Plasma lipoproteins: genetic influences and clinical implications

Robert A. Hegele

Abstract | Susceptibility to the growing global public health problem of cardiovascular disease is associated with levels of plasma lipids and lipoproteins. Several experimental strategies have helped us to clarify the genetic architecture of these complex traits, including classical studies of monogenic dyslipidaemias, resequencing, phenomic analysis and, more recently, genome-wide association studies and analysis of metabolic networks. The genetic basis of plasma lipoprotein levels can now be modelled as a mosaic of contributions from multiple DNA sequence variants, both rare and common, with varying effect sizes. In addition to filling gaps in our understanding of plasma lipoprotein metabolism, the recent genetic advances will improve our ability to classify, diagnose and treat dyslipidaemias.

Lipid
A member of a diverse group of hydrophobic compounds with many biological functions, such as structural components of cell membranes, energy storage sources and intermediates in signalling pathways.

Lipoprotein
A molecular complex containing proteins (apolipoproteins) and lipids (cholesterol or TG), which allow the lipid component to be soluble in plasma. TG-carrying lipoproteins are CMs and VLDL, whereas cholesterol-carrying lipoproteins are LDL and HDL.

Robarts Research Institute and Schulich School of Medicine and Dentistry, University of Western Ontario, 406-100 Perth Drive, London, Ontario, Canada N6A 5K8.
e-mail: hegele@robarts.ca
doi:10.1038/nrg2481

Levels of certain plasma lipids and lipoproteins are key risk factors for cardiovascular disease (CVD). The search for the genetic contributors to variation in plasma lipid and lipoprotein levels began more than 25 years ago, using classical linkage analysis, association studies and animal models[1,2]. Progress was slow and, for some time, underwhelming. Although the genetic basis of most Mendelian dyslipidaemias was solved before the turn of the millennium, the genetic causes of less severe but more common lipoprotein variation, which is closer to the median of the population distribution, remained elusive. In 2004, a small but important group of genetic determinants of lipoprotein variation was uncovered by resequencing genomic DNA from individuals with extreme lipoprotein phenotypes[3]. Since 2008, genome-wide association (GWA) studies have implicated common variants in numerous loci and genes as being the genetic influences underlying the variation that is closer to the population median. Although the clinical application of these new data is uncertain, the biological implications might be far-reaching — several novel pathways involved in lipoprotein metabolism, in mechanisms underlying disease risk and in potential therapeutic strategies have been identified.

This Review focuses on recent discoveries of the genetic determinants of variation in plasma lipoprotein levels in humans and animal models, how these have enhanced existing knowledge and how they have identified new avenues for investigation. These recent genetic advances will broaden our understanding of basic metabolic pathways and will improve classification, diagnosis and treatment strategies.

Lipids and lipoproteins

Lipids. Clinically, the most important plasma lipids are cholesterol and triglyceride (TG; also called triacylglycerol)[4]. Cholesterol has numerous roles: it is a component of cell membranes; the precursor for steroid hormones and vitamin D, and for oxysterols and bile acids (both of which are activators of nuclear hormone receptors involved in sterol metabolism); and it is required for the activation of neuronal signalling molecules[4–6]. Only a small amount of circulating cholesterol originates from the diet; ~80% is derived from endogenous synthesis, for which 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMGCR) catalyses the rate-limiting step. Most cholesterol in the circulation is esterified, with free cholesterol constituting a minor fraction. TG is a key energy source that is made up of free fatty acids (FFAs) that are ester-linked to a glycerol backbone. TG is synthesized in intestinal and liver cells and is then transported through the plasma and, after lipolysis at the endothelial surface, delivers FFA to peripheral cells for β-oxidation or storage.

Lipoproteins. The insolubility of cholesterol and TG in plasma requires that they are transported in spheroidal macromolecules called lipoproteins, which have a hydrophobic core containing phospholipid, fat-soluble antioxidants and vitamins, and cholesteryl ester, and a hydrophilic coat that contains free cholesterol, phospholipid and apolipoprotein molecules. The main TG-carrying lipoproteins are chylomicron (CM) and very low-density lipoprotein (VLDL). The main cholesterol-carrying lipoproteins are low-density lipoprotein (LDL) and high-density lipoprotein (HDL).

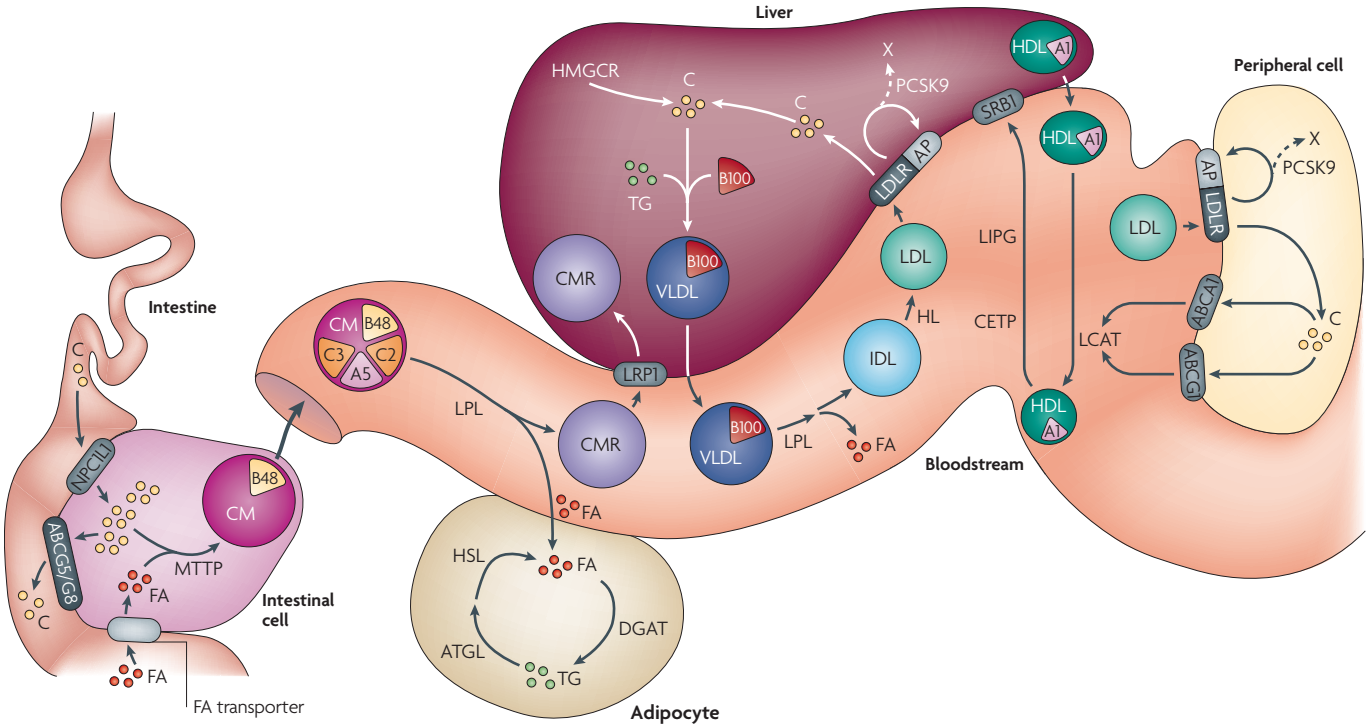


Figure 1 | **Overview of lipoprotein metabolism.** The main lipids in lipoproteins are free and esterified cholesterol (C) and triglyceride (TG). The metabolism of TG, low-density lipoprotein (LDL) cholesterol and high-density lipoprotein (HDL) cholesterol is shown. In TG metabolism, hydrolysed dietary fats enter intestinal cells (enterocytes) via fatty acid (FA) transporters. Reconstituted TG is packaged with C ester and the apolipoprotein B (APOB) isoform B48 into chylomicrons (CMs) by microsomal TG-transfer protein (MTTP) through a vesicular pathway. CMs, secreted via the lymphatic system, enter the vena cava and circulate until they interact with lipoprotein lipase (LPL), the secretion of which depends on lipase-maturation factor 1 (not shown), and which is secured to endothelium by proteoglycans and glycosylphosphatidylinositol-anchored HDL-binding protein 1 (not shown). CMs contain apoliproteins, including APOA5 (A5), APOC2 (C2) and APOC3 (C3). Released free FAs incompletely enter peripheral cells. In adipocytes, enzymes including acyl CoA:diacylglycerol acyltransferase (DGAT) resynthesize TG, which is hydrolysed by adipose TG lipase (ATGL) and hormone sensitive lipase (HSL). CM remnants (CMRs) are taken up by hepatic LDL receptor (LDLR), in the absence of LDLR they are taken up by LDLR-related protein-1 (LRP1). In liver cells (hepatocytes), TG is packaged with cholesterol and the APOB isoform B100 into very low-density lipoprotein (VLDL); the TG contained in VLDL is hydrolysed by LPL, releasing FAs and VLDL remnants (IDL) that are hydrolysed by hepatic lipase (HL), thereby yielding LDL. In LDL cholesterol metabolism, sterols in the intestinal lumen enter enterocytes via the Niemann-Pick C1-like 1 (NPC1L1) transporter and some are resecreted by heterodimeric ATP-binding cassette transporter G5/G8 (ABCG5/G8). In enterocytes, cholesterol is packaged with TG into CM. In hepatocytes, cholesterol is recycled or synthesized *de novo*, with 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMGCR) being rate-limiting. LDL transports cholesterol from the liver to the periphery. LDL is endocytosed by peripheral cells and hepatocytes by LDLR, assisted by an adaptor protein (AP). Proprotein convertase subtilisin/kexin type 9 (PCSK9), when complexed to LDLR, short-circuits recycling of LDLR from the endosome, leading to its degradation (X). In HDL cholesterol metabolism, HDL, via APOA-I (A1), mediates reverse cholesterol transport by interacting with ATP-binding cassette A1 (ABCA1) and ABCG1 transporters on non-hepatic cells. Lecithin-cholesterol acyltransferase (LCAT) esterifies cholesterol so it can be used in HDL cholesterol, which, after remodelling by cholesterol ester transfer protein (CETP) and by endothelial lipase (LIPG), enters hepatocytes via scavenger receptor class B type I (SRB1).

