Published in final edited form as:

Curr Opin Lipidol. 2011 April ; 22(2): 100–105. doi:10.1097/MOL.0b013e328342a375.

Combining genome-wide data from humans and animal models of dyslipidemia and atherosclerosis

Stela Z. Berisha¹ and Jonathan D. Smith^{1,2}

¹Department of Cell Biology, Lerner Research Institute, Cleveland Clinic, Cleveland, OH

²Department of Molecular Medicine, Cleveland Clinic Lerner College of Medicine of Case Western Reserve University, Cleveland, OH

Abstract

Purpose of review—Comparative genomics allows researchers to combine genome wide association data from humans with studies in animal models in order to assist in the identification of the genes and the genetic variants that modify susceptibility to dyslipidemia and atherosclerosis.

Recent findings—Association and linkage studies in human and rodent species have been successful in identifying genetic loci associated with complex traits, but have been less robust in identifying and validating the responsible gene and/or genetic variants. Recent technological advancements have assisted in the development of comparative genomic approaches, which rely on the combination of human and rodent datasets and bioinformatics tools, followed by the narrowing of concordant loci and improved identification of candidate genes and genetic variants. Additionally, candidate genes and genetic variants identified by these methods have been further validated and functionally investigated in animal models, a process that is not feasible in humans.

Summary—Comparative genomic approaches have lead to the identification and validation of several new genes, including a few not previously implicated, as modifiers of plasma lipid levels and atherosclerosis, yielding new insights into the biological mechanisms of these complex traits.

Keywords

comparative genomics; cross-species QTLs; cross-species eQTLs; dyslipidemia; atherosclerosis

Introduction

Plasma lipoprotein levels and the susceptibility to atherosclerosis are examples of complex traits, which are affected by the contribution of many different genes, gene-gene interactions, and gene-environment interactions. The recent success of ever larger consortiums performing genome wide association (GWA) studies in humans, which are based on the “common disease-common variant hypothesis, have recently identified 95 loci significantly associated with plasma levels of high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), total cholesterol (TC) or triglyceride (TG) levels [1**,2,3]. These findings have provided valuable insight in the biological mechanisms and pathways involved and have improved our understanding of interactions that underlie the pathology of dyslipidemia and other related complex diseases including atherosclerosis/

Address correspondence to: J.D.S. at: Cleveland Clinic, NC10, 9500 Euclid Avenue, Cleveland, OH 44195, USA, Telephone: (216) 444-2248, Fax: (216) 444-9404, smithj4@ccf.org.

The authors declare no conflict of interests.

coronary artery disease (CAD). However, despite yielding higher resolution than that obtained via family-based linkage analysis, GWA studies using haplotype-based tag single nucleotide polymorphisms (SNPs) have had little success in identifying the causal genetic variant, particularly when the best associated SNPs fall in intergenic regions [4,5]. Another limitation of GWA studies is that they often account for a small fraction of the heritable variance. Although, this might seem to support the “common disease- rare variant” hypothesis, a recent analysis has determined that most of the heritable variance for height is in fact captured by the common SNPs, although not by the statistical methods currently applied to ascertain genome wide significance [6**].

Model organisms for human complex traits, in which one can control for diet, environment, and other factors impossible to control for in human studies, have been developed and used for the identification of the susceptibility and modifier genes. The mammalian species most frequently used for these studies are mouse and rat, whose genomes show a high degree of synteny with the human genome. Strain intercrosses have been performed to identify quantitative trait loci (QTLs) linked to the levels of plasma HDL-C, LDL-C, TC, and TG, and for atherosclerosis lesion burden. Next, QTLs can be validated and their respective chromosomal intervals narrowed by creating and analyzing interval-specific congenic strain. Many QTLs for different complex traits have been identified and can be queried online at http://www.informatics.jax.org/searches/allele_form.shtml; however, relatively few of the responsible quantitative trait genes (QTGs) have been validated. QTL analysis has its own limitations associated with the prolonged and costly processes of narrowing down the QTL to a region of a few testable candidate genes.

