

Despite this early success, using exome sequencing to discover novel inherited causes of Mendelian disorders can be difficult in instances where inheritance patterns are not recessive. A recent study performed whole exome sequencing on 213 individuals from 41 families with extreme LDL-C or HDL-C levels. The study successfully identified putative pathogenic variants in nine (22%) families, all in known lipid-mediated genes [20]. Although novel variants were not discovered in this study, the study highlighted potential limitations of detecting rare variants in familial settings. These limitations include complex inheritance patterns and the over representation of shared alleles within families, making identification of casual alleles indistinct [20].

Exome sequencing to discover rare mutations associated with complex disease

Recently, large-scale efforts have applied exome sequencing to study rare variants (frequency < 1%) in population-based samples. By studying the extremes of quantitative traits such as LDL-C, Lange et al. showed that carrying rare variants in patatin-like phospholipase domain-containing protein 5 (*PNPLA5*) increased LDL-C by 43.1 mg/dl compared to non-carriers [21*]. This finding was unique in that it was only observed in African Americans. Hence, population-specific signals can be captured by incorporating samples from different ethnicities in the study design of RVAS.

Exome sequencing efforts have also been applied to individuals who present with myocardial infarction (MI) at a young age. Do et al. sequenced more than 9,700 exomes from samples with early-onset MI and MI-free controls. From this effort, rare variant association signals were discovered for early-onset MI in two genes – 1) low-density lipoprotein receptor [*LDLR*] and 2) and apolipoprotein A-V [*APOA5*]. In this study, we demonstrated that by focusing on rare protein-altering variants, one can implicate the causal variant and gene, unlike CVAS, where it's often difficult to point to a creditable set of variants [22]. For example, in this study, *APOA5* was established as a *bona fide* MI gene. Although common variants in the *apolipoprotein A-1* [*APOA1*]-*apolipoprotein C-III* [*APOC3*]-*apolipoprotein A-4* [*APOA4*]-*apolipoprotein A-5* [*APOA5*] (abbreviated: *APOA1-C3-A4-A5*) locus have been associated with plasma triglycerides and MI risk [23–25], the region contains a complex linkage disequilibrium pattern and several established lipid genes including *APOA1*, *APOC3*, and *APOA4*. Therefore, previously, it had been unclear if *APOA5* was a MI gene in humans. Furthermore, this study showed that SNVs predicted to severely disrupt gene function showed the largest magnitude of effect for both *LDLR* and *APOA5*. This includes complete LoF variants as well as the set of variants that are predicted to be deleterious across multiple *in silico* functional prediction algorithms. This study provided proof of principle that exome sequencing can be used to discover rare mutations that contribute to the genetic risk to a complex disease.

SEQUENCING AS A TOOL TO PRIORITIZE MOLECULAR TARGETS IN DRUG DEVELOPMENT

Human genetics has the potential to identify new therapeutic targets and provide evidence for validation of targets with regard to their impact on plasma lipids and MI. A case-in-point is the discovery of the *PCSK9* gene as a regulator of plasma LDL-C. Abifadel et al. studied large pedigrees using classical linkage analysis and discovered that gain-of-function mutations in *PCSK9* led to autosomal dominant familial hypercholesterolemia (FH)[26]. Other studies of *PCSK9* also suggest causality with rare variants in proprotein convertase subtilisin/kexin type 9 preproprotein (*PCSK9*) in population-based settings. Cohen et al showed that individuals with nonsense mutations in *PCSK9* (C679X and Y142X) had much lower total cholesterol and LDL-C compared to those without the mutation [27]. Furthermore, when investigating probands in trios with LDL-C below the 10 percentile, Cohen et al. report one family with the p.Y142X null mutation in which all individuals carrying the substitution had LDL-C levels below the 5th percentile [27]. Based in part on the human genetics evidence, pharmaceutical companies have concluded that targeting

PCSK9 can lead to an effective therapeutic. Recently, the U.S Food and Drug Administration has approved two *PCSK9* inhibitors (Praluent and Repatha) for use in individuals who require additional lowering of LDL-C.

