Published in final edited form as:

Curr Cardiovasc Risk Rep. 2014 January ; 8: 372. doi:10.1007/s12170-013-0372-3.

How Gene Networks Can Uncover Novel CVD Players

Laurence D Parnell¹, Patricia Casas-Agustench², Lakshmanan K Iyer³, and Jose M Ordovas^{1,2}

¹Nutritional Genomics Laboratory, JM-USDA Human Nutrition Research Center on Aging at Tufts University, 711 Washington Street, Boston, MA 02111

²Instituto Madrileño de Estudios Avanzados (IMDEA) Alimentación, CEI UAM+CSIC, C/Faraday, 7, 1^a planta D1.11, Ciudad Universitaria de Cantoblanco, Ctra. de Colmenar Km.15, Madrid, 28049, Spain

³Tufts Center for Neuroscience Research, Tufts University School of Medicine, 136 Harrison Ave, Boston, MA 02111, Molecular Cardiology Research Institute, Tufts Medical Center, 15 Kneeland Street, Boston, MA 02111

Abstract

Cardiovascular diseases (CVD) are complex, involving numerous biological entities from genes and small molecules to organ function. Placing these entities in networks where the functional relationships among the constituents are drawn can aid in our understanding of disease onset, progression and prevention. While networks, or interactomes, are often classified by a general term, say lipids or inflammation, it is a more encompassing class of network that is more informative in showing connections among the active entities and allowing better hypotheses of novel CVD players to be formulated. A range of networks will be presented whereby the potential to bring new objects into the CVD milieu will be exemplified.

Keywords

cardiovascular disease; genetics; interactions; network; risk

Introduction

Risk factors for CVD include many different genetic variants, small molecules, lipid particles and proteins, but also such more encompassing terms as foam cell activation, family history, habitual diet and gut microbiome composition. Cholesterol particles found in serum have long been measured and their levels considered as indicators of CVD risk. In this regard though, circulating high-density lipoprotein cholesterol (HDL-C) and low-density lipoprotein cholesterol particles (LDL-C) are not homogeneous, but rather consist of several interacting entities, including proteins, various classes of lipids and microRNAs. Each of these entities performs specific functions within the lipoprotein particle or when that

Contact: Laurence.parnell@ars.usda.gov, 617-556-3089, 617-556-3211 (F).

Compliance with Ethics Guidelines

Human and Animal Rights and Informed Consent

This article does not contain any studies with human or animal subjects performed by the author.

Conflict of Interest

Laurence D Parnell declares that he has no conflict of interest.

Patricia Casas-Agustench declares that she has no conflict of interest.

Lakshmanan K Iyer declares that he has no conflict of interest.

Jose M Ordovas declares that he has no conflict of interest.

particle is delivered to a cell. These functions, centered primarily on the transport of lipid moieties to peripheral tissues and in reverse to the liver, and their roles in CVD risk are well characterized. However, those functions should be considered within the context of all entities present and their interactions among themselves. Thus, the HDL-C particle, for example, can be thought of as a collection of interacting entities that together perform certain functions. Just as these well known lipid particles are dynamic, so too are the biological networks that can be described as operating within each particle. The connections are fluid – stronger under certain conditions, such as the postprandial state or at a given time of day, and weaker or non-existent in other cases, for example when certain genetic variants prevent an interaction.

Networks are composed of nodes connected to each other by edges, which represent a biological interaction (Examples are sketched in Figure 1). Such networks of interacting entities are often called interactomes and these networks are highly applicable to CVD because relationships and information on different nodes can be represented, such as diet, disease, gene and genetic variation, and small molecules like food or drugs. The edge connecting two nodes can be without direction or weight, or can be annotated to include effect (increases, decreases), weight (degree of effect, or number of reports indicating that relationship), or other conditions (in a certain cell type, or with regard to a specific condition). When clear and uncluttered, such networks can depict a wealth of information that has the potential to inspire formulation of new, testable hypotheses.

While it may be well understood that clinical parameters of CVD risk are not typically evaluated in isolation, common practice shows otherwise. A single risk factor value beyond an accepted threshold is often used to assign a patient to a therapeutic regimen. What ought to happen instead is an evaluation of several measures of disease status being taken and considered under certain relationships to each other, such as co-occurrence (low HDL-C with high serum triglyceride (TG) levels), medication, or with regard to family history (ie, genetics). These relationships are amenable to network analysis. Furthermore, the many genes that either encode proteins responsible for CVD-related functions or were identified by genome-wide scans for genetic variants associating with disease risk or incidence (genome-wide association studies, or GWAS) can often be depicted in networks with their inter-connecting links displayed. Many such networks pertinent to blood lipids and CVD are known. These could be augmented with other data to identify novel candidates within the larger CVD gene framework. Such an integrative approach was employed to identify genes involved in mediating the APOE ϵ 4-associated pathway that promotes late-onset Alzheimer's disease (LOAD)[1]. Differential co-expression correlation network analysis of APOE4 carriers and LOAD patients identified several modulators central to functions of amyloid beta A4 precursor protein. Such an approach, for example, as applied to lipids, CVD in general, or the even the elements, such as genes and metabolites, that are sensitive to the diet also can be effective in identifying players in CVD. These could be new or show new levels of importance.