Dyslipidaemia
A group of biochemical disorders characterized by quantitative disturbances in plasma lipids and lipoproteins, usually defined by deviations from age- and sex-specific normal ranges. Dyslipidaemia is a risk factor for CVD, such as stroke or heart attack.

Lipoproteins are distinguished from each other by size, density, electrophoretic mobility, composition and function (Supplementary information S1 (table)). Lipoprotein metabolism is a complex network of assembly, secretion, processing and catabolism (FIG. 1). This Review examines genetic influences on the level of the plasma lipoproteins LDL and HDL and of the plasma lipid TG, which have been connected to clinical disorders, especially CVD.

Plasma lipids and lipoproteins and CVD risk. Chronic excess of plasma LDL interferes with arterial relaxation, and LDL particles that are not catabolized through regulated receptor-mediated endocytosis are recognized by scavenger receptors on arterial-wall macrophages[7,8]. Lipids that are engulfed by macrophages become oxidized, generating toxic intermediates, which induce cytokine production and chemotaxis of inflammatory cells[7,8]. LDL-loaded arterial-wall macrophages become

Box 1 | **Phenotyping of dyslipidaemia**

Dyslipidaemia is broadly defined as deviation of one or more plasma lipoprotein traits across a biochemical threshold, often an age- and sex-defined percentile value. Although non-genetic factors, such as an imbalance between caloric intake and expenditure, excessive alcohol intake, diabetes and certain medications, are associated with dyslipidaemia, genetic susceptibility must also be important as not all individuals who are exposed equally to these stressors develop dyslipidaemia.

In the pre-genomic era, a commonly used system for classifying dyslipidaemia — the Fredrickson or World Health Organization International Classification of Diseases (WHO ICD) lipoprotein phenotypes — was based on patterns of lipoprotein fractions seen in families with hyperlipoproteinaemia (HLP; summarized in TABLE 1). Although patient subgroups in particular HLP types have discrete monogenic diseases, most display only the biochemical deviation, without syndromic manifestations, until cardiovascular disease onset. No WHO ICD HLP phenotype is primarily characterized by deviations of HDL cholesterol. By contrast, five of the six phenotypes include elevated triglyceride (TG) in their definitions. Although it has served an invaluable role over the last 40 years, recent studies have simply evaluated LDL and HDL cholesterol and TG as continuous variables[20–24], without reference to WHO ICD HLP phenotype.

foam cells, which are building blocks for atherogenic plaques. Occlusive plaques, often the result of atherogenic plaque rupture compounded by thrombosis, result in CVD, such as coronary heart disease (CHD) or stroke[7,8].

Most plasma lipids and lipoproteins follow a right-skewed normal distribution in the general population. Median levels vary by age and sex: in general, older age and male sex are associated with a less favourable lipid profile. There are also differences in median levels across geographical ancestries, and these might contribute to differences in CVD risk. CVD is directly related to the total plasma level of cholesterol and the level of LDL cholesterol[9], and is inversely related to the level of plasma HDL cholesterol[10].

In men, a rise in total cholesterol from 5.2 to 6.2 mmol L^{-1} (200 to 240 mg dL^{-1}) is associated with a threefold increased risk of death from CHD[11]. Lowered levels of HDL cholesterol to <0.9 mmol L^{-1} (<35 mg dL^{-1}) is an independent CHD risk factor and is the most common lipid disturbance seen in CHD patients under 60 years of age[12]. But although the inverse association between HDL levels and CVD is indisputable, a causal relationship remains uncertain, partly because not all genetic disorders causing very low HDL levels are associated with CVD. This is in contrast to the uniform presence of CVD in genetic disorders that cause high LDL levels. Plasma LDL cholesterol and apolipoprotein B concentration are directly correlated[13], as are plasma HDL cholesterol and apolipoprotein A-I concentration[14] (see Supplementary information S1 (table) for a summary of apolipoproteins). The ratio of apolipoprotein B to apolipoprotein A-I has been advocated as the strongest predictor of CHD risk[15], although the ratio of total cholesterol to HDL cholesterol seems to be equally predictive[16]. The relationship between plasma TG and CHD has been confounded by the association of elevated TG with depressed HDL cholesterol. Nonetheless, elevated TG (especially non-fasting TG[17,18]) and familial hypertriglyceridaemias[19] are independently associated with CHD risk.

Owing to the importance of plasma lipoprotein levels to CHD risk, dissecting the genetic determinants of these levels has been the focus of extensive research. The main strategies are discussed in the following sections, with GWA studies from 2008 (REFS 20–24) and 2009 (REFS 25–27) having provided particularly significant advances.

Early studies of plasma lipoprotein variation

From approximately 1982 to 2004, small case–control or cohort-based association studies and linkage studies of discrete and quantitative lipoprotein traits using candidate genes or genome-wide marker sets were the principal tools used to study genetic factors influencing the level of plasma lipoproteins[1]. Hundreds of such studies were reported; the results of most of them have never been replicated. Some early candidate gene studies hinted at the polygenic nature of these traits, showing contributions of multiple small-effect gene variants[28]. Also, some linkage peaks from these early studies have subsequently helped to map important genes and loci, including upstream regulatory factor 1 (*USF1*) as a determinant of hyperlipoproteinaemia (HLP) type 2B[29] (BOX 1), and more recently WW-domain-containing oxidoreductase (*WWOX*) as a determinant of HDL cholesterol[30]. However, convincing data are lacking for significant metabolic roles for *USF1*, *WWOX* or numerous other genes found this way. Furthermore, most such genes have not appeared in the results from recent genome-wide scans[20–27,31].

Despite the substantial investment of time and resources, only a few of the markers that were discovered in these studies have been replicated. Recent meta-analyses do, however, support the association of variation in LDL cholesterol with apolipoprotein E (*APOE*)[32], and of HDL cholesterol with cholesteryl ester transfer protein (*CETP*)[33]. Also, the chromosome 11 *APOA1–C3–A4–A5* gene cluster has been consistently associated with plasma TG variation[34]. Their limitations notwithstanding[35], carefully planned linkage and association studies will continue to complement more recently applied approaches and methods, such as resequencing, genome-wide studies and animal models, which are discussed in subsequent sections of this Review.

Monogenic disorders of lipoprotein metabolism

Several important genetic determinants of lipoprotein levels have been identified through studies of monogenic disorders. Monogenic disorders comprise a rare patient subgroup that is found at the extremes of population-specific lipoprotein distributions. Careful extensive phenotypic evaluation, sometimes called phenomics, allows the phenotypic and molecular genetic attributes of these disorders to be evaluated together (FIG. 2). The molecular basis of many monogenic high- and low-lipoprotein syndromes was elucidated using biochemical approaches or classical linkage analysis. Studies of these diseases have helped to define key pathways, such as receptor-mediated endocytosis through the LDL receptor[4] and sterol efflux from cells through ATP-binding cassette proteins[36]. Several causative genes with rare mutations have resurfaced as loci with common small-effect genetic variants that underlie lipoprotein variation.

Genome-wide association (GWA) study
A powerful experimental approach for gene mapping that uses SNP markers across the human genome to identify genetic regions that are statistically associated with quantitative or qualitative traits in samples of unrelated individuals.

Apolipoprotein
The specific name for the protein component of a lipoprotein. There are at least 13 different apolipoproteins, which have a variety of roles, including as enzyme cofactors and receptor ligands.

Phenomics
The objective and systematic acquisition of high-quality phenotypic data (that is, deep phenotyping), allowing for phenotypic features to be analysed on a continuum together with molecular data, such as gene expression profiles or causative genomic mutations.