As another example, RVAS have been used to validate the main target effect of the drug ezetimibe. Ezetimibe has been shown to lower plasma levels of LDL-C. However, until recently, it was unclear whether this drug would also lower risk for CHD. Ezetimibe functions by inhibiting the protein product of Neimann-Pick C1 Like 1 (*NPC1L1*), a gene that encodes a protein essential in intestinal cholesterol absorption [28]. By sequencing the exons of 7,364 CHD cases and 14,728 controls, a study showed that heterozygote carriers for complete LoF mutations in *NPC1L1* had 12 mg/dL lower LDL-C compared to non-carriers. Subsequently, carriers were shown to have a 55% decreased risk for CHD [29**]. This finding supports the hypothesis that *NPC1L1* inhibition may decrease CHD risk. Findings from this study were consistent with those from a very recent study, The Improved Reduction of Outcomes: Vytorin Efficacy International Trial (IMPROVE-IT), which showed that ezetimibe added to statin therapy had a stronger effect on CHD risk compared to just statin therapy alone [30].

Finally, an antisense oligonucleotide inhibiting *APOC3* has been developed based on evidence from human genetics that shows that a null mutation in *APOC3*, p.R19X, leads to lower plasma triglycerides. However, it was unknown whether inhibiting *APOC3* would lower CHD risk. A genetic study by Crosby et al. described rare loss-of function and missense mutations in *APOC3* [31**], where carriers of rare *APOC3* mutations had considerably lower serum triglyceride (39%) and circulating APOC3 (43%) levels compared to non-carriers. Carriers of *APOC3* variants also had significantly higher HDL and lower LDL cholesterol. These carriers subsequently had as much as 40% reduced risk of developing CHD [31**]. This work suggests that therapeutically targeting *APOC3* will reduce risk of CHD.

RARE VARIANTS CAN UNCOVER NOVEL BIOLOGY

In addition to prioritizing new targets for drug discovery, such gene discovery can often reveal previously unappreciated biologic mechanisms. With *PCSK9*, an entirely new mechanism for regulation of plasma LDL-C was identified [32, 33]. Similarly, the discovery of ATP-binding cassette sub-family A member 1 (*ABCA1*) as the causal gene for Tangier disease [34], a disorder (originally discovered on Tangier Island off the coast of Virginia) characterized by very low HDL-C, has led to a new field of research worldwide in cholesterol efflux.

Recent RVAS using exome sequencing technology have highlighted specific genetic mechanisms involved in triglyceride (TG)-rich lipoprotein metabolism and the etiology of CHD risk. Plasma TG concentration is a complex polygenic trait that integrates chylomicrons synthesized by the intestine in the postprandial state, very-low-density lipoproteins (VLDL) synthesized by the liver, and remnant lipoproteins of both intestinal and hepatic origin. Epidemiological studies have established associations between plasma TG and increased risk to CHD events [35]; however, until recently, it has been unknown

whether this association represents causal processes. Recent Mendelian randomization studies have suggested that plasma TG [36, 37] and remnant cholesterol are causal risk factors for CHD. However, it has been suggested that this causal relationship may be driven by only a subset of genes regulating TG-rich lipoproteins contribute to CHD risk.

The aforementioned RVAS on *APOC3* and *APOA5* have implicated these genes as specific mechanisms for altering TG-rich lipoproteins and impacting CHD risk. Furthermore, rare complete LoF mutations in the gene encoding lipoprotein lipase (*LPL*) cause chylomicronemia, severe hypertriglyceridemia, very low HDL-C, and premature CHD [38, 39]. In addition, a gain-of-function (GoF) variant in *LPL* is associated with lower TG, higher HDL-C, and reduced CHD risk [40, 41]. These three genes suggest that at least some of the genes in the LPL pathway can decrease CHD risk by altering TG, beyond lowering LDL-C. We note that not all genes in the LPL pathway do not show definitive evidence of altering CHD risk. For example, apolipoprotein C-II is a critical co-factor for the activation of LPL. Furthermore, lipase maturation factor 1 and glycosylphosphatidylinositol anchored high density lipoprotein binding protein 1 play key functional roles in regulating LPL activity [42]. Complete LoF mutations have been observed in each of these three genes and have been shown to lead to deficiency in LPL activity and severe hypertriglyceridemia. It's unclear why some genes in the LPL pathway appear to be influence CHD risk while others do not. One possibility may perhaps be that *APOC3*, *APOA5* and *LPL* may be capturing specific atherogenic processes. Another may be statistical where current RVAS for MI are - even at ~10,000 cases and controls - are statistically underpowered and therefore it remains unclear whether complete LoF mutations in these genes alter CHD risk [16, 43**].