This review will highlight different types of networks relevant to CVD and how data integration and network analysis has the potential to uncover novel players in CVD. Types of networks considered here include – expression quantitative trait loci (eQTL), RNA-based networks, gene-metabolite networks, protein-lipid networks, nuclear hormone networks, lipidomics, inflammation, and genetics networks and their modulation by lifestyle factors.

Networks describing control of gene activity

eQTL

An expression quantitative trait locus, or eQTL, is a genetic marker, typically a SNP, that associates with mRNA levels of a nearby (*cis*) or distant (*trans*) gene. GWAS have identified many variants associating with disease risk and by far the emerging trend in this field is to describe the mechanism by which those variants confer risk. Allele-specific gene expression differences increasingly are noted as drivers of the observed phenotypic differences [2–4]. For example, genetic variants associating with LDL-C and VLDL levels have been found that map within and in the vicinity of *PSRC1* and *SORT1* on chromosome 1p13 [5,6]. Detailed analysis of this region showed that variant rs12740374 creates a binding site for the CCAAT/enhancer binding protein transcription factor and affects the hepatic expression of the *SORT1* [6]. Interestingly, this locus contains variants that are under positive selection, implying a role of environmental, perhaps dietary, factors in shaping the genetic architecture that regulates expression of this gene implicated in coronary heart disease and myocardial infarction [7].

Large-scale mapping of eQTL onto SNPs previously assigned to disease risk associations has shown recently that eQTL are drivers of disease phenotypes and associations [8]. In this regard it is satisfying to see that SNPs associating with the disease phenotype(s) also associate with expression of genes known to be involved in that disease or its underlying functional networks. In order to provide the data needed to accomplish such disease-gene expression links, the GTEx Consortium was formed. This project aims to build a resource database and associated tissue bank for the study of the links between genetic variation and gene expression in over 45 different human tissues [9]. Overall, this approach shows great promise in clarifying the role of genetic variation and the risk of cardiometabolic disorders but has yet to be applied to epistatic and gene-environment interactions.

RNA-based networks

MicroRNAs, or miRs, are short RNA entities typically 22–26 nt (nucleotides) in length that bind to mRNAs and direct their decay. This regulatory event serves to down-regulate expression of the mRNA's encoded protein. Long, non-coding RNAs, or lncRNAs, are > 200 nt and have been reported in some cases to function as a decoy, competing for miRs that regulate mRNA translation into protein. lncRNAs can also function as an anti-sense message and interfere with mRNA stability and translation into a polypeptide. At the same time, many genetic variants identified by GWAS and other analyses as relevant to disease status and risk are increasingly seen as altering gene expression [8,10,11].

Cholesterol transport via ABCA1 protein and regulation of this process via SREBP1 and SREBP2, sterol regulatory element binding transcription factors, showcase an elegant mRNA-miR network. In this network, miR-33B, which is encoded within intron 16 of the primary SREBF1 transcript, and miR-33A, which is likewise encoded within SREBF2, offer balancing fine-tuning of the control of expression of a number of genes participating in cholesterol transport and reverse cholesterol transport, such as ABCA1, ABCG1, NPC1 and ABC11 [12–14]. In this scenario, SREBF1/2 promotes transcription of the aforementioned genes while miR-33A/B interfere with the translation of these mRNAs. Together, these genes and the two miRs can be placed into an RNA network that has cholesterol homeostasis as its output. Validation of other miR33A/B targets can identify novel functional elements within the cholesterol homeostasis pathway, or bridges to other pathways that are at least in part co-regulated with cholesterol homeostasis. In this regard, the discovery that miR-33A/B regulate IRS2 in hepatic cells adds a link between cholesterol homeostasis and insulin signaling [15]. With regard to the application of miRs as therapeutic

agents, a very recent report indicated that subcutaneous delivery of a type of octomer anti-miR to obese and insulin-resistant nonhuman primates resulted in relieving suppression of miR-33 targets [16]. Specifically, this included ABCA1 and an increase in circulating HDL cholesterol [16]. A second example of this network type brings together the lipoprotein lipase gene *LPL*, its SNP rs13702, miR-410 and lipid traits [17]. Meta-analysis of ten cohorts showed that SNP rs13702, which was shown to disrupt an interaction between *LPL* mRNA and miR-410, associated with HDL-cholesterol and triacylglycerol (TAG) levels, with interactions between intake of polyunsaturated fatty acid and TAG being detected.

Gene-metabolite networks

Some genes encode enzymes where reactants yield products and these products, or metabolites, can elicit certain biological effects. For other proteins, their activity is modulated by interactions with metabolites. Clinically important examples include C-peptide (from processing of pro-insulin to its active form), estrone sulfate, homocysteine and even dietary sodium with its connections to blood pressure via sodium pumps in the glomeruli of the kidney.