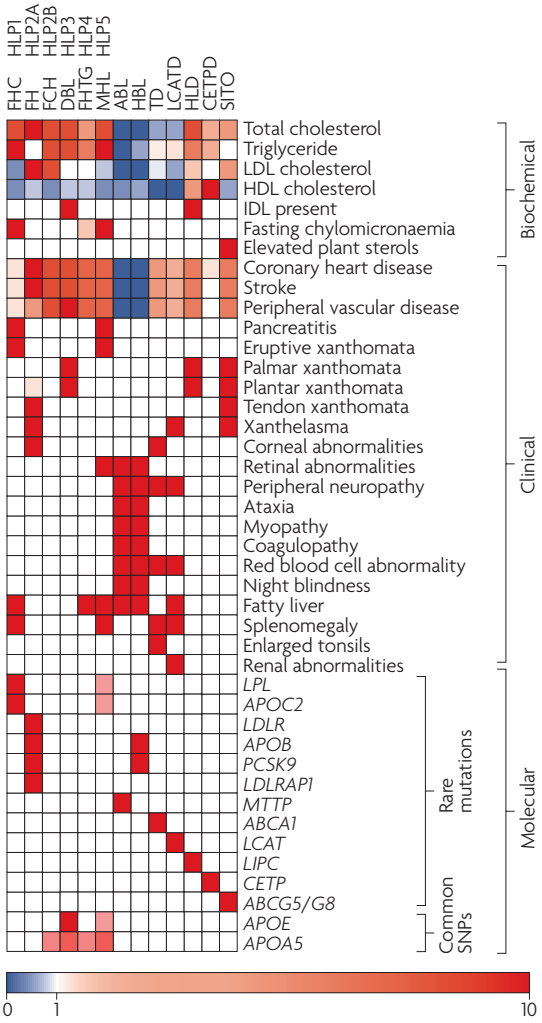


Figure 2 | **Phenomic and molecular description of selected dyslipidaemias.** Defining features and dyslipidaemia names are listed in rows and columns, respectively. For biochemical traits, the fold-change from normal is indicated by the colour intensity from the legend: white indicates no difference from normal, red indicates a fold-increase above normal (maximum 10) and blue a fold-decrease below normal (minimum 0). For susceptibility to coronary heart disease, stroke and peripheral vascular disease, white indicates no difference from normal, red the fold-increase risk above normal and blue the fold-decrease risk below normal. For qualitative clinical features, white indicates the normal state (absence of feature) and red indicates presence of the feature. For rare mutations, white indicates no role, red a major etiologic role, and pink a minor etiologic role for the gene. For common polymorphisms, white indicates no role, and gradations of red reflect risk associated with the genotype. *ABCA1*, ATP-binding cassette A1; *ABCG5/G8*, ATP-binding cassette transporter G5/G8; ABL, abetalipoproteinaemia; *APO*, apolipoprotein; *CETP*, cholesteryl ester transfer protein; *CETPD*, CETP deficiency; DBL, dysbetalipoproteinaemia (also called HLP3); FCH, familial combined hyperlipidaemia (also called HLP2B); FH, familial hypercholesterolaemia (a subset of HLP2A); FHC, familial hyperchylomicronaemia (essentially identical to HLP1); FHTG, familial hypertriglyceridaemia (also called HLP4); HBL, hypobetalipoproteinaemia; HDL, high-density lipoprotein; HLD, hepatic lipase deficiency; HLP, hyperlipoproteinaemia type; IDL, intermediate-density lipoprotein; *LCAT*, lecithin cholesterol acyltransferase; LCATD, LCAT deficiency; LDL, low-density lipoprotein; *LDLR*, lipoprotein lipase receptor; *LDLRAP1*, low-density lipoprotein receptor adaptor protein 1; *LIPC*, hepatic lipase; *LPL*, lipoprotein lipase; MHL, mixed hyperlipidaemia (also called severe hypertriglyceridaemia, a subset of HLP5); *MTTP*, microsomal triglyceride transfer protein; *PCSK9*, proprotein convertase subtilisin/kexin type 9; SITO, sitosterolaemia; TD, Tangier disease.

Altered levels of LDL cholesterol. Some of the classical World Health Organization (WHO) HLP phenotypes have a monogenic basis (BOX 1; TABLE 1). For instance, HLP type 2A is defined by LDL cholesterol above the ninety-fifth percentile of the normal population distribution, but ~10% of these subjects have a discrete monogenic syndrome[36]. Such syndromes include heterozygous familial hypercholesterolaemia (FH), or the phenotypically similar familial defective apolipoprotein B and autosomal dominant hypercholesterolaemia, which is due to gain of function in the convertase gene *PCSK9*. Homozygous FH is a much rarer and more severe autosomal recessive condition (TABLE 1). These monogenic syndromes are distinguished from general HLP type 2A by the severity of LDL cholesterol elevation, by an autosomal inheritance pattern, by lipid deposits in the eyes, skin and tendons, and by markedly elevated CVD risk. Most HLP type 2A patients have no defined monogenic defect. Interestingly, different homozygous loss-of-function (LOF) mutations in the *APOB* or *PCSK9* genes cause a monogenic syndrome called homozygous hypobetalipoproteinaemia (HHBL), in which almost no LDL cholesterol is present. Homozygous mutations in microsomal triglyceride transfer protein (*MTTP*), cause a similar disease called abetalipoproteinaemia (ABL)[37].

Altered levels of HDL cholesterol. Monogenic disorders have also been identified that alter levels of HDL cholesterol. For example, a few individuals with a HDL cholesterol level below the fifth percentile have extremely rare monogenic disorders including Tangier disease (TD), which is due to homozygous mutations in the ATP-binding cassette gene *ABCA1* (REF. 38) or homozygous deficiencies of apolipoprotein A-I or lecithin-cholesterol acyltransferase (*LCAT*)[39]. A few individuals with HDL cholesterol levels above the ninety-fifth percentile have homozygous deficiencies of *CETP* or hepatic lipase (*LIPC*)[40].

Altered TG levels. Finally, a few individuals with plasma TG levels above the ninety-fifth percentile have rare monogenic disorders resulting from homozygous LOF mutations, including in the lipoprotein lipase (*LPL*), *APOC2* and *APOA5* genes[41–44]. Gene therapy with *LPL* can benefit patients with LPL deficiency[45]. The frequency of individuals with these elevated TG syndromes is less than 1 in 100,000 and they fit the definition of

Table 1 | **Classical hyperlipoproteinaemia phenotypes**

WHO ICD and OMIM numbers	Frederickson HLP phenotype	Elevated lipid	Elevated lipoprotein	Genetics
E78.3 238600	HLP type 1 (also known as familial chylomicronaemia or LPL deficiency)	TG	CM	Primarily paediatric presentation, but also in young adults; monogenic; autosomal recessive form is due to mutant *LPL* or *APOC2*; other forms involve mutant *APOA5*, *LMF1* and possibly *GPIHBP1*
E78.0 143890	HLP type 2A (also known as heterozygous and homozygous familial hypercholesterolaemia); severe molecular subtypes	TC (top fifth percentile)	LDL (top fifth percentile)	Many forms are polygenic, ~10% are monogenic; heterozygous form is due to mutant *LDLR*, *APOB* or *PCSK9*; homozygous form is due to mutant *LDLR* or *LDLRAP1*
E78.4 144250	HLP type 2B (also known as combined hyperlipoproteinaemia)	TC, TG (both top fifth percentile)	VLDL, LDL (top fifth percentile)	Polygenic, multiple etiologies; some cases are due to *USF1*, *APOB* or *LPL*; ~35% of subjects show *APOA5* S19W or –1131T>C
E78.2 107741	HLP type 3 (also known as dysbetalipoproteinaemia)	TC, TG (both top fifth percentile)	IDL	Polygenic; mutant *APOE* or homozygosity for E2 allele of *APOE* is necessary but not sufficient; dominant form is due to mutant *APOE*; ~40% of subjects show *APOA5* S19W or –1131T>C
E78.1 144600, 145750	HLP type 4 (also known as primary hypertriglyceridaemia)	TG (top fifth percentile)	VLDL	Polygenic; ~35% of subjects show *APOA5* S19W or –1131T>C
E78.3 144650	HLP type 5 (also known as mixed hyperlipidaemia)	TC, TG (both top fifth percentile)	VLDL, CM	Polygenic; ~10% of cases have mutant *LPL*, *APOC2* and *APOA5*; ~55% of subjects show *APOA5* S19W or –1131T>C; also small effects from *APOE*, *TRIB1*, *CHREBP*, *GALNT2*, *GCKR* and *ANGPTL3*

–1131T>C, a T to C conversion at position –1131; *ANGPTL3*, angiopoietin-like 3; *APOA5*, apolipoprotein A-V; *APOB*, apolipoprotein B; *APOC2*, apolipoprotein C-II; *APOE*, apolipoprotein E; *CHREBP*, carbohydrate response element binding protein (also known as *MLXIPL*); CM, chylomicron; *GALNT2*, UDP-*N*-acetyl-alpha-*D*-galactosamine:polypeptide *N*-acetylgalactosaminyltransferase 2; *GCKR*, glucokinase regulator; HLP, hyperlipoproteinaemia; ICD, International Classification of Diseases; IDL, intermediate-density lipoprotein; LDL, low-density lipoprotein; *LDLRAP1*, LDL receptor adaptor protein 1 (also known as *ARH*); LPL, lipoprotein lipase; OMIM, Online Mendelian Inheritance in Man; *PCSK9*, proprotein convertase subtilisin/kexin type 9; S19W, serine to tryptophan conversion at amino acid 19; TG, triglyceride; *TRIB1*, tribbles homologue 1 (*Drosophila*); *USF1*, upstream transcription factor 1; VLDL, very low-density lipoprotein; WHO, World Health Organization.