The limitations of human GWA studies and rodent QTL analysis in identifying modifier genes have led to the development of new approaches using comparative genomics and co-expression network analysis to combine human GWA studies with rodent genetic and genomic studies. The use of this methodology to identify candidate genes involved in dyslipidemia and atherosclerosis will be the focus of this review, along with testing these candidate genes in animal models. The rapid advancement in this field has been made possible by technological advances including SNP microarrays to genotype millions of SNPs per sample, gene expression microarrays to obtain genome-wide transcriptome profiles, next generation sequencing to yield strain specific whole genome sequences, and innovative bioinformatic, statistical, and systems biology analyses that allow one to interpret this data.

Comparative genomics and bioinformatic tools to identify causal genes

Human GWA studies have led to the identification of clusters of SNPs, in linkage disequilibrium with each other, that reveal regions of usually < 50 kb that are associated with a complex trait such as lipoprotein levels, atherosclerosis, etc. Despite this level of resolution, it is often difficult to identify the causative variant; and, when the best associated SNPs are in close proximity to several genes or in an intergenic region, it may be difficult to even know which gene is responsible [7]. To better resolve a GWA locus and identify the best candidate genes and/or regulatory regions one can use a comparative genomics approach that relies on the assumption that regulation or variation in the same causative gene in concordant regions across multiple species can modify a phenotypic trait. Thus comparison of a human GWA locus with concordant QTLs identified in mouse or rat strain intercrosses may facilitate the detection of functionally conserved genes and regulatory regions [8]. Rodent strain intercrosses can identify loci associated with a complex trait through the QTL method; however, this method leads to large intervals that can contain hundreds of potential causative genes. Thus, both human GWA and rodent QTL methods have limitations, but the integration of these two types of analysis may lead to the rapid

identification of strong candidate genes relevant to human complex traits such as dyslipidemia and atherosclerosis.

Paigen and colleagues proposed a “toolbox” of comparative genomic and bioinformatics analyses that can be applied to rodent QTL studies to help identify the causal gene and variants (Figure 1) [9,10]. Mouse QTLs for a complex trait are sometimes mapped to the same locus in multiple strain intercrosses, which strengthens the evidence for a causative gene, and often helps narrow the likely interval. These mouse QTLs may overlap with syntenic regions identified as rat QTLs or human GWA or linkage-identified loci associated with a similar trait [11,12*]. These are called “concordant loci” and several studies have shown a remarkable concordance between human and mouse loci for HDL-C (93%), LDL-C (100%), TG (80%) levels, and atherosclerosis (63%) [9,13–16]. However, one should keep in mind that because QTLs for complex multifactorial disease are large and numerous it is possible that the observed concordant loci may be due to chance [8,11]. Comparative mapping of the concordant loci across species can be used to narrow the region of interest, although, if the concordance hypothesis for a specific locus is not true, it could lead to discarding the region in which the QTL resides.

Combining data from mouse and human studies has led to the recent identification of several candidate genes for dyslipidemia and related diseases, and has provided insight into their mechanism of action. Su *et al.* identified HDL-C QTLs from two mouse intercrosses on chromosomes 1, 3, 5, 6, 8, 9, 17, 18, and 19; and, these QTLs were located in regions homologous to human QTLs for HDL-C identified by family based linkage studies in various populations [13,17,18]. Comparative genomics helped narrow the mouse HDL-C QTL interval on chromosome 17 (*Hdlq56*) from 48 to 1.7 Mb, and in a similar fashion reduced the QTL interval of HDL QTLs on chromosome 3 (*Hdlq21*), 5 (*Hdlq52*), 8 (*Hdlq44*), and 18 (*Hdlq57*) [17]. The comparative genomics step was followed by mouse strain haplotype analysis, which is based upon sequence analysis that has determined that the genomes of the common mouse strains are composed of mosaic blocks from several ancestral strains; and, these regions of ancestral identity among strains have been mapped across the genome [19**]. The haplotype blocks that are shared between a strain pair are eliminated from further consideration to reduce the QTL interval, based on the assumption that the causative variant is derived from distinct ancestral origins. Although 97% of strain differences have been attributed to ancestral strain origins [20], more recent mutations within the shared haplotype blocks have occurred, thus this step has the possibility of eliminating the causative variant. When this method was applied to the HDL QTLs it led to a narrowing of the QTL interval to several dispersed genomic regions that have different ancestral origin in the high HDL strains vs. the low HDL strains [17].