The flavin containing monooxygenase 3 (FMO3) catalyzes the NADPH-dependent oxygenation of various xenobiotics and is the primary such enzyme in the liver. FMO3 catalyzes the production of trimethylamine N-oxide (TMAO), which is derived from dietary phosphatidylcholine, and plasma levels of TMAO and choline are CVD risk indicators [18]. In mice, hepatic *Fmo3* synthesizes TMAO from trimethylamine. In vivo over-expression or silencing of *Fmo3* increases or decreases plasma TMAO levels, respectively [19]. Notably, *Fmo3* expression is repressed by testosterone and induced by bile acids via the NR1H4 nuclear receptor also known as FXR [19]. Furthermore, natural variation in TMAO levels contributes to atherosclerosis susceptibility in mice [19]. Another study elegantly demonstrated that omnivorous humans made more TMAO than vegans or vegetarians after ingestion of L-carnitine, and the TMAO was produced in a microbiome-dependent manner [20]. Thus, thinking of a network comprised of dietary intake of red meat, L-carnitine, gut microbiota, TMAO and the FMO3 enzyme assists in the understanding of CVD risk.

Recent years have seen the advent of the human serum and urine metabolome databases [21,22] and increased application of metabolite profiling of human samples. With resources such as high-throughput metabolomics accompanied by thorough data analysis and comparison to the aforementioned databases, it is now possible to identify dysregulated pathways and potential new biomarkers of disease status. In this regard, such studies have indicated that plasma levels of cystine, urea cycle intermediates and the dibasic amino acids citrulline, lysine and ornithine correlate with the LDL-C lowering response of simvastatin [23]. Connections between branched-chain amino acids, their interacting proteins and CVD have been investigated by many. One example of work in this area showed that increased levels of three amino acids – tyrosine, phenylalanine and isoleucine – significantly correlated with incidence of carotid IMT, the occurrence of a plaque larger than 10 mm² and inducible ischemia [24].

Protein-lipid networks

Certainly, there are many important protein-lipid interactions with regard to CVD. Well known examples include the chylomicron and associated lipoprotein-cholesterol particles such as HDL, adipocyte lipid droplets and their surface-binding proteins such as perilipins, acyl-CoA synthetase long-chain family members and the SPDR gene product. Here, we briefly summarize other emerging, CVD-relevant networks at the interface of lipids and proteins.

The v-src avian sarcoma viral oncogene homolog encodes the SRC tyrosine protein kinase with established roles in oncogenesis and metabolic homeostasis. Importantly, the SRC protein has many substrates to which it adds a regulatory phosphate group and is itself a highly networked molecule. Those interactions though are not limited to proteins. For example, SRC action occurs according to membrane constituent phospholipids in adipocytes – environment specific events and connections with consequences in signaling [25]. SRC also acts as one of many inputs regulating gluconeogenesis and glycogenolysis in the liver [26], which are both elevated in metabolic syndrome. These findings strongly suggest that the constitution of fat in the diet has the potential to influence trafficking in the plasma lipid bilayer of this important signaling molecule. Thus, the more complete view of this important protein involves its network of interacting proteins and membrane lipids.

Lastly, endogenous endocannabinoids, a family of compounds consisting of fatty acids and alcohols attached to various polar head moieties, are affiliated with lipoproteins with effects on signaling through SMO, or smoothened [27]. One action of SMO is to promote insulin-independent glucose uptake in skeletal muscle and brown adipose [28]. Importantly, the endocannabinoids do not appear to act through binding to known receptors such as PPARA and TRPV1 [27]. In addition, it has been found that targeted knock down of mouse *Abhd6* (a 2-arachidonoylglycerol hydrolase important in brain endocannabinoid signaling) shielded mice from high-fat diet-induced obesity among other metabolic dysfunctions [29].

Nuclear hormone networks

Nuclear hormones and their receptors form networks with important roles in CVD. Estrogen, for example, has been shown to have vasculoprotective effects in premenopausal women. However, it also has adverse effects in older women with menopausal symptoms. Recent findings have shown that the smooth muscle mineralocorticoid receptor (NR3C2) has important roles in blood pressure control and vascular aging [30]. Two recent studies have identified key pathways and genes that are involved in mediating the effects of nuclear hormones in vasculature with an aim to characterize their roles in CVD [31,32]. In one of those studies, the aldosterone-regulated transcript profile identified the oxidative stress pathway and genes such as *PIGF*, *MT1*, *MT2*, *CTGF*, *BIRC2* and *SGK*, operating via NR3C2, at the site of vascular injury [31]. The second report examined the effects of estrogen treatment on blood vessels, finding that this treatment affected expression preferentially of such functional pathways as gene expression, lipid metabolism, cell-cell signaling, cell growth/proliferation and cellular movement [32]. The mechanism of action of estrogen and other nuclear ligands in cardiovascular systems requires further investigation in order to get a better understanding [33].

Lipidomics networks

When the call for thorough lipid analyses, beyond the standard measures of triglycerides and cholesterol, is put forth both for diagnostic purposes and for checking the efficacy of prescribed therapy [34], this necessitates the development and use of data, which link those lipids to proteins that act as transporters or catalytic modifiers of the lipids, and then to genes and genetic loci of disease risk, followed by cell types or tissues where those lipid moieties are likely to be incorporated. The human serum metabolome lists over 3300 lipid and fatty acid compounds [21], and while >99% have links to a metabolic enzyme, other information compiled within any database on implications in disease or connections to food or gut microbiome sources are lacking. Here more work is needed and more so because a good metabolomics profile will give some information on fatty acids and other lipids detected in the sample. Along the lines of identifying lipids from tandem mass spectroscopy data LipidBlast has been developed covering 119,200 compounds from 26 lipid classes [35].