HLP type 1, with skin and eye abnormalities and often pancreatitis[41]. Very low TG is found in patients with ABL (homozygous *MTTP* mutations), in some patients with HHBL (homozygous *APOB* mutations)[46] and in patients with CM retention disease (homozygous mutations in the *SAR1* homologue *SAR1B*)[47].

Identifying genes that cause monogenic dyslipoproteinaemias has furthered our understanding of which proteins act together to regulate LDL and HDL cholesterol and TG metabolism (FIG. 1). For instance, plasma LDL cholesterol levels depend crucially on LDL receptor function, which in turn requires proper binding of apolipoprotein B, the presence of LDL receptor accessory protein 1 (LDLRAP1) as a chaperone through the early phase of endocytosis, and regulated intracellular receptor degradation by the convertase PCSK9. Despite the fairly complete understanding of monogenic dyslipoproteinaemias, some patients do not have mutations in any known genes, suggesting the presence of other causative genes that, once identified, might in turn illuminate even newer metabolic pathways.

Resequencing tails of the lipoprotein distribution

High-throughput resequencing of larger numbers of selected individuals, such as those outside the threshold values of lipoprotein traits, has been used to detect mutations in candidate genes[48]. An excess of mutations among individuals at one extreme of the distribution compared with individuals at the opposite extreme is considered to be strong evidence for genetic association[49]. Most candidate genes so far evaluated by resequencing were selected because they caused monogenic disorders involving the trait of interest.

For instance, resequencing of *ABCA1*, *APOA1* and *LCAT* in individuals from a population-based study with HDL cholesterol levels below the fifth percentile (cases) or above the ninety-fifth percentile (controls) showed that heterozygous LOF mutations were eight times more common in cases than in controls (16% versus 2%)[3]. Similarly, an excess of *PCSK9* LOF mutations was found in subjects with low LDL cholesterol compared with those with high LDL cholesterol (2% versus 0.1%)[50]. Furthermore, an excess of mutations in *NPC1L1*, which encodes a Niemann–Pick C1-like protein, was found in subjects with low compared with high intestinal sterol absorption (10% versus 2%)[51]; two-thirds of these mutations had compromised *in vitro* expression[52], confirming that the resequencing strategy identified dysfunctional variants. Other examples include a 50-fold increased odds of rare heterozygous LOF mutations in *LPL*, *APOC2* or *APOA5* genes in patients with severe hypertriglyceridaemia (HTG, also known as HLP type 5; TG > 10 mmoL L^{-1}) compared with controls (10% versus 0.2%)[53]. Rare heterozygous mutations in endothelial lipase (*LIPG*) were detected in a few subjects with high HDL levels[54].

Whole-population samples have also been resequenced, as in the case of angiopoietin-like 4 (*ANGPTL4*), which was selected as a candidate gene because *Angptl4* induced-mutant mice have altered lipoprotein levels[55]. *ANGPTL4* mutations accumulated in subjects at the low end of TG distribution[56], although the middle of the TG distribution also contained patients with rare mutations.

A current limitation is that resequencing can only target candidate genes that have been identified by monogenic disease or GWA studies (described in the following section). However, as whole-genome sequencing becomes a reality, significant clusters of mutations in subjects at distribution extremes will enable the detection of new and unsuspected genes that influence lipoprotein levels.

GWA studies of plasma lipids and lipoproteins

The success of GWA studies has resulted from convergence of several technical advances, which include: the HapMap Project; microarrays that allow detection of 10^5 to over 10^6 SNPs; microarrays with SNPs selected to maximize information using linkage disequilibrium data[57]; and very large sample sizes and replication samples, which require multi-site collaborations. A clear tipping point occurred in late 2007, when the essentially negative GWA analysis performed in participants of the Framingham Heart Study[58] was followed in early 2008 by several positive GWA studies[20–24]. The negative study evaluated ~10^5 SNPs in a single cohort of ~10^3 participants, whereas the positive studies evaluated >3×10^5 SNPs in multiple discovery and replication samples involving >5×10^4 participants. The general implications of GWA studies for identifying the genetic determinants of disease traits, including the technical benefits of large samples and dense SNP arrays, have been discussed in a number of recent papers[59–62].

Between-study similarities have increased confidence in GWA findings, as has the identification by the 2008 GWA studies of loci containing genes that were previously associated with lipoprotein levels. For LDL cholesterol, approximately half of the associated genes found in these studies had been identified previously, for example *APOE*, LDL receptor (*LDLR*), *APOB*, *PCSK9* and *HMGCR*. However, loci containing sortilin 1 (*SORT1*), cartilage intermediate layer protein 2 (*CILP2*), basal cell adhesion molecule (*BCAM*) or the translocase gene *TOMM40* had no prior connection to lipoprotein metabolism but had equally strong associations. For HDL cholesterol, approximately three-quarters of significant SNPs were in loci harbouring known genes such as *CETP*, *LIPC*, *LPL*, *ABCA1*, endothelial lipase (*LIPG*) and *LCAT*. Only *GALNT2* and the *MVK–MMAB* locus had no previous connection to HDL metabolism. For TG, approximately one-third of genes in significantly associated loci were known, such as *APOA5*, *LPL*, *LIPC*, *APOB* and *ANGPTL3*, whereas loci harbouring *CILP2*, *TRIB1*, *GCKR*, *CHREBP* (also called *MLXIPL*) and *GALNT2* had minimal prior connection to TG metabolism. Genes in three of these newly identified loci link TG with carbohydrate metabolism (this is relevant as the most common lipid disturbance observed in diabetes is elevated plasma TG) *GCKR*[21] encodes glucokinase regulatory protein; *CHREBP*[63] encodes carbohydrate response element binding protein, a glucose-responsive transcription factor that is active in hepatic glycolysis, lipogenesis and VLDL secretion; and *GALNT2* encodes UDP-*N*-acetyl-alpha-*D*-galactosamine:polypeptide *N*-acetylgalactosaminyltransferase. Scrutiny of the 2009 GWA studies[25–27] shows only a modest increment in new loci over the 2008 GWA studies[20–24], indicating that the 'low-hanging fruit' has been harvested. Furthermore, despite claims that these loci underlie dyslipidaemias[26], study samples were normolipidaemic; GWA studies in truly dyslipidaemic patients are not yet reported.

Most of the significant loci from GWA studies harbour several genes. Although the signal for a locus might arise from a familiar gene, functional corroborative evidence is still required to prove that this gene is involved. Furthermore, for new loci that harbour several genes, any one might be the functional basis of the association signal. But because GWA studies detected well-known genes with a proven role in lipoprotein metabolism, previously unappreciated genes identified through GWA might also be important[60]. Finally, as GWA studies identified several genes that were known to cause monogenic diseases, novel loci might contain rare disease-causing mutations in patients with no mutations in known genes.

But although GWA studies using SNPs have created tremendous excitement, it is important to recognize a few limitations. First, two of these studies used the same sample of diabetic subjects at crucial filtering stages, so the similarity of associations is not that surprising and might be specific to the metabolic context of diabetes[20,24]. Second, because the study samples were not actually dyslipidaemic, some loci underlying more severe traits might have been missed. Third, the effect sizes were small — 5 to 10% cumulatively — suggesting that other genes and loci remain to be found, as the known variants account for only a small portion of lipoprotein variation. Interestingly, many associated regions seem also to be associated with differences in mRNA expression of regional genes, which if confirmed might signal an important new general mechanism to explain the relationship between common SNPs and common traits[26].

Even though SNP studies in lipoprotein metabolism have been informative, a focus on single nucleotides fails to capture the complete range of genetic variability. Recent reports have demonstrated the surprising ubiquity of larger copy number variations (CNVs) in apparently healthy people[64], adding to the complexity of the 'normal' genome, and evidence is emerging that CNV can contribute to complex traits. For instance, disease-causing CNVs in *LDLR* and *LPL* genes are present in a small but important fraction of patients with HLP types 2A and 1, respectively[65]. Also, CNV of the lipoprotein Lp(a) gene (*LPA*) is the key determinant of plasma Lp(a) concentration[65]. As methodology becomes more robust and standardized, CNVs will probably be shown to be determinants of plasma lipoproteins.