In Table 2, we summarize human genetic evidence for each of the genes in the LPL pathway that show definitive genetic evidence in influencing CHD risk. Common, low-frequency and/or rare variants of genes encoding LPL and its modulating proteins are associated with differences in plasma TG and/or HDL-C, and in some instances, with risk of CHD [31**, 40, 43**, 44]. These human genetic findings highlight the importance of the LPL pathway in triglyceride-rich lipoprotein metabolism, and the etiology of CHD.

FUTURE DIRECTIONS AND CONCLUSIONS

The field has made tremendous progress in the genomics of lipoprotein metabolism in the context of rare variants. Addressing current limitations of exome sequencing such as increasing sample size of RVAS to increase statistical power will accelerate gene discovery efforts. Fortunately, large collaborations and the use of large health care systems with hundreds of thousands of individuals with available electronic medical records can potentially overcome sample size limitations. Furthermore, recruitment and follow-up of carriers of rare mutations can be used to understand biological function using a reverse genetics approach. Unlike traditional forward genetics where a phenotype-genotype relationship is inferred from observing a phenotype first, reverse genetics uses known genotype information to further probe specific physiological processes. Such approaches have primarily been used in non-human model organisms but are now being used to generate genotype-first prospective human cohorts [45]. Much like the progress seen for genetic mapping in CVAS studies [46], we expect that as sequencing becomes more cost-effective

and the scalability of RVAS to very large sample sizes becomes feasible, rare variant discovery of human disease, and subsequent translation of these discoveries to biologic and clinical insights will be accomplished.

Acknowledgments

Financial support and sponsorship

GMP is supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health under Award Number K01HL125751. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Abbreviations

RVAS	rare variant association studies
CHD	coronary heart disease
SNVs	single nucleotide variants
LoF	loss of function
CVAS	common variant association studies
GWAS	genome-wide association studies
ANGPTL3	angiopoietin-like 3
PNPLA5	patatin-like phospholipase domain-containing protein 5
LDLR	low-density lipoprotein receptor
APOA5	apolipoprotein A-V
APOA1	apolipoprotein A-I
APOC3	apolipoprotein C-III
PCSK9	proprotein convertase subtilisin/kexin type 9 preproprotein
NPC1L1	Neimann-Pick C1 Like 1
ABCA1	ATP-binding cassette sub-family A member 1
LDL-C	low-density lipoprotein cholesterol
HDL-C	high-density lipoprotein cholesterol
TG	triglycerides
LPL	lipoprotein lipase

REFERENCES

1. Auton A, Brooks LD, et al. Genomes Project C. A global reference for human genetic variation. *Nature*. 2015; 526:68–74. [PubMed: 26432245]
2. Hardy J, Singleton A. Genomewide association studies and human disease. *The New England journal of medicine*. 2009; 360:1759–1768. [PubMed: 19369657]
3. Huang J, Howie B, McCarthy S, et al. Improved imputation of low-frequency and rare variants using the UK10K haplotype reference panel. *Nature communications*. 2015; 6:8111.