This is one such tool that will be very useful as lipidomics and lipid-centric networks gain in relevance to CVD research and diagnostics. On the other hand, development and use of a comprehensive database that links food items to biochemical and bioactive compounds detected within that item would offer an alternative to correct lipidomic imbalances by other means – say by combined nutrition and pharmacological approaches [36].

Inflammation networks

The impact of inflammation on CVD risk and outright occurrence has been investigated for some time [37], with inflammation being invoked to characterize key links between CVD and, for example, health of the gut microbiome [38], sickle cell disease [39], adiposity [40] and atherosclerosis [41,42]. By the very nature of the molecules involved – cytokines and their receptors on various cell types – inflammation as a precursor of or response to cardiovascular diseases lends itself to a network or systems view. Within adipose, for example, the homeostatic relationship between the adipocyte and the M2 macrophage is altered and increased communication with the M1 macrophage elicits a series of pro-inflammatory events signifying the metabolic syndrome condition. Increased levels of available saturated fatty acids, coupled with increases in inflammatory cytokines such as IL6, IL1B and TNF, and a loss of repression of NFKB1, eventually lead to an uptick in the recruitment of monocytes, via coordinate increases from both adipocytes and macrophages in MCP-1 (CCL2), SAA, and S100A8/9 [43].

Working within the Framingham Heart Study, researchers were able to identify a sub-network of inflammation of genes whose expression in platelets and megakaryocytes is statistically tied to expression of the well known inflammation biomarkers CRP and IL6 [44]. These genes included *ALOX5*, *CD163*, *IFIT1*, *IL6*, *NFKB1* and *TNFRSF1B*. Subsequent research along these lines, for example, could look at the relationship between genetic variants associating with CRP and IL6 expression and the activity of the identified inflammation genes, perhaps in ways modulated by lifestyle choices such as the diet or exercise. Lastly, the discovery of new anti- and pro-inflammatory pathways that connect lipid and inflammation biology has set the stage for exploring a new set of diverse therapeutic targets for coronary artery disease [42].

Genetic networks

Studies defining genetic loci of importance to CVD have been ongoing for quite some time and many genetic loci conferring increased risk of CVD have been described. Over the last seven or so years, the application of GWAS has accelerated the rate of discovery of loci associating with either CVD occurrence or clinical measures of elevated risk of CVD. These latter include reduced levels of HDL, increased levels of total cholesterol, triglycerides or LDL, and anthropometrics of obesity and adiposity. The focus in this part of this review will be on two aspects: genetic networks involving epistasis and genetic risk scores, and gene-environment interactions.

Genetic risk score and CVD

Epistasis generally refers to the interaction between different genes. The focus of GWAS for CVD traits on main effects has left open the question of the extent of contribution to trait variance by interactions with other genes and environmental factors (GxE interactions). Identification of epistatic interactions may be key in understanding complex diseases such as CVD as epistasis is one of several factors contributing to the missing heritability [45]. In support of epistasis an interaction between *HMGCR* and *LIPC* with an effect on HDL cholesterol levels has been described and validated in a number of independent multiple-ethnic populations [46], which provides insight into the genetic architecture of this complex

trait. Also with regard to HDL, use of biobanked samples and electronic medical records in a novel approach identified 11 epistatic interactions, eight of which were validated in a second cohort [47].

In parallel, the identification of GxE interactions is particularly important for CVD because such strong non-genetic risk factors as diet and physical activity have roles in this major preventable disease. However, GxE interactions consider only single variants. Combining multiple loci with modest effect into a global genetic risk score (GRS) is thought to improve identification of individuals at risk of developing CVD and individual risk assessment. The use of a GRS in several prospective cohort studies improved the ability to predict incident CHD beyond that afforded by traditional non-genetic risk factors [48,49], or modestly improved CHD risk prediction in the Atherosclerosis Risk in Communities study but not in the Rotterdam and Framingham Offspring studies [50]. Thus, application of the GRS to environmental interactions should be at the forefront in attempts to characterize GRS-CVD connections with great potential in the implementation of personalized nutrition and health monitoring. This is supported by several studies assessing GRSxE interactions related to CHD or CVD risk factors including diabetes and obesity.