The next step in this bioinformatics process involves looking for nonsynonymous SNPs and using gene expression differences to identify specific candidate genes. Su *et al.*, used these methods to identify *Lpl* (lipoprotein lipase), a gene known for its function in regulating HDL levels in mice and humans, as the candidate gene for the *Hdlq44* QTL [17]. They also identified *Nr2f2* (nuclear receptor subfamily 2, group F, member 2) as a candidate gene for chromosome 7 (*Hdlq58*) QTL [17]. *Nr2f2* gene was found to regulate the transcription of *Apoa2* and polymorphisms in this gene are found to be associated with HDL levels in a GWA study [21], showing cross species concordance.

Furthermore, in a different study Su *et al.*, identified the partially overlapping *Farp2* (FERM, RhoGEF and pleckstrin domain protein 2) and *Stk25* genes (serine/threonine kinase 25) as candidate genes for an HDL QTL on chromosome 1 [22]. Although they did not use cross-species comparative genomics to narrow down the interval, they observed that both genes are located on human chromosome 2q37.2 where a human QTL for HDL was previously

identified [23]. Also, a GWA study found polymorphisms in the *FARP2* gene associated with variations of human HDL levels [21], thus providing strong evidence that *Farp2* and/or the adjacent *Stk25* gene identified in mice are likely to play a role in regulating HDL levels in humans as well.

Another successful example of the comparative genomics approach is the narrowing of a mouse HDL QTL on chromosome 18 from 9 Mb to 6.9 Mb and the identification of *Lipg* (endothelial lipase) as the most likely candidate gene for this QTL by haplotype, gene expression, and sequence variation analyses [24**]. Several recent GWA studies have found *LIPG* SNPs associated with HDL-C levels in humans [2,3,25–27]. Also, a rare loss-of-function mutation in the *LIPG* gene (*Asn396Ser*) was found to be significantly associated with increased HDL-C levels in human subjects [28].

Application of the abovementioned set of bioinformatic tools [10], including comparative genomics, identified *Scarb1* (scavenger receptor class B, member 1) and *Acads* (acyl-Coenzyme A dehydrogenase, short chain) as candidate genes for HDL QTLs on mouse chromosome 5 [29*]. *Acads* in mice is adjacent to the *Mvk* and *Mmab* genes, which were found to be present in a cluster of five genes positioned in a concordant human HDL QTL located at 12q24.23. Recently, two GWA studies have shown that this genomic region contains SNPs significantly associated with HDL levels in humans [26,27]. Additional studies, including generation of knockdown, knockout and/or transgenic mice, are necessary to investigate the relationship of the *Mvk*, *Mmab* and *Acads* genes with HDL metabolism.

Cross-species expression QTLs and co-expression networks

Gene expression profiling using microarrays can simultaneously assess the levels of thousands of transcripts. These datasets have provided valuable insights into the functional relationships among genes and in understanding human complex traits. Due to the difficulty in collecting the correct tissue type relevant to a complex trait in humans, these studies are often performed in a rodent species, with the caveat that not all mechanisms are conserved between humans and rodents. Wang *et al.*, developed a new method to analyze co-expressed networks; and, by comparing human, mouse, and rat liver co-expression networks they found that ~10% of the gene-gene co-expression network relationships were conserved [30**]. The authors identified a co-regulated network of genes across species that were over represented in cell signaling, cell adhesion, and sterol biosynthesis pathways; and, some of these genes were also found as candidates in a human GWA study for plasma lipid levels. By integrating GWA study candidate loci, human *cis*-eQTLs, and conserved co-expression modules between human and rodent species, the authors identified *Sort1*, *Fads1*, *Fads2* and *Galnt2* as the genes that are most likely to be responsible for regulating plasma lipid levels [30**].