4. Boyko AR, Williamson SH, Indap AR, et al. Assessing the evolutionary impact of amino acid mutations in the human genome. *PLoS genetics*. 2008; 4:e1000083. [PubMed: 18516229]
5. McPherson R, Pertsemlidis A, Kavaslar N, et al. A common allele on chromosome 9 associated with coronary heart disease. *Science*. 2007; 316:1488–1491. [PubMed: 17478681]
6. Helgadottir A, Thorleifsson G, Manolescu A, et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science*. 2007; 316:1491–1493. [PubMed: 17478679]
7. Wellcome Trust Case Control C. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007; 447:661–678. [PubMed: 17554300]
8. Samani NJ, Erdmann J, Hall AS, et al. Genomewide association analysis of coronary artery disease. *The New England journal of medicine*. 2007; 357:443–453. [PubMed: 17634449]
9. Teslovich TM, Musunuru K, Smith AV, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*. 2010; 466:707–713. [PubMed: 20686565]
10. Kathiresan S, Willer CJ, Peloso GM, et al. Common variants at 30 loci contribute to polygenic dyslipidemia. *Nature genetics*. 2009; 41:56–65. [PubMed: 19060906]
11. Willer CJ, Schmidt EM, et al. Global Lipids Genetics C. Discovery and refinement of loci associated with lipid levels. *Nature genetics*. 2013; 45:1274–1283. [PubMed: 24097068]
12. Surakka I, Horikoshi M, Magi R, et al. The impact of low-frequency and rare variants on lipid levels. *Nature genetics*. 2015; 47:589–597. [PubMed: 25961943]
13. Consortium CAD. A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nature genetics*. 2015; 47:1121–1130. [PubMed: 26343387]
14. Musunuru K, Strong A, Frank-Kamenetsky M, et al. From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. *Nature*. 2010; 466:714–719. [PubMed: 20686566]
15. Bamshad MJ, Ng SB, Bigham AW, et al. Exome sequencing as a tool for Mendelian disease gene discovery. *Nature reviews. Genetics*. 2011; 12:745–755. [PubMed: 21946919]
16. Do R, Kathiresan S, Abecasis GR. Exome sequencing and complex disease: practical aspects of rare variant association studies. *Human molecular genetics*. 2012; 21:R1–R9. [PubMed: 22983955]
17. Musunuru K, Pirruccello JP, Do R, et al. Exome sequencing, ANGPTL3 mutations, and familial combined hypolipidemia. *The New England journal of medicine*. 2010; 363:2220–2227. [PubMed: 20942659]
18. Koishi R, Ando Y, Ono M, et al. Angptl3 regulates lipid metabolism in mice. *Nature genetics*. 2002; 30:151–157. [PubMed: 11788823]
19. Minicocci I, Montali A, Robciuc MR, et al. Mutations in the ANGPTL3 gene and familial combined hypolipidemia: a clinical and biochemical characterization. *The Journal of clinical endocrinology and metabolism*. 2012; 97:E1266–E1275. [PubMed: 22659251]
20. Stitzel NO, Peloso GM, Abifadel M, et al. Exome sequencing in suspected monogenic dyslipidemias. *Circulation. Cardiovascular genetics*. 2015; 8:343–350. [PubMed: 25632026]
21. Lange LA, Hu Y, Zhang H, et al. Whole-exome sequencing identifies rare and low-frequency coding variants associated with LDL cholesterol. *American journal of human genetics*. 2014; 94:233–245. [PubMed: 24507775] This is one of the first large-scale exome sequencing studies for LDL-C. Furthermore, the authors perform a wide range of association tests and ultimately discover several signals for a burden of low-frequency coding variants influencing LDL-C in two distinct populations.
22. Hormozdiari F, Kostem E, Kang EY, et al. Identifying causal variants at loci with multiple signals of association. *Genetics*. 2014; 198:497–508. [PubMed: 25104515]
23. Pennacchio LA, Olivier M, Hubacek JA, et al. An apolipoprotein influencing triglycerides in humans and mice revealed by comparative sequencing. *Science*. 2001; 294:169–173. [PubMed: 11588264]
24. Sarwar N, et al. Triglyceride Coronary Disease Genetics C, Emerging Risk Factors C. Triglyceride-mediated pathways and coronary disease: collaborative analysis of 101 studies. *Lancet*. 2010; 375:1634–1639. [PubMed: 20452521]