With regard to diet, a Western dietary pattern interacted with a GRS based on ten diabetes risk SNPs in a case-control study from the Health Professionals Follow-Up Study (HPFS). Intakes of the Western dietary pattern were significantly associated with increased diabetes risk among men with a higher GRS, but not among those with a lower GRS [51]. Sugar-sweetened beverage intake interacted with a GRS, calculated on the basis of 32 BMI-associated loci, in relation to BMI and incident obesity in the Nurses' Health Study and the HPFS, with subsequent replication. The genetic association with adiposity was more pronounced with greater intake of sugar-sweetened beverages [52]. Recently, association analyses were tested between a GRS comprising 13 obesity-related SNPs and dietary intakes of fat, carbohydrates, protein, fiber and total energy intake, as well as interactions on BMI and related traits among non-diabetic Malmö Diet and Cancer Study participants. Although no significant interactions were observed between GRS and total energy intake or macronutrient intakes on BMI or any associated traits, individuals with higher GRS showed lower total energy intake and higher intake levels of protein and fiber [53]. In relation to physical activity, significant interactions were observed between a 12-SNP obesity GRS with physical activity on obesity risk and BMI over time in the European Prospective Investigation of Cancer Norfolk cohort. Living a physically active lifestyle was associated with a 40% reduction in the genetic predisposition to common obesity [54]. These interactions were replicated in a meta-analysis of 11 cohorts of European ancestry, although these effects appear to stem from the North American cohorts [55]. In two large prospective cohorts of US populations, TV watching and physical activity interacted significantly with an obesity GRS on 32 SNPs, whereas prolonged TV watching may accentuate the predisposition to elevated adiposity, greater leisure time physical activity may attenuate the genetic association [56]. In sum, the application of a GRS as a type of network for a given trait is a developing area of CVD research.

Networks of gene-environment interactions

Much more experimental is the application of systems biology to gene-environment interactions resulting in biological networks based on either individual CVD-related traits or a single environmental factor that modifies the genotype-phenotype association. We have described such an approach with cardiometabolic traits, GxE interactions and PPARA and PPARG networks [57]. Networking individual GxE interactions based on a common attribute holds promise in characterizing how the genome interprets a particular dietary

macronutrient with associative effects on measures of CVD risk. That such networks could incorporate allele-specific edges between the biological entities is of particular interest.

Nutrition, exercise and pharmaceutical agents modulate gene networks

In applying a network perspective to disease, it follows that disease can arise from dysfunctional regulatory networks, which themselves are the result of altered edges between nodes [58]. These nodes can represent any of a variety of biological or physiological entities (Figure 1) and insults to the network away from homeostasis, such as a non-ideal diet or exercise regimen, have the effect of promoting metabolic syndrome and leading to heart failure [59]. However, progression from an excess of calories in the diet and a lack of physical activity to atherosclerosis and cardiac arrest is in part determined by nodes (eg, genetic polymorphisms) and edges (eg, unique relationships among the affected systems) that are personal to a given individual. Progression to heart failure is complex at least partially because of the crosstalk of the affected organs, pathways, genes and metabolites [59]. Further complexity arises from the personal nature of relationships within the nodes of the heart failure network. Genetic variation, lifestyle choices (including diet and pharmaceuticals) and the GxE and epistatic interactions all contribute to the flux through the network [60]. This is based on a large body of work showing that disease progression, even in twin studies, is not uniform. At the same time, while homeostasis is a powerful force, intervention – with a proper and healthy diet or an active and regular exercise regimen – will have different effects on different persons. Thus, quantification of the relationships (ie, edges) between disease-relevant entities and their composite parts can contribute to defining first, all the nodes within a CVD network and second, the personal network attributes unique to an individual.

Conclusion

Indeed, it may present both mental and computational challenges to consider networks for cardiovascular health and cardiovascular diseases that contain entities ranging from genes, proteins, lipids, small molecules including dietary components and drugs, and gut microbial species numbers, and all overlaid with genetic variation. However, this is indeed the direction in which much research is advancing. Large amounts of data are collected and integrated for the purpose of depicting functional relationships and pinpointing new actors. The topics broached in this short review are indicative of the complexities of cardiovascular diseases, but with that complexity comes the challenge to integrate these data in a way that highlights the relationships among entities and to leverage such to identify novel CVD factors as either new relationships among known entities or newly identified as relevant to CVD.

Acknowledgments

This work is supported in part by National Institutes of Health (5R21HL114238-02) to LDP; National Institutes of Health (1R21AR055228-01A1, HL54776), the National Institute of Diabetes and Digestive and Kidney Diseases (DK075030) and the US Department of Agriculture Research Service (53-K06-5-10 and 58-1950-9-001) to JMO. This research has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. PIOF-GA-2010-272581 to PC-A. Tufts Center for Neuroscience Research P30 NS047243 provided support to LKI. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. The USDA is an equal opportunity provider and employer.

References

1. Rhinn H, Fujita R, Qiang L, Cheng R, Lee JH, Abeliovich A. Integrative genomics identifies APOE ε4 effectors in Alzheimer's disease. *Nature*. 2013; 500:45–50. [PubMed: 23883936]