Model organisms

Studying lipoprotein metabolism in model organisms has proven useful for identifying new genetic determinants, for investigating established genetic determinants through *in vivo* study of relevant pathways, and for validating the data obtained from human genetic studies[66]. Recently, mapping the defect in the dyslipidaemic *cld* (combined lipase deficiency) mouse strain identified that lipase maturation factor 1 (*LMF1*) is essential for the

Framingham Heart Study
A widely cited longitudinal study of CVD based in Framingham, Massachusetts, that began in 1948 with 5,209 adult subjects and is currently studying its third generation of participants. It has greatly improved our understanding of risk factors for heart disease.

processing and secretion of both LPL and hepatic lipase[42]. In addition, the defect in an induced-mutant dyslipidaemic mouse strain showed that *GPIHBP1*, which encodes an endothelial anchoring protein, is essential for LPL stabilization and intravascular function[67]. A homozygous nonsense mutation in *LMF1* was found in 1 in 11 patients with severe HTG[42], whereas a homozygous missense *GPIHBP1* mutation with minimal *in vitro* evidence for dysfunction[68] was found in only 1 in 600 patients with severe HTG, suggesting that *GPIHBP1* might not be a major cause of human HTG. *APOM* was inferred from both null and transgenic mice to be an important component of HDL and to have a role in atherosclerosis[69]. Manipulation of *Angptl3* and *Angptl4* in mice[70,71] suggested that the products of both genes regulate TG levels; these findings were subsequently confirmed by GWA (for *ANGPTL3*) and resequencing studies (for *ANGPTL4*)[20,24,56].

Many genetic findings involving human lipoproteins have been validated and enhanced through studies of animal models. For instance, after mutations in two genes encoding ATP-binding cassette proteins, *ABCG5* and *ABCG8*, were discovered in patients with sitosterolaemia[72,73], their role in sterol excretion from intestinal cells (enterocytes) was delineated in knockout mice[74]. Also, bioinformatic-based discovery of *NPC1L1* — using a search strategy based on tissue expression and the presence of functional domains — was followed by evaluation of *Npc1l1*-null mice and definitive demonstration of the role of NPC1L1 in intestinal sterol transport and as the target for a new class of drugs that inhibit sterol absorption, of which ezetimibe is the first member[75]. Human mutations in *NPC1L1* were then noted to affect intestinal sterol absorption[52] and the plasma LDL cholesterol response to ezetimibe[76].

Other gene products that were defined functionally using knockout mice or mice with spontaneous mutations include several adipocyte-based genes: *Atgl* (also known as *Pnpla2*), encoding adipose triglyceride lipase, which is involved in the intracellular lipolysis of TG[77]; *Abhd5*, which encodes CGI-58 (REF. 78), the cofactor for ATGL; *Lpin1,* encoding lipin 1, which directs lipids to adipose storage sites[79]; and *Plin*, encoding perilipin, which provides a scaffold that coordinates access of enzymes to lipid droplets in adipocytes[80]. However, among the human homologues of these genes, mutations were found only for *ABDH5* in Chanarin–Dorfman syndrome[78], in which TG accumulates in the skin, liver and muscle of affected patients, but not in the blood or adipose tissue.

Animal models — mainly mice but also invertebrates, rats, pigs and primates — have helped to both identify and validate genes in lipoprotein metabolism, and will continue to have an important role, especially with technical improvements such as tissue-specific gene targeting and the alteration of multiple genes and of large chromosomal regions. However, given the known limitations of animal model systems, it is important to appreciate that experimental results from these models must be interpreted within the totality of experimental evidence, especially results of human studies.

Sitosterolaemia
Also known as phytosterolaemia, this extremely rare autosomal recessive disorder is characterized by intestinal hyperabsorption and decreased biliary excretion of dietary sterols leading to hypercholesterolaemia, lipid deposits in the skin and tendons (xanthomas), and accelerated CVD.

Ezetimibe
A lipid-lowering drug that is the first member of the cholesterol absorption inhibitor class, which targets NPC1L1 on intestinal epithelial cells and in hepatocytes. This stimulates expression of cell-surface LDL receptors, resulting in increased clearance of LDL from the bloodstream.

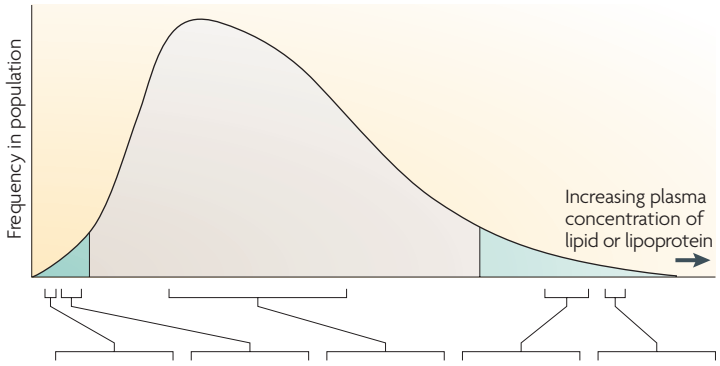
Mendelian randomization
The random assignment of alleles from parents to offspring that occurs during gamete formation. It is the underlying concept of a method to genetically stratify individuals in a large population sample and then evaluate phenotypic differences based on a pre-specified genotype.

Relationships between variants and phenotypes

As next-generation sequencing becomes more cost-effective, genome-wide sequence analysis will usurp genotyping in disease gene-mapping studies owing to its ability to identify hundreds of putative disease-causing mutations in an individual sample. For most mutations discovered by genome-wide sequence analysis, functional data will be absent. Traditional concerns surrounding proof of causation, such as differentiating biologically relevant dysfunctional mutations from incidental non-functional mutations, will be become more significant with the large increase in DNA sequence data. Rapid and robust approaches to impute a functional consequence will be required; Mendelian randomization and analysis of networks might be helpful in this regard.

Mendelian randomization. Mendelian randomization[81] has been proposed as an approach to assess whether genetically determined intermediate traits, such as lipoprotein levels, are causally related to end points, such as CVD. Proponents of Mendelian randomization suggest that the association of CVD risk with plasma lipoproteins can be diluted by non-genetic factors that alter plasma lipoproteins, whereas the association of CVD with the genetic determinants of lipoprotein levels is more direct, robust and less likely to be confounded by non-genetic influences[82]. For instance, heterozygosity for one of two nonsense mutations in *PCSK9* — involving tyrosine at position 142 (Y142X) or cysteine at position 679 (C679X) — which have a combined frequency of ~2% in subjects of African ancestry, was associated with reductions of 28% in mean LDL cholesterol and of 88% in CHD risk[83]. This level of CHD reduction was approximately three times greater than predicted from short-term studies of LDL cholesterol-lowering therapies[84]. Furthermore, meta-analysis of studies involving >120,000 participants showed LDL cholesterol variation of ~20% that tracked with a 40% variation in CHD risk across *APOE* genotypes[32], indicating an amplified CHD risk relative to expectations from LDL cholesterol differences in short-term clinical trials. These findings suggest that CHD risk is more closely related to the product of genetically determined LDL cholesterol levels and years of exposure than to a measurement of LDL at a single time point[85].

Mendelian randomization can also disprove the importance of genetic determinants that are initially hypothesized to be relevant. For instance, a large Danish study showed that heterozygotes for four rare *ABCA1* LOF mutations had ~20% reduced plasma HDL cholesterol, which would be expected to increase CHD risk by ~60% or more. However, no increase in CHD risk was observed[86], although the number of events was small. Also, the importance of *APOM* as a genetic determinant of HDL cholesterol level and of CHD end points was similarly challenged by minimal association of both traits with genetic variation in a large population sample[87]. This approach is therefore complementary to other methods and can help to rapidly validate the effect of new genetic determinants.



LDL cholesterol	Rare (ho):	Rare (he):	Common:	Rare (he):	Rare (ho):
	MTTP	*APOB*	*APOE*	*APOB*	*APOB*
	APOB	*PCSK9*	*LDLR*	*PCSK9*	*PCSK9*
	PCSK9		*PCSK9*	*LDLR*	*LDLR*
			FADS2,3		*LDLRAP1*
			APOB		
			HMGCR		
			SORT1		
			CILP2		
HDL cholesterol	Rare (ho):	Rare (he):	Common:	Rare (he):	Rare (ho):
	ABCA1	*ABCA1*	*CETP*	*CETP*	*CETP*
	APOA1	*APOA1*	*ABCA1*	*LIPC*	*LIPC*
	LCAT	*LCAT*	*LIPC*	*LIPG*	
		Common:	*LIPG*		
		WWOX	*LPL*		
			GALNT2		
			PLTP		
			MVK		
Triglyceride	Rare (ho):	Rare (he):	Common:	Rare (he):	Rare (ho):
	MTTP	*APOB*	*APOA5*	*LPL*	*LPL*
	PSCK9	*PSCK9*	*LPL*	*APOC2*	*APOC2*
		ANGPTL3	*GCKR*	*APOA5*	*APOA5*
		APOC3	*TRIB1*	Common:	*LMF1*
			CHREBP	*APOA5*	*GPIHBP1*
			GALNT2	*LPL*	
			ANGPTL3	*GCKR*	
			DOCK7	*TRIB1*	
			FADS1,2,3	*CHREBP*	
			LIPC	*GALNT2*	
			APOB	*ANGPTL3*	
			CILP2		