25. De Caterina R, Talmud PJ, Merlini PA, et al. Strong association of the APOA5-1131T>C gene variant and early-onset acute myocardial infarction. *Atherosclerosis*. 2011; 214:397–403. [PubMed: 21130994]
26. Abifadel M, Varret M, Rabes JP, et al. Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nature genetics*. 2003; 34:154–156. [PubMed: 12730697]
27. Cohen J, Pertsemlidis A, Kotowski IK, et al. Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. *Nature genetics*. 2005; 37:161–165. [PubMed: 15654334]
28. Davis HR Jr, Zhu LJ, Hoos LM, et al. Niemann-Pick C1 Like 1 (NPC1L1) is the intestinal phytosterol and cholesterol transporter and a key modulator of whole-body cholesterol homeostasis. *The Journal of biological chemistry*. 2004; 279:33586–33592. [PubMed: 15173162]
29. Stitzel NO, Won HH, et al. Myocardial Infarction Genetics Consortium I. Inactivating mutations in NPC1L1 and protection from coronary heart disease. *The New England journal of medicine*. 2014; 371:2072–2082. [PubMed: 25390462] This study was the largest study to date to show that a burden of rare complete loss of function mutations in *NPC1L1*, the gene target of ezetimibe, are associated with decreased CHD risk.
30. Cannon CP, Blazing MA, Giugliano RP, et al. Ezetimibe Added to Statin Therapy after Acute Coronary Syndromes. *The New England journal of medicine*. 2015; 372:2387–2397. [PubMed: 26039521]
31. Blood I, et al. Tg, Hdl Working Group of the Exome Sequencing Project NHL. Loss-of-function mutations in APOC3, triglycerides, and coronary disease. *The New England journal of medicine*. 2014; 371:22–31. [PubMed: 24941081] This study is noteworthy as it identifies an association of rare mutations in *APOC3* are associated with decreased plasma triglycerides and decreased CHD risk. This study suggests that therapeutically targeting *APOC3* can reduce CHD risk.
32. Horton JD, Cohen JC, Hobbs HH. Molecular biology of PCSK9: its role in LDL metabolism. *Trends in biochemical sciences*. 2007; 32:71–77. [PubMed: 17215125]
33. Maxwell KN, Fisher EA, Breslow JL. Overexpression of PCSK9 accelerates the degradation of the LDLR in a post-endoplasmic reticulum compartment. *Proceedings of the National Academy of Sciences of the United States of America*. 2005; 102:2069–2074. [PubMed: 15677715]
34. Brooks-Wilson A, Marcil M, Clee SM, et al. Mutations in ABC1 in Tangier disease and familial high-density lipoprotein deficiency. *Nature genetics*. 1999; 22:336–345. [PubMed: 10431236]
35. Sarwar N, Danesh J, Eiriksdottir G, et al. Triglycerides and the risk of coronary heart disease: 10,158 incident cases among 262,525 participants in 29 Western prospective studies. *Circulation*. 2007; 115:450–458. [PubMed: 17190864]
36. Do R, Willer CJ, Schmidt EM, et al. Common variants associated with plasma triglycerides and risk for coronary artery disease. *Nature genetics*. 2013; 45:1345–1352. [PubMed: 24097064]
37. Varbo A, Benn M, Tybjaerg-Hansen A, et al. Remnant cholesterol as a causal risk factor for ischemic heart disease. *Journal of the American College of Cardiology*. 2013; 61:427–436. [PubMed: 23265341]
38. Hoffer MJ, Bredie SJ, Snieder H, et al. Gender-related association between the-93T-->G/D9N haplotype of the lipoprotein lipase gene and elevated lipid levels in familial combined hyperlipidemia. *Atherosclerosis*. 1998; 138:91–99. [PubMed: 9678774]
39. Johansen CT, Wang J, Lanktree MB, et al. Excess of rare variants in genes identified by genome-wide association study of hypertriglyceridemia. *Nature genetics*. 2010; 42:684–687. [PubMed: 20657596]
40. Rip J, Nierman MC, Ross CJ, et al. Lipoprotein lipase S447X: a naturally occurring gain-of-function mutation. *Arteriosclerosis, thrombosis, and vascular biology*. 2006; 26:1236–1245.
41. Reymer PW, Gagne E, Groenemeyer BE, et al. A lipoprotein lipase mutation (Asn291Ser) is associated with reduced HDL cholesterol levels in premature atherosclerosis. *Nature genetics*. 1995; 10:28–34. [PubMed: 7647785]