2. Erbilgin A, Civelek M, Romanoski CE, Pan C, Hagopian R, Berliner JA, Lusis AJ. Identification of CAD candidate genes in GWAS loci and their expression in vascular cells. *J Lipid Res.* 2013; 54:1894–905. [PubMed: 23667179]
3. Zhang X, Johnson AD, Hendricks AE, Hwang SJ, Tanriverdi K, Ganesh SK, et al. Genetic associations with expression for genes implicated in GWAS studies for atherosclerotic cardiovascular disease and blood phenotypes. *Hum Mol Genet.* 2013 in press. 10.1093/hmg/ddt461
4. Stranger BE, Raj T. Genetics of human gene expression. *Curr Opin Genet Dev.* 2013; 23:627–34. [PubMed: 24238872]
5. Kathiresan S, Melander O, Guiducci C, Surti A, Burt NP, Rieder MJ, et al. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat Genet.* 2008; 40:189–97. [PubMed: 18193044]
6. 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.* 2010; 466:714–9. [PubMed: 20686566]
7. Parnell LD, Lee YC, Lai CQ. Adaptive genetic variation and heart disease risk. *Curr Opin Lipidol.* 2010; 21:116–22. [PubMed: 20154611]
8. Westra HJ, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet.* 2013; 45:1238–43. [PubMed: 24013639]
- 9•. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet.* 2013; 45:580–5. This project will make significant contributions to describing genes whose expression in any of >40 tissues is regulated in part by common genetic variation. [PubMed: 23715323]
10. Rung J, Cauchi S, Albrechtsen A, Shen L, Rocheleau G, Cavalcanti-Proença C, et al. Genetic variant near IRS1 is associated with type 2 diabetes, insulin resistance and hyperinsulinemia. *Nat Genet.* 2009; 41:1110–5. [PubMed: 19734900]
11. Richardson K, Lai CQ, Parnell LD, Lee YC, Ordovas JM. A genome-wide survey for SNPs altering microRNA seed sites identifies functional candidates in GWAS. *BMC Genomics.* 2011; 12:504. [PubMed: 21995669]
12. Najafi-Shoushtari SH, Kristo F, Li Y, Shioda T, Cohen DE, Gerszten RE, Näär AM. MicroRNA-33 and the SREBP host genes cooperate to control cholesterol homeostasis. *Science.* 2010; 328:1566–9. [PubMed: 20466882]
13. Rayner KJ, Suárez Y, Dávalos A, Parathath S, Fitzgerald ML, Tamehiro N, et al. miR-33 contributes to the regulation of cholesterol homeostasis. *Science.* 2010; 328:1570–3. [PubMed: 20466885]
14. Allen RM, Marquart TJ, Albert CJ, Suchy FJ, Wang DQ, Ananthanarayanan M, et al. miR-33 controls the expression of biliary transporters, and mediates statin- and diet-induced hepatotoxicity. *EMBO Mol Med.* 2012; 4:882–95. [PubMed: 22767443]
15. Dávalos A, Goedeke L, Smibert P, Ramírez CM, Warriar NP, Andreo U, et al. miR-33a/b contribute to the regulation of fatty acid metabolism and insulin signaling. *Proc Natl Acad Sci U S A.* 2011; 108:9232–7. [PubMed: 21576456]
- 16•. Rottiers V, Obad S, Petri A, McGarrah R, Lindholm MW, Black JC, et al. Pharmacological Inhibition of a MicroRNA Family in Nonhuman Primates by a Seed-Targeting 8-Mer AntimiR. *Sci Transl Med.* 2013; 5:212ra162. Several groups have been investigating gene-microRNA-cholesterol networks. This report describes the latest in exploiting those networks for therapeutic purposes.
17. Richardson K, Nettleton JA, Rotllan N, Tanaka T, Smith CE, Lai CQ, et al. Gain-of-function lipoprotein lipase variant rs13702 modulates lipid traits through disruption of a microRNA-410 seed site. *Am J Hum Genet.* 2013; 92:5–14. [PubMed: 23246289]
18. Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, Dugar B, et al. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature.* 2011; 472:57–63. [PubMed: 21475195]
19. Bennett BJ, de Aguiar Vallim TQ, Wang Z, Shih DM, Meng Y, Gregory J, et al. Trimethylamine-N-oxide, a metabolite associated with atherosclerosis, exhibits complex genetic and dietary regulation. *Cell Metab.* 2013; 17:49–60. [PubMed: 23312283]