Figure 3 | **Genetic determinants of the plasma lipoprotein distribution.** A schematic representation of the population distribution of three lipoprotein traits: low-density lipoprotein (LDL) cholesterol and high-density lipoprotein (HDL) cholesterol, and triglyceride. Plasma concentration is on the x axis and the frequency of the trait is on the y axis. Shaded areas represent the bottom and top fifth percentiles of the distribution. Below the graph are genes that determine lipoprotein concentrations in specific segments of the distribution. Genes were determined by classical genetic or biochemical methods (no shading), by resequencing (orange shaded boxes) or by genome-wide association analysis (blue shaded boxes). The genes with common variants identified from genome-wide association represent several genes contained in the associated loci; some of these (for example, sortilin 1 (*SORT1*) or cartilage intermediate layer protein (*CILP*)) might not ultimately prove to be source of the association signal. The far extremes of the distribution are comprised of individuals with monogenic syndromes, with the prevalence of affected individuals being less that 1 in 10^5. Less extreme, with a prevalence of less than 1 in 100, are affected individuals with heterozygous mutations in many of the same genes as in the monogenic disorders. The recently reported *APOC3* nonsense mutation, found in Amish people with depressed plasma triglyceride, was only seen in heterozygotes[105]. Shown in the green boxes are genes with small to moderate effect sizes associated with severe hypertriglyceridaemia (also called hyperlipoproteinaemia type 5). Towards the median of the distribution, common polymorphisms in several genes make modest, but significant, contributions. For small-effect genes with common SNPs the relative size of their contribution is shown by the shading of the blue box, with darkest shading for the largest contribution. ho, homozygous; he, heterozygous. *ABCA1*, ATP-binding cassette A1; *ANGPTL3*, angiopoietin-like 3; *APO*, apolipoprotein; *CETP*, cholesteryl ester transfer protein; *CHREBP*, carbohydrate response element binding protein; *DOCK7*, dedicator of cytokinesis 7; *FADS*, fatty acid desaturase; *GALNT2*, UDP-*N*-acetyl-alpha-*D*-galactosamine:polypeptide *N*-acetylgalactosaminyltransferase; *GCKR*, glucokinase (hexokinase 4) regulator; *GPIHBP1*, glycosylphosphatidylinositol anchored high density lipoprotein binding protein 1; *HMGCR*, 3-hydroxy-3-methylglutaryl coenzyme A reductase; *LCAT*, lecithin cholesterol acyltransferase; *LDLR*, lipoprotein lipase receptor; *LDLRAP1*, low-density lipoprotein receptor adaptor protein 1; *LIPC*, hepatic lipase; *LIPG*, endothelial lipase; *LMF1*, lipase maturation factor 1; *LPL*, lipoprotein lipase; *MTTP*, microsomal triglyceride transfer protein; *MVK*, mevalonate kinase; *PCSK9*, proprotein convertase subtilisin/kexin type 9; *PLTP*, phospholipid transfer protein; *TRIB1*, tribbles homologue 1 (*Drosophila*); *WWOX*, WW-domain-containing oxidoreductase.

Network analysis
A blanket term for a range of computational methods to analyse complex sets of gene expression or related data in order to develop models of functionality, such as gene regulatory network models.

Network analysis of lipoprotein metabolism. Network analyses that are based on integrative genomic approaches combine large-scale genotypic and gene expression results in model systems, such as segregating mouse populations, to reconstruct reliable genetic networks underlying complex traits, such as plasma lipoproteins[88]. Computational network-based approaches can enhance traditional biochemical or animal experiments to implicate new causative genes and mutations, through mapping their involvement in molecular networks that are anchored around known susceptibility loci. For example, plasma HDL cholesterol was linked to promoter variants in vanin 1 (*VNN1*) using global profiles of RNA expression in blood lymphocytes[89]. Also,

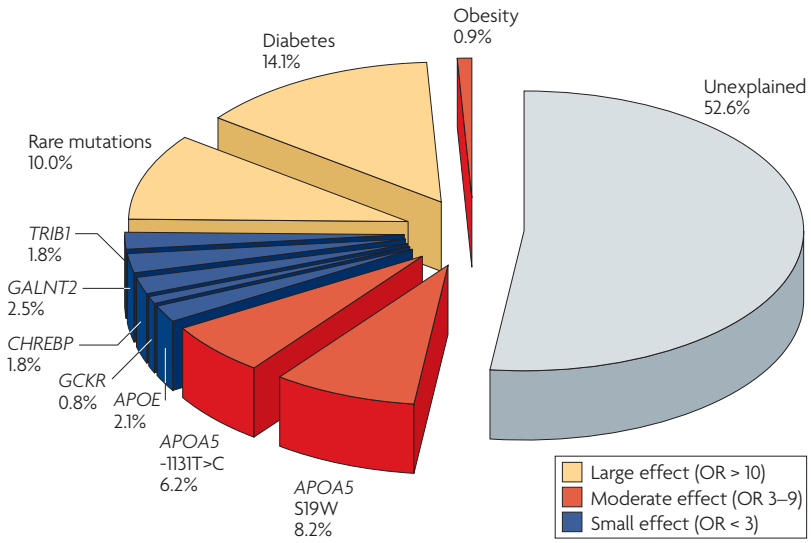


Figure 4 | **Relative contributions of genetic variants to severe hypertriglyceridaemia (HTG; also called hyperlipoproteinaemia type 5).** Genetic variants so far account for one-third of the susceptibility to HTG. In this figure, the percent contribution of several variables to severe HTG, estimated from linear regression analysis, is reflected in the size of the segment (with the percentages indicated), and the effect size, which is estimated from odds ratios (ORs) in logistic regression analysis, is reflected by the colour of the segment. Thus, rare mutations account for ~10% of the genetic component of severe HTG, but the OR is extremely high (~50) because these mutations are very rare in the general population. By contrast, a high proportion of subjects (~60%) with severe HTG carry common apolipoprotein A-V (*APOA5*) variants that have less of an influence on the phenotype (based on data from REFS 53, 93). Diabetes and obesity probably have their own genetic determinants, which are not shown here. The clinical and genetic features shown in this figure correctly classified individuals with HTG ~90% of the time, compared with 50% of the time when the classification is left to chance[93]. –1131T>C, T to C conversion at nucleotide position –1131; *APO*, apolipoprotein; *CHREBP*, carbohydrate response element binding protein; *GALNT2*, UDP-*N*-acetyl-alpha-*D*-galactosamine:polypeptide *N*-acetylgalactosaminyltransferase; *GCKR*, glucokinase (hexokinase 4) regulator; S19W, serine to tryptophan conversion at amino acid 19; *TRIB1*, tribbles homologue 1 (*Drosophila*).

global studies of expression in murine liver and adipose tissue imputed a causal relationship between a novel macrophage-enriched gene network and obesity and related traits such as TG and HDL[90]. Specifically, three genes identified by network-based analysis, namely *Lpl*, lactamase beta (*Lactb*) and protein phosphatase 1-like (*Ppm1l*), were validated as previously unknown obesity genes, which strengthened the association between a newly identified biological network and a complex metabolic trait[90]. Thus, metabolic traits such as plasma lipoproteins, which are products of molecular networks that are modulated by sets of complex genetic loci and environmental factors, would lend themselves to evaluation by such advanced computational approaches.

Mosaic model case study: plasma TGs

According to the 'common disease–common variant' model, genetic predisposition is comprised of multiple common genetic variants that each have a small to moderate effect on phenotypes, either alone or in combination with modifier genes or environmental factors[91]. Alternatively, the 'heterogeneity' model maintains that genetic contribution to complex traits is comprised of numerous rare genetic variants[92]. In reality, a blend of these models — in which a mosaic of many variants, frequent and rare, with effect sizes ranging from small to large — seems to explain the genetic basis of plasma lipoproteins (FIG. 3).

Plasma TG variation illustrates the genetic mosaic model. The most extreme HTG segment (that is, the top 0.001% of the population) contains individuals with a monogenic illness (HLP type 1) who carry rare homozygous large-effect mutations[41,42,44] (FIG. 3). By contrast, near the median TG level, GWA studies identified approximately eight common associated SNPs with small effect sizes (1–2% per gene)[24,58]. When these common SNPs were studied in HTG patients with TG >10 mmol L^{-1} (that is, the top 0.2% of the population)[53], almost all of these SNPs showed moderate associations (odds ratio of 2–10) and cumulatively contributed to ~30% of the disease susceptibility[93] (FIG. 4). Combining markers increased HTG risk: those with six risk alleles had an almost 'monogenic' odds ratio of 25 (FIG. 5), but they made up only ~1% of the population; and ~15% of HTG subjects had seven risk alleles compared with 0% of controls. In addition, rare heterozygous LOF mutations were strongly associated with HTG (odds ratio of ~50), albeit in only ~10% of patients[53]. Thus, common small to moderate-effect variants and rare large-effect mutations can both be used to determine TG levels, albeit in different segments of the distribution. If this model can be validated, it might also apply to other phenotypes and traits.