42. Brahm AJ, Hegele RA. Chylomicronaemia--current diagnosis and future therapies. *Nature reviews. Endocrinology*. 2015; 11:352–362.
43. Do R, Stitzel NO, Won HH, et al. Exome sequencing identifies rare LDLR and APOA5 alleles conferring risk for myocardial infarction. *Nature*. 2015; 518:102–106. [PubMed: 25487149] This work was the first large-scale RVAS study for MI using exome sequencing. The study identified for the first time, a burden of rare mutations in two genes that are associated with increased risk for MI at the population level.
44. Kastelein JJ, Groenemeyer BE, Hallman DM, et al. The Asn9 variant of lipoprotein lipase is associated with the -93G promoter mutation and an increased risk of coronary artery disease. The Regress Study Group. *Clinical genetics*. 1998; 53:27–33. [PubMed: 9550358]
45. Stessman HA, Bernier R, Eichler EE. A genotype-first approach to defining the subtypes of a complex disease. *Cell*. 2014; 156:872–877. [PubMed: 24581488]
46. Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. *Science*. 2008; 322:881–888. [PubMed: 18988837]

Key points

1. Recently, several large-scale rare variation associations studies have been performed using next generation sequencing coupled with targeted capture of the protein-coding regions (“exome sequencing”).
2. We have acquired important biologic and clinical insights into lipoprotein metabolism and coronary heart disease from these studies.
3. First, RVAS can facilitate gene discovery by identifying novel rare protein-altering variants in specific genes that are associated with disease.
4. Second, rare variant associations can provide supportive evidence for putative drug targets for novel therapies.
5. Finally, rare variants can uncover new pathways and reveal new biologic mechanisms.

Table 1

Advantages and disadvantages of rare variant association studies and common variant association studies.

Rare variant association studies		Common variant association studies	
Advantages	Disadvantages	Advantages	Disadvantages
<ul style="list-style-type: none"> • Can discover mutations at very low frequencies • Captures the complete spectrum of genetic variation across all types (copy number variation, indels, SNVs) • Provides direct biological insight 	<ul style="list-style-type: none"> • Sequencing is cost-prohibitive in large numbers of individuals • Statistical power of RVAS is still limited due to costs of sequencing large numbers of individuals • Expensive computational costs of storing sequence data for large-scale RVAS • So far, few novel gene discoveries have been made 	<ul style="list-style-type: none"> • Targets variation seen in many individuals • CVAS have now scaled up to hundreds of thousands of individuals • Analysis to conduct CVAS have largely been standardized • Several novel gene discoveries have been made • Genotyping arrays are cost-effective (comparatively) 	<ul style="list-style-type: none"> • Miss variation unique to individuals or seen in a few individuals • Often difficult to know the gene and variant involved in disease

Table 2

Coding variants in lipoprotein lipase (LPL)-/-regulating genes associated with blood lipid levels and their relationship to coronary heart disease.

Gene	Protein change	Lipid Phenotype	CHD risk	Frequency	Ref
LPL	p.S447X	↑HDL-C; ↓TG	↓ (OR=0.4)	Common	[40]
	p.D9N, p.N291S	↑TG	↑ (OR=5)	Common	[44]
APOA5	multiple	↓HDL-C; ↑TG	↑ (OR=2.2)	Rare	[43**]
APOC3	p.R19X, IVS2+1G>A, IVS3+1G>T, p.A43T	↓TG; ↓apoCIII	↓ (OR=0.6)	Rare	[31**]