20. Koeth RA, Wang Z, Levison BS, Buffa JA, Org E, Sheehy BT, et al. Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nat Med.* 2013; 19:576–85. [PubMed: 23563705]
21. Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, Liu Y, et al. HMDB 3.0--The Human Metabolome Database in 2013. *Nucleic Acids Res.* 2013; 41:D801–7. [PubMed: 23161693]
22. Bouatra S, Aziat F, Mandal R, Guo AC, Wilson MR, Knox C, et al. The human urine metabolome. *PLoS One.* 2013; 8:e73076. [PubMed: 24023812]
- 23•. Trupp M, Zhu H, Wikoff WR, Baillie RA, Zeng ZB, Karp PD, et al. Metabolomics reveals amino acids contribute to variation in response to simvastatin treatment. *PLoS One.* 2012; 7:e38386. Metabolomics is an expanding wing of omics research. This study showcases the potential benefits of integrating small molecule measures in a study of a commonly prescribed hypolipidemic therapeutic. [PubMed: 22808006]
24. Magnusson M, Lewis GD, Ericson U, Orho-Melander M, Hedblad B, Engström G, et al. A diabetes-predictive amino acid score and future cardiovascular disease. *Eur Heart J.* 2013; 34:1982–9. [PubMed: 23242195]
25. Holzer RG, Park EJ, Li N, Tran H, Chen M, Choi C, et al. Saturated fatty acids induce c-Src clustering within membrane subdomains, leading to JNK activation. *Cell.* 2011; 147:173–84. [PubMed: 21962514]
26. Liu HY, Collins QF, Moukdar F, Zhuo D, Han J, Hong T, et al. Suppression of hepatic glucose production by human neutrophil alpha-defensins through a signaling pathway distinct from insulin. *J Biol Chem.* 2008; 283:12056–63. [PubMed: 18347011]
27. Khaliullina H, Bilgin M, Sampaio JL, Shevchenko A, Eaton S. Lipoproteins carry endocannabinoids that inhibit the Hedgehog pathway. *bioRxiv.* 201310.1101/000570
28. Teperino R, Amann S, Bayer M, McGee SL, Loipetzberger A, Connor T, et al. Hedgehog partial agonism drives Warburg-like metabolism in muscle and brown fat. *Cell.* 2012; 151:414–26. [PubMed: 23063129]
29. Thomas G, Betters JL, Lord CC, Brown AL, Marshall S, Ferguson D, et al. The serine hydrolase ABHD6 is a critical regulator of the metabolic syndrome. *Cell Rep.* 2013; 5:508–20. [PubMed: 24095738]
30. McCurley A, Pires PW, Bender SB, Aronovitz M, Zhao MJ, Metzger D, et al. Direct regulation of blood pressure by smooth muscle cell mineralocorticoid receptors. *Nat Med.* 2012; 18:1429–33. [PubMed: 22922412]
31. Newfell BG, Iyer LK, Mohammad NN, McGraw AP, Ehsan A, Rosano G, et al. Aldosterone regulates vascular gene transcription via oxidative stress-dependent and -independent pathways. *Arterioscler Thromb Vasc Biol.* 2011; 31:1871–80. [PubMed: 21617142]
32. Schnoes KK, Jaffe IZ, Iyer L, Dabreo A, Aronovitz M, Newfell B, et al. Rapid recruitment of temporally distinct vascular gene sets by estrogen. *Mol Endocrinol.* 2008; 22:2544–56. [PubMed: 18787042]
33. Howard BV, Rossouw JE. Estrogens and cardiovascular disease risk revisited: the Women's Health Initiative. *Curr Opin Lipidol.* 2013; 24:493–9. [PubMed: 24184944]
34. Quehenberger O, Dennis EA. The human plasma lipidome. *N Engl J Med.* 2011; 365:1812–23. [PubMed: 22070478]
35. Kind T, Liu KH, Lee do Y, DeFelice B, Meissen JK, Fiehn O. LipidBlast in silico tandem mass spectrometry database for lipid identification. *Nat Methods.* 2013; 10:755–8. [PubMed: 23817071]
- 36•. Weir JM, Wong G, Barlow CK, Greeve MA, Kowalczyk A, et al. Plasma lipid profiling in a large population-based cohort. *J Lipid Res.* 2013; 54:2898–908. The identified relationships between different lipids and disease, sex and age do much to demystify the complexity of this class of molecule. [PubMed: 23868910]
37. Tracy RP, Lemaitre RN, Psaty BM, Ives DG, Evans RW, Cushman M, et al. Relationship of C-reactive protein to risk of cardiovascular disease in the elderly. Results from the Cardiovascular Health Study and the Rural Health Promotion Project. *Arterioscler Thromb Vasc Biol.* 1997; 17:1121–7. [PubMed: 9194763]
38. Obin M, Parnell LD, Ordovas JM. The emerging relevance of the gut microbiome in cardiometabolic health. *Curr Cardiovasc Risk Rep.* 2013; 7:425–426.