Implications of lipoprotein genetic studies

Biological or clinical utility or both? To a clinician, the variation in lipoprotein levels that can be attributed to individual variants that have been identified by GWA studies (1–2%) might be considered to be too small to be meaningful. These effects cause alterations in levels within the normal range, and would not merit diagnosis or treatment. The extent of such variation, even cumulatively across several genes, is small compared with variation from other sources, such as laboratory error in measuring lipoprotein levels or the effects of exercise or diet. The major short-term effect from resequencing and GWA studies might therefore be more relevant to biological researchers than to clinicians, as the newly identified pathways can immediately be studied. The potential role of these new molecules and pathways in the diagnosis and treatment of patients is not immediately apparent and will require extensive study[59]. However, the impact of some genetic discoveries will be felt at the clinical level eventually, including prediction of disease risk, drug development, nosology and response to treatment.

Will genetic data add clinically useful information other than simple biochemical determination of lipoprotein levels? Genotype might be a superior index of lifelong exposure to an intermediate trait than a single biochemical test[85,94], and genetic risk stratification is possible at a young age, when dyslipidaemia is not clinically apparent. Pilot studies have shown that

Receiver operating characteristic (ROC) curve
A graphical plot, based on signal detection theory, which plots sensitivity along the y axis and 1 – specificity along the x axis for a binary classifier system — such as a genetic test or indeed any clinical test. ROC curve analysis can help select optimal diagnostic models and is fundamental to cost versus benefit analysis of diagnostic decision making.

Pleiotropy
Occurs when a single gene influences multiple phenotypic traits.

Epistasis
The interaction between genes when the action of one gene is modified by one or more other genes, which are sometimes called modifier genes.

Statin
A member of the class of lipid-lowering drugs that are used to treat people who have, or are at risk of, CVD. By inhibiting HMGCR — a key enzyme in cholesterol synthesis — statins ultimately stimulate the expression of cell-surface LDL receptors, resulting in an increased clearance of LDL from the bloodstream.

SNP data can predict CVD: individuals with more deleterious variants experience CVD at a younger age[24,95]. Furthermore, the relationships between genotype and CVD risk seem to be independent of single lipoprotein measurements, suggesting that subtle effects might be lost in a measurement made at a single point in time and are instead integrated over a patient's lifetime.

A related question is whether genetic testing can improve the ability to predict CVD, using standard criteria such as increased area under the receiver operating characteristic (ROC) curve, a standard approach to evaluate the clinical usefulness of a laboratory test. This has been demonstrated for some SNPs that are associated with CHD[96], although larger and more extensive replication cohorts, validation studies and demonstration of the consistency of markers across populations are required before such tests can be considered clinically.

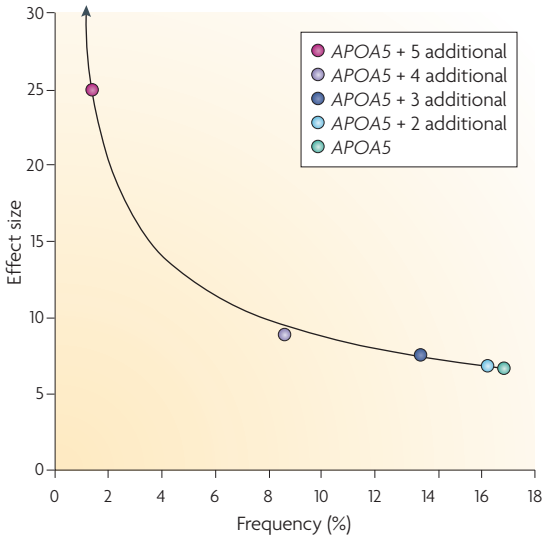


Figure 5 | **Polygenicity of severe hypertriglyceridaemia (HTG).** The effect size is the odds ratio for affected status with severe HTG (also called hyperlipoproteinaemia type 5) and is shown on the y axis. The frequency of carriers of particular combinations of alleles in the normal population is shown along the x axis (data is taken from REF. 93). As two apolipoprotein A-V (*APOA5*) SNPs — S19W (serine to tryptophan conversion at amino acid 19) and –1131T>C (T to C conversion at nucleotide position –1131) — are very strongly associated with severe HTG, the presence of either served as the primary genetic predictor: the odds ratio (OR) was 6.93 (95% confidence interval 4.44–10.8) and the carrier frequency in normal controls was ~17%. Adding any one or two of the other SNP genotypes from FIG. 4 did not substantially change this OR. However, adding three, four or five additional common SNPs from FIG. 4 to the presence of either *APOA5* marker sequentially increased the OR up to 25, so that individuals with six genetic risk markers had an OR for severe HTG that was almost in the monogenic range. However, only ~1.5% of individuals in the general population would be unfortunate enough to carry all of these markers[93]. Notably, 15.2% (20 in 132) of HTG patients carried seven genetic risk alleles, compared with 0% (0 in 351) of normolipidaemic controls, which prohibits the calculation of an OR but indicates that these rare individuals have an enormous risk of developing HTG[93].

Despite the attractiveness of genetic markers, established biomarkers such as plasma LDL cholesterol level already represent the integration of many genetic and environmental effects — many of which are unknown and immeasurable. Even with 'complete' genomic information, confounders such as genetic pleiotropy, epistasis, influence of genetic background, epigenetic mechanisms (such as methylation) and gene–environment interactions need to be considered. Despite the current limitations of using genetic factors to diagnose or predict the plasma lipid profile, it is likely that the identification of genetic determinants will have several practical implications.

Drug development and design. The association signal of *HMGCR* from GWA studies of LDL cholesterol illustrates the validity of this approach for the identification of molecular targets and pathways. Although common *HMGCR* SNPs had only a ~2–3% effect on LDL cholesterol in GWA studies, pharmacologic targeting of this enzyme with statin drugs can reduce LDL cholesterol by 50% or more[59]. Furthermore, inter-individual differences in response to statin drugs are associated with a common alternatively spliced variant of *HMGCR*[97]. Thus, other genes from GWA studies with comparably small effects at the population level might similarly point to targets for drugs that will have significant clinical efficacy[59].

GWA studies showed that *APOB* and *PCSK9* SNPs are associated with small variations in LDL cholesterol; each of these genes is currently being targeted for LDL cholesterol reduction using RNA-based interference strategies[98,99]. In addition, the association of *CETP* with HDL cholesterol in GWA studies confirmed the relevance of the pathway involving CETP; indeed, several inhibitors of CETP are being developed to raise HDL cholesterol, although not without concerns about their efficacy and risk[100]. Although the decisions to develop inhibitors of APOB, PCSK9 and CETP preceded the recent GWA results, the consistency of data across monogenic, resequencing, animal and now GWA studies increases confidence in these targets. The findings also raise the possibility that targeting other novel loci found in GWA studies might be effective for modulating plasma lipoproteins. Finally, resequencing studies that identify LOF mutations in new genes in patients with favourable lipoproteins should identify therapeutic targets for pharmacologic inhibition or knockdown strategies.

Nosology of lipoprotein disorders. The identification of molecular genetic determinants could help to revise current biochemical-based classification systems, such as the WHO International Classification of Diseases (ICD) phenotypes (BOX 1; TABLE 1). For instance, 40–60% of subjects with HLP type 2B, 3, 4 or 5 carry one of two *APOA5* variants — S19W (serine to tryptophan conversion at amino acid 19) or -1131T>C (T to C conversion at nucleotide position –1131) — whereas only ~15% of

normolipidaemic individuals carry these variants[101]. The HLP type 2B, 3, 4 and 5 phenotypes seem to be disparate at the biochemical level, but they have much more in common genetically than was previously suspected[101]. A different *APOA5* variant, G185C (glycine to cysteine conversion at amino acid 185), showed similar associations with TG in individuals of Asian ancestry[102]. Thus, HTG could be subdivided on the basis of *APOA5* genotype, especially if this affects prognosis or response to intervention. Furthermore, resequencing studies that find LOF mutations in new genes in dyslipidaemic patients will suggest new etiologies, helping to refine diagnosis or treatment.

Public health implications of genetic findings. Mendelian epidemiological findings, especially those related to *PCSK9* mutations and reduced CVD, have been interpreted to suggest that early initiation and long-term maintenance of LDL-lowering treatment will produce important public health benefits[85,94]. This logic is appealing, especially given the metabolic problems and associated long-term CVD risks faced by today's youth. However, before treating children with statins, it is important to clarify a few points: whether pleiotropic effects of the *PCSK9* variants on non-LDL pathways contribute to lifelong CVD risk reduction; whether pharmacological inhibition of HMGCR is medically or biologically equivalent to a germline mutation in *PCSK9*; and whether exposing large numbers of currently healthy individuals to drug treatment will create other problems, such as off-target or side effects, rapid consumption of resources, and unforeseen consequences of long-term exposure. Because diet is a key determinant of lipoprotein levels[103], early dietary interventions would probably be the safest, most economical and most effective strategy for CVD prevention.