In an attempt to validate mouse candidate genes as modifiers of aortic lesions, Yang *et al.*, performed cross-species analyses using mouse adipose and liver eQTLs along with aortic lesion QTLs from a strain intercross on the *ApoE*^{-/-} background and compared them to relevant human GWA study results [31**]. The authors identified five human genes from the GWA studies (*ATP10D*, *SLC2A9*, *PIK3CG*, *LOXL1* and *KIAA1598*) that overlapped with their list of mouse candidate genes, suggesting these genes as the strongest candidate modifiers for aortic lesions in mice [31**].

Use of mouse models to test candidate loci identified from human GWA studies

Human GWA studies have consistently identified variants in the 9p21 region as the SNPs most highly associated with increased risk for coronary artery disease (CAD) and myocardial infarction (MI) [32–34]. CAD associated SNPs in this region are positioned within a 58-Kb linkage disequilibrium block not closely associated with protein coding genes. *ANRIL*, a large non-coding RNA is found in this block; although, there is no evidence to date that this transcript is expressed in mice. In order to find out the mechanism(s) by which these SNPs confer an increased risk for CAD and MI, Visel *et al.*, identified the 70 Kb mouse syntenic region on chromosome 4 with conserved flanking genes [35**]. Using Cre-Lox methodology, they created mice lacking this region (Chr4^{Δ70kb/Δ70kb}), which did not have significant abnormalities in urine and blood chemistry markers; however, 45% of them developed neoplasms of various types [35**]. The expression levels of the flanking cyclin dependent kinase *Cdkn2a* and *Cdkn2b* genes decreased by more than tenfold in the hearts of the Chr4^{Δ70kb/Δ70kb} mice compared to the wild type controls [35**]. In heterozygous mice, allelic imbalance in *Cdkn2b* gene expression proved that this 70 Kb mouse region acts as a *cis*-acting enhancer in the heart and aorta [35**]. They also observed that primary aortic smooth muscles and mouse embryonic fibroblasts cell cultures from chr4^{Δ70kb/Δ70kb} mice proliferated excessively and were resistant to senescence compared to wild type cells, which could explain the excessive neoplasms. Thus, comparative genomics and functional testing in mice revealed that the most likely mechanism for the association of human genetic variation in the 9p21 with CAD is through the regulation of *Cdkn2a/b* gene expression.

As demonstrated by the above study and others, testing human candidate genes or genetic variants identified by GWA studies in animal models is useful for validation and can provide insight into the biological mechanisms responsible for the pathophysiology of human complex diseases. A recent plasma lipid/lipoprotein GWA study with >100,000 individuals of European ancestry was performed that identified 95 loci significantly associated with plasma concentrations of TC, LDL-C, HDL-C and TG, 59 of which were novel findings [1**]. The candidate genes *GALNT2* (UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 2), *PPP1R3B* (protein phosphatase 1, regulatory subunit 3B) and *TTC39B* (tetratricopeptide repeat domain 39B) were tested in mouse models [1**]. Each one of these genes was separately validated as causal for plasma lipid levels either through liver specific over expression via a viral vector or the knockdown of endogenously expressed gene through the delivery of short hairpin RNAs. Over expression of the mouse ortholog *Galnt2* resulted in significantly lower plasma HDL-C levels compared to wild type mice; and, knocking down its expression by ~95% resulted in higher HDL-C levels, thus validating human *GALNT2* genetic variation as causal in regulating lipid levels [1**]. The two other candidate genes, *PPP1R3B* and *TTC39B*, were identified from human liver eQTL analysis. Their validation in regulating lipid metabolism was demonstrated by over expression of the mouse ortholog *Ppp1r3b* that resulted in reduced HDL-C and TC levels, and by the knock down of the expression of the mouse ortholog *Ttc39b* that resulted in increased HDL-C levels [1**].

GWA studies identified several SNPs in the 1p13 region that are associated with CAD/MI and LDL-C levels; and, Musunuru *et al.*, found that the minor allele of one of these SNPs was associated with increased liver expression levels of *SORT1* (sortilin 1, also known as neurotensin receptor 3), a nearby gene not previously implicated in lipoprotein metabolism, and decreased levels of LDL-C, [36**]. Liver-specific over expression of the mouse ortholog *Sort1* via a viral vector was found to be associated with a marked reduction in plasma TC and LDL-C, a result observed consistently in the *Apobec1*^{-/-}; *APOB* Tg mice

and three other background strains [36**]. Furthermore, *Sort1* over expressing mice were found to have a large decrease in the rate of VLDL and TG secretion [36**]. Meanwhile, liver specific knockdown of *Sort1* in *Apobec^{-/-}*; *APOB* Tg mice resulted in a large increase in TC, LDL-C, and VLDL levels [36**]. Results obtained from *Sort1* knockout mice were found to be consistent with liver-specific knockdown mice, and with the inverse correlation between hepatic SORT1 expression and LDL-C levels observed in humans.