39. Jacobs AS, Ayinde HO, Lee DL. Inflammatory Biomarkers and Cardiovascular Complications in Sickle Cell Disease: A Review. *Curr Cardiovasc Risk Rep.* 2013; 7:368–377.
40. Ordovás JM, Robertson R, Cléirigh EN. Gene-gene and gene-environment interactions defining lipid-related traits. *Curr Opin Lipidol.* 2011; 22:129–36. [PubMed: 21311326]
41. Ordovas-Montanes JM, Ordovas JM. Cholesterol, Inflammasomes, and Atherogenesis. *Curr Cardiovasc Risk Rep.* 2012; 6:45–52. [PubMed: 22368729]
42. Weber C, Noels H. Atherosclerosis: current pathogenesis and therapeutic options. *Nat Med.* 2011; 17:1410–22. [PubMed: 22064431]
43. Metabolic Syndrome ePoster. *Nat Med.* 2011; 17
44. McManus DD, Beaulieu LM, Mick E, Tanriverdi K, Larson MG, Keaney JF Jr, et al. Relationship among circulating inflammatory proteins, platelet gene expression, and cardiovascular risk. *Arterioscler Thromb Vasc Biol.* 2013; 33:2666–73. [PubMed: 23968978]
45. Okser S, Pahikkala T, Aittokallio T. Genetic variants and their interactions in disease risk prediction – machine learning and network perspectives. *BioData Min.* 2013; 6:5. [PubMed: 23448398]
46. Ma L, Brautbar A, Boerwinkle E, Sing CF, Clark AG, Keinan A. Knowledge-driven analysis identifies a gene-gene interaction affecting high-density lipoprotein cholesterol levels in multi-ethnic populations. *PLoS Genet.* 2012; 8:e1002714. [PubMed: 22654671]
47. Turner SD, Berg RL, Linneman JG, Peissig PL, Crawford DC, Denny JC, et al. Knowledge-driven multi-locus analysis reveals gene-gene interactions influencing HDL cholesterol level in two independent EMR-linked biobanks. *PLoS One.* 2011; 6:e19586. [PubMed: 21589926]
48. Ripatti S, Tikkanen E, Orho-Melander M, Havulinna AS, Silander K, Sharma A, et al. A multilocus genetic risk score for coronary heart disease: case-control and prospective cohort analyses. *Lancet.* 2010; 376:1393–400. [PubMed: 20971364]
49. Morrison AC, Bare LA, Chambless LE, Ellis SG, Malloy M, Kane JP, et al. Prediction of coronary heart disease risk using a genetic risk score: the Atherosclerosis Risk in Communities Study. *Am J Epidemiol.* 2007; 166:28–35. [PubMed: 17443022]
50. Brautbar A, Pompeii LA, Dehghan A, Ngwa JS, Nambi V, Virani SS, et al. A genetic risk score based on direct associations with coronary heart disease improves coronary heart disease risk prediction in the Atherosclerosis Risk in communities (ARIC), but not in the Rotterdam and Framingham Offspring, Studies. *Atherosclerosis.* 2012; 223:421–6. [PubMed: 22789513]
51. Qi L, Cornelis MC, Zhang C, van Dam RM, Hu FB. Genetic predisposition, Western dietary pattern, and the risk of type 2 diabetes in men. *Am J Clin Nutr.* 2009; 89:1453–8. [PubMed: 19279076]
52. Qi Q, Chu AY, Kang JH, Jensen MK, Curhan GC, Pasquale LR, et al. Sugar-sweetened beverages and genetic risk of obesity. *N Engl J Med.* 2012; 367:1387–96. [PubMed: 22998338]
53. Rukh G, Sonestedt E, Melander O, Hedblad B, Wirfält E, Ericson U, Orho-Melander M. Genetic susceptibility to obesity and diet intakes: association and interaction analyses in the Malmö Diet and Cancer Study. *Genes Nutr.* 2013; 8:535–47. [PubMed: 23861046]
54. Li S, Zhao JH, Luan J, Ekelund U, Luben RN, Khaw KT, et al. Physical activity attenuates the genetic predisposition to obesity in 20,000 men and women from EPIC-Norfolk prospective population study. *PLoS Med.* 2010; 7:e1000332. [PubMed: 20824172]
55. Ahmad S, Rukh G, Varga TV, Ali A, Kurbasic A, Shungin D, et al. Gene $\times$ physical activity interactions in obesity: combined analysis of 111,421 individuals of European ancestry. *PLoS Genet.* 2013; 9:e1003607. [PubMed: 23935507]
56. Qi Q, Li Y, Chomistek AK, Kang JH, Curhan GC, Pasquale LR, et al. Television watching, leisure time physical activity, and the genetic predisposition in relation to body mass index in women and men. *Circulation.* 2012; 126:1821–7. [PubMed: 22949498]
57. Lee YC, Lai CQ, Ordovas JM, Parnell LD. A Database of Gene-Environment Interactions Pertaining to Blood Lipid Traits, Cardiovascular Disease and Type 2 Diabetes. *J Data Mining Genomics Proteomics.* 2011; 2:106. [PubMed: 22328972]
58. de la Fuente A. From ‘differential expression’ to ‘differential networking’ - identification of dysfunctional regulatory networks in diseases. *Trends Genet.* 2010; 26:326–33. [PubMed: 20570387]

59. Lusis AJ, Attie AD, Reue K. Metabolic syndrome: from epidemiology to systems biology. *Nat Rev Genet.* 2008; 9:819–30. [PubMed: 18852695]
60. Civelek M, Lusis AJ. Systems genetics approaches to understand complex traits. *Nat Rev Genet.* 2013; 15:34–48. This important review explores the means by which complex traits can be mapped and dissected for the quantitative understanding of biology and disease. [PubMed: 24296534]

Fig 1.

General aspects of gene networks. A) Networks are composed nodes and edges. Nodes can be genes, microRNAs, proteins, metabolites, food items, dietary patterns, diseases, or clinical measures of disease risk, among other entities. Edges are the connections between nodes. B) The relationship between two nodes, for example between two genes, can be altered according to the state or condition of a node. For example, different genetic variants (1*, 1**) have different and diverse effects on the nature of the interaction (edge) between node 1 and node 2, which are indicated by edges of different thicknesses or weights. C) Interactions within a network can be depicted in a response-specific manner. Here, several studies conducted in adipose after a high-fat meal have shown that node 3 is known to regulate expression of node 4, increasing its expression, while only one study shows that node 5 responds with increased expression. Thickness of the edge represents the amount of evidence while the coordinated increased expression of nodes 4 and 5 allow us to form an inferred relationship, albeit weak, between those two nodes. D) Known CVD genes (shaded) are networked via a common attribute to other non-CVD genes (unshaded). The more highly linked non-CVD genes become candidates in CVD research. E) Two CVD-relevant pathways are depicted where the shading of the node indicates strength of the relationship to CVD. A node (square) that bridges or links the two pathways becomes a CVD candidate.