Independently, Kjolby *et al.* explored the *Sort1* gene as a potential contributor to lipoprotein metabolism in mice [37*]. They observed that recombinant adenovirus mediated *Sort1* over expression led to increased TC and apoB-100 plasma levels, while the *Sort1^{-/-}* deficient mice on the C57BL/6 *Ldlr^{-/-}* background had lower plasma levels of apoB-100, TC, IDL and LDL when compared with *Ldlr^{-/-}* controls [37*]. Therefore, this study found a positive correlation between *Sort1* expression levels and LDL-C levels. Thus, the two above studies came to opposite conclusions regarding the relationship between the levels of *Sort1* expression and LDL-C. This discrepancy might be due to several factors including, but not limited to, differences in strain background, mouse genetic models, and diet.

A diagram of many of the cross-species and bioinformatic approaches described in this review is presented in Figure 2.

Conclusion

Taking advantage of the concordance of human and rodent loci associated with similar complex traits, as well as the synteny between their genomes, researchers have developed strategies to combine human and rodent genetic and genomic studies to help identify the genes that modify these clinical traits, and the responsible genetic variants. These investigations can be initiated either from human GWA findings or from rodent QTL studies, and yield increased power through cross-species comparisons. Although these methods may yield time-saving shortcuts in modifier gene identification, one must be cautious of the many assumptions in specific bioinformatic steps that may not be universally valid. Nevertheless, these comparative genomic strategies offer many advantages and we anticipate increased usage of these cross-species methods to find genes and variants associated with common complex traits.

Acknowledgments

This work was supported by National Institute of Health Grant HL098193 to J.D. Smith.

References

Papers of particular interest, published within the annual period of review, have been highlighted as:

* of special interest

** of outstanding interest

- **1. Teslovich TM, Musunuru K, Smith AV, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*. 2010; 466:707–713. GWA study of more than 100,000 individuals of European ancestry that identified 95 common variants associated with human plasma levels of HDL, LDL, TC and TG. This study also validated three candidate genes in animal models demonstrating their role in regulating plasma lipid levels. [PubMed: 20686565]
2. 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; 41:47–55. [PubMed: 19060911]

3. Kathiresan S, Willer CJ, Peloso GM, et al. Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat Genet.* 2009; 41:56–65. [PubMed: 19060906]
4. Altshuler D, Daly MJ, Lander ES. Genetic Mapping in Human Disease. *Science.* 2008; 322:881–8. [PubMed: 18988837]
5. Frazer KA, Murray SS, Schork NJ, Topol EJ. Human genetic variation and its contribution to complex traits. *Nat Rev Genet.* 2009; 10:241–251. [PubMed: 19293820]
- **6. Yang J, Benyamin B, McEvoy BP, et al. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet.* 2010; 42:565–569. Demonstrates that most of the missing heritability in GWA studies is attributed to common variation, but that current statistical tests are not robust in identifying the responsible variants. [PubMed: 20562875]
7. Manolio TA. Genomewide association studies and assessment of the risk of disease. *N Engl J Med.* 2010; 363:166–176. [PubMed: 20647212]
8. Moreno C, Lazar J, Jacob HJ, Kwitek AE. Comparative genomics for detecting human disease genes. *Adv Genet.* 2008; 60:655–697. [PubMed: 18358336]
9. DiPetrillo K, Wang X, Stylianou IM, Paigen B. Bioinformatics toolbox for narrowing rodent quantitative trait loci. *Trends Genet.* 2005; 21:683–692. [PubMed: 16226337]
10. Burgess-Herbert SL, Cox A, Tsaih SW, Paigen B. Practical applications of the bioinformatics toolbox for narrowing quantitative trait loci. *Genetics.* 2008; 180:2227–2235. [PubMed: 18845850]
11. Stoll M, Kwitek-Black AE, Cowley AW Jr, et al. New target regions for human hypertension via comparative genomics. *Genome Res.* 2000; 10:473–482. [PubMed: 10779487]
- *12. Cox A, Sheehan SM, Kloting I, et al. Combining QTL data for HDL cholesterol levels from two different species leads to smaller confidence intervals. *Heredity.* 2010; 105:426–432. Demonstrates the possibility of narrowing down a QTL interval for HDL cholesterol levels by combining QTL data from rat and mouse. [PubMed: 20551980]
13. Rollins J, Chen Y, Paigen B, Wang X. In search of new targets for plasma high-density lipoprotein cholesterol levels: promise of human-mouse comparative genomics. *Trends Cardiovasc Med.* 2006; 16:220–234. [PubMed: 16980179]
14. Wang X, Ishimori N, Korstanje R, Rollins J, et al. Identifying novel genes for atherosclerosis through mouse-human comparative genetics. *Am J Hum Genet.* 2005; 77:1–15. [PubMed: 15931593]
15. Wang X, Paigen B. Genetics of variation in HDL cholesterol in humans and mice. *Circ Res.* 2005; 96:27–42. [PubMed: 15637305]
16. Wang X, Paigen B. Genome-wide search for new genes controlling plasma lipid concentrations in mice and humans. *Curr Opin Lipidol.* 2005; 16:127–137. [PubMed: 15767852]
17. Su Z, Wang X, Tsaih SW, et al. Genetic basis of HDL variation in 129/SvImJ and C57BL/6J mice: importance of testing candidate genes in targeted mutant mice. *J Lipid Res.* 2009; 50:116–125. [PubMed: 18772481]
18. Ishimori N, Li R, Kelmenson PM, et al. Quantitative trait loci analysis for plasma HDL-cholesterol concentrations and atherosclerosis susceptibility between inbred mouse strains C57BL/6J and 129S1/SvImJ. *Arterioscler Thromb Vasc Biol.* 2004; 24:161–166. [PubMed: 14592847]
- **19. Tsaih SW, Korstanje R. Haplotype association mapping in mice. *Methods Mol Biol.* 2009; 573:213–222. Presents haplotype association mapping in mice, a novel phenotype-driven approach similar to the GWA studies in humans. The strategy helps identify genetic loci associated with the complex trait under study by using inbred mouse strains as individuals. [PubMed: 19763930]
20. Frazer KA, Wade CM, Hinds DA, et al. Segmental phylogenetic relationships of inbred mouse strains revealed by fine-scale analysis of sequence variation across 4.6 mb of mouse genome. *Genome Res.* 2004; 14:1493–1500. [PubMed: 15289472]
21. Saxena R, Voight BF, et al. Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University, and Novartis Institutes of BioMedical Research. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science.* 2007; 316:1331–1336. [PubMed: 17463246]

22. Su Z, Cox A, Shen Y, et al. Farp2 and Stk25 are candidate genes for the HDL cholesterol locus on mouse chromosome 1. *Arterioscler Thromb Vasc Biol.* 2009; 29:107–113. [PubMed: 18988887]
23. Elbein SC, Hasstedt SJ. Quantitative trait linkage analysis of lipid-related traits in familial type 2 diabetes: evidence for linkage of triglyceride levels to chromosome 19q. *Diabetes.* 2002; 51:528–535. [PubMed: 11812765]
- **24. Su Z, Ishimori N, Chen Y, et al. Four additional mouse crosses improve the lipid QTL landscape and identify *Lipg* as a QTL gene. *J Lipid Res.* 2009; 50:2083–2094. Demonstrates that combining data from four different mouse crosses and using comparative genomic, haplotype, genomic sequence variation, and gene expression analyses can help identify the causal gene for the HDL QTL on the distal end of chromosome 18. [PubMed: 19436067]
25. Heid IM, Boes E, Muller M, et al. Genome-wide association analysis of high-density lipoprotein cholesterol in the population-based KORA study sheds new light on intergenic regions. *Circ Cardiovasc Genet.* 2008; 1:10–20. [PubMed: 20031538]
26. Kathiresan S, Melander O, Guiducci C, et al. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat Genet.* 2008; 40:189–197. [PubMed: 18193044]
27. Willer CJ, Sanna S, Jackson AU, et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet.* 2008; 40:161–169. [PubMed: 18193043]
28. Edmondson AC, Brown RJ, Kathiresan S, et al. Loss-of-function variants in endothelial lipase are a cause of elevated HDL cholesterol in humans. *J Clin Invest.* 2009; 119:1042–1050. [PubMed: 19287092]
- *29. Su Z, Leduc MS, Korstanje R, Paigen B. Untangling HDL quantitative trait loci on mouse chromosome 5 and identifying *Scarb1* and *Acads* as the underlying genes. *J Lipid Res.* 2010; 51:2706–2713. Demonstrates the use of genetic and genomic resources to unravel a complex HDL QTL to identify candidate genes in multiple mouse crosses. [PubMed: 20562441]
- **30. Wang K, Narayanan M, Zhong H, et al. Meta-analysis of inter-species liver co-expression networks elucidates traits associated with common human diseases. *PLoS Comput Biol.* 2009; 5:e1000616. Demonstrates that systematic comparisons of co-expression networks among human and rodent species can help illuminate common mechanisms of complex diseases such as hyperlipidemia and type 2 diabetes. [PubMed: 20019805]
- **31. Yang X, Peterson L, Thieringer R, et al. Identification and validation of genes affecting aortic lesions in mice. *J Clin Invest.* 2010; 120:2414–2422. compares the candidate genes for aortic lesions in mice identified through gene expression analysis with findings from human GWA studies of CAD and related traits. [PubMed: 20577049]
32. 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]
33. 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]
34. Samani NJ, Erdmann J, Hall AS, et al. Genomewide association analysis of coronary artery disease. *N Engl J Med.* 2007; 357:443–453. [PubMed: 17634449]
- **35. Visel A, Zhu Y, May D, et al. Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. *Nature.* 2010; 464:409–412. Identifies the mechanism for the intergenic 9p21 CAD associated locus by creating a mouse deletion in the orthologous region. The authors demonstrated that this region controls the expression of the neighboring *Cdkn2a/b* genes in cardiac tissue via *cis*-acting mechanism. [PubMed: 20173736]
- **36. 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. Demonstrates that an intergenic polymorphism discovered through GWA studies is associated with hepatic expression of the *SORT1* gene. Over expression and knock down of the mouse ortholog was shown to alter plasma LDL and VLDL levels. [PubMed: 20686566]
- *37. Kjolby M, Andersen O, Breiderhoff T, et al. *Sort1*, encoded by the cardiovascular risk factor 1p13.3, is a regulator of hepatic lipoprotein export. *Cell Metabolism.* 2010; 12:213–223. Demonstrates that over expression and knockout of the mouse *Sort1* gene is associated with altered plasma LDL, TC and apoB-100 levels. [PubMed: 20816088]

Key Points

- Human and rodent association and linkage studies often give rise to concordant loci for dyslipidemia and other complex traits.
- Bioinformatic analyses can be performed to combine findings from different species and help narrow the interval to a few candidate genes.
- Gene expression analyses and cross-species expression QTLs can also help identify promising candidate genes.
- Animal models can be used to characterize the mechanisms responsible for human genetic variation associated with complex traits, and to validate dyslipidemia and atherosclerosis modifier genes discovered through human GWA studies.

Figure 1. Use of bioinformatic tools including comparative genomics to refine a QTL interval. Hypothetical example of a mouse complex trait QTL with the shaded regions representing the QTL region of interest through stepwise analyses as denoted. The respective size (in megabases) and number of genes are shown on the right. Black lines within each of the individual and combined mouse QTLs represent the QTL peaks. Adapted from *Burgess-Herbert SL, et al.* [10]

Figure 2.

Diagram of various strategies to combine animal and human studies to help identify candidate genes and regulatory regions for complex traits. Although we refer to mouse studies, these strategies may be applied to other species as well.