Conclusions

Progress in understanding the genes that determine plasma lipoprotein levels has rapidly accelerated thanks to high-throughput automated DNA sequencing and genotyping, which complement traditional approaches such as candidate gene association studies, linkage studies and the use of animal models. Phenomic analysis (also called deep phenotyping), systems biology and network approaches might help us to further integrate experimental data. Newly identified genes are potential new drug targets, so it is crucial to define the metabolic roles of the gene products. Given the probable diminishing returns from further GWA studies of normolipidaemic samples, future GWA studies should focus on samples covering a wide range of geographical ancestries, and should also be performed in patients with dyslipidaemias. Next-generation sequencing of the whole genome will provide an unbiased approach for the identification of additional causative genes in patients with extreme lipoprotein phenotypes — an approach that does not depend on prior knowledge. Accounting for the genetic determinants of such intimately associated metabolic states as obesity and diabetes will also be essential. Validation of new findings will require multi-centre collaborations and rapid translation across technological and experimental platforms. Also, because drugs and diet play a key part in the management of dyslipidaemia, rational pharmacogenetic and nutrigenetic studies should be undertaken[104]. The definition of new pathways and targets will inform new drug design and eventually lead to evidence-based changes in clinical practice. However, predicting the precise evolution and consequences of dyslipidaemia in any single individual might remain elusive, because of the confounding influence of the environment, non-linear interactions between genes and environment, and stochastic effects in these networks of pathways.

1. Breslow, J. L. Genetics of lipoprotein abnormalities associated with coronary artery disease susceptibility. *Annu. Rev. Genet.* **34**, 233–254 (2000).
2. Lusis, A. J. Genetic factors affecting blood lipoproteins: the candidate gene approach. *J. Lipid Res.* **29**, 397–429 (1988).
3. Cohen, J. C. *et al.* Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science* **305**, 869–872 (2004).
An early demonstration of how resequencing individuals at the extremes of the distribution of lipoprotein levels can uncover genetic determinants.
4. Brown, M. S. & Goldstein, J. L. A receptor-mediated pathway for cholesterol homeostasis. *Science* **232**, 34–47 (1986).
5. Incardona, J. P. & Roelink, H. The role of cholesterol in Shh signaling and teratogen-induced holoprosencephaly. *Cell. Mol. Life Sci.* **57**, 1709–1719 (2000).
6. Repa, J. J. & Mangelsdorf, D. J. The role of orphan nuclear receptors in the regulation of cholesterol homeostasis. *Annu. Rev. Cell Dev. Biol.* **16**, 459–481 (2000).
7. Lusis, A. J. Atherosclerosis. *Nature* **407**, 233–241 (2000).
8. Rader, D. J. & Daugherty, A. Translating molecular discoveries into new therapies for atherosclerosis. *Nature* **451**, 904–913 (2008).
9. Grundy, S. M. Promise of low-density lipoprotein-lowering therapy for primary and secondary prevention. *Circulation* **117**, 569–573; discussion 573 (2008).
10. Lewis, G. F. & Rader, D. J. New insights into the regulation of HDL metabolism and reverse cholesterol transport. *Circ. Res.* **96**, 1221–1232 (2005).
11. Stamler, J. *et al.* Relationship of baseline serum cholesterol levels in 3 large cohorts of younger men to long-term coronary, cardiovascular, and all-cause mortality and to longevity. *JAMA* **284**, 311–318 (2000).
12. Genest, J. J. Jr *et al.* Familial lipoprotein disorders in patients with premature coronary artery disease. *Circulation* **85**, 2025–2033 (1992).
13. Sniderman, A. D. Applying apoB to the diagnosis and therapy of the atherogenic dyslipoproteinemias: a clinical diagnostic algorithm. *Curr. Opin. Lipidol.* **15**, 433–438 (2004).
14. Barter, P. J. & Rye, K. A. The rationale for using apoA-I as a clinical marker of cardiovascular risk. *J. Intern. Med.* **259**, 447–454 (2006).
15. McQueen, M. J. *et al.* Lipids, lipoproteins, and apolipoproteins as risk markers of myocardial infarction in 52 countries (the INTERHEART study): a case–control study. *Lancet* **372**, 224–233 (2008).
16. van der Steeg, W. A. *et al.* Role of the apolipoprotein B–apolipoprotein A-I ratio in cardiovascular risk assessment: a case–control analysis in EPIC-Norfolk. *Ann. Intern. Med.* **146**, 640–648 (2007).
17. Nordestgaard, B. G., Benn, M., Schnohr, P. & Tybjaerg-Hansen, A. Nonfasting triglycerides and risk of myocardial infarction, ischemic heart disease, and death in men and women. *JAMA* **298**, 299–308 (2007).
18. Bansal, S. *et al.* Fasting compared with nonfasting triglycerides and risk of cardiovascular events in women. *JAMA* **298**, 309–316 (2007).
19. Austin, M. A. *et al.* Cardiovascular disease mortality in familial forms of hypertriglyceridemia: a 20-year prospective study. *Circulation* **101**, 2777–2782 (2000).
20. Kathiresan, S. *et al.* Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nature Genet.* **40**, 189–197 (2008).
21. Kooner, J. S. *et al.* Genome-wide scan identifies variation in *MLXIPL* associated with plasma triglycerides. *Nature Genet.* **40**, 149–151 (2008).
22. Sandhu, M. S. *et al.* LDL-cholesterol concentrations: a genome-wide association study. *Lancet* **371**, 483–491 (2008).
23. Wallace, C. *et al.* Genome-wide association study identifies genes for biomarkers of cardiovascular disease: serum urate and dyslipidemia. *Am. J. Hum. Genet.* **82**, 139–149 (2008).
24. Willer, C. J. *et al.* Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nature Genet.* **40**, 161–169 (2008).
25. Aulchenko, Y. S. *et al.* Genome-wide association study in 16 European population cohorts: major loci influencing lipid levels and coronary heart disease risk. *Nature Genet.* **41**, 47–55 (2009).
26. Kathiresan, S. *et al.* Common DNA sequence variants at thirty genetic loci contribute to polygenic dyslipidemia. *Nature Genet.* **41**, 56–65 (2009).

27. Sabatti, C. *et al.* Genomewide association analysis of metabolic traits in a birth cohort from a founder population. *Nature Genet.* **41**, 35–46 (2009). **References 20–27 are GWA studies that identify new loci for plasma lipoproteins.**
28. Hegele, R. A., Brunt, J. H. & Connelly, P. W. Multiple genetic determinants of variation of plasma lipoproteins in Alberta Hutterites. *Arterioscler. Thromb. Vasc. Biol.* **15**, 861–871 (1995).
29. Pajukanta, P. *et al.* Familial combined hyperlipidemia is associated with upstream transcription factor 1 (*USF1*). *Nature Genet.* **36**, 371–376 (2004).
30. Lee, J. C. *et al.* WW-domain-containing oxidoreductase is associated with low plasma HDL-C levels. *Am. J. Hum. Genet.* **83**, 180–192 (2008).
31. Saxena, R. *et al.* Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* **316**, 1331–1336 (2007).
32. Bennet, A. M. *et al.* Association of apolipoprotein E genotypes with lipid levels and coronary risk. *JAMA* **298**, 1300–1311 (2007).
33. Thompson, A. *et al.* Association of cholesteryl ester transfer protein genotypes with CETP mass and activity, lipid levels, and coronary risk. *JAMA* **299**, 2777–2788 (2008).
34. Shoulders, C. C., Jones, E. L. & Naoumova, R. P. Genetics of familial combined hyperlipidemia and risk of coronary heart disease. *Hum. Mol. Genet.* **13**, R149–R160 (2004).
35. Pollex, R. L. & Hegele, R. A. Genetic determinants of plasma lipoproteins. *Nature Clin. Pract. Cardiovasc. Med.* **4**, 600–609 (2007).
36. Rader, D. J., Cohen, J. & Hobbs, H. H. Monogenic hypercholesterolemia: new insights in pathogenesis and treatment. *J. Clin. Invest.* **111**, 1795–1803 (2003).
37. Tarugi, P. *et al.* Molecular diagnosis of hypobetalipoproteinemia: an ENID review. *Atherosclerosis* **195**, e19–e27 (2007).
38. Brooks-Wilson, A. *et al.* Mutations in *ABC1* in Tangier disease and familial high-density lipoprotein deficiency. *Nature Genet.* **22**, 336–345 (1999).
39. Assmann, G., von Eckardstein, A. & Funke, H. High density lipoproteins, reverse transport of cholesterol, and coronary artery disease. Insights from mutations. *Circulation* **87**, III28–III34 (1993).
40. Hovingh, G. K. *et al.* Inherited disorders of HDL metabolism and atherosclerosis. *Curr. Opin. Lipidol.* **16**, 139–145 (2005).
41. Fojo, S. S. & Brewer, H. B. Hypertriglyceridaemia due to genetic defects in lipoprotein lipase and apolipoprotein C-II. *J. Intern. Med.* **231**, 669–677 (1992).
42. Peterfy, M. *et al.* Mutations in *LMF1* cause combined lipase deficiency and severe hypertriglyceridemia. *Nature Genet.* **39**, 1483–1487 (2007).
43. Young, S. G. *et al.* GPIHBP1: an endothelial cell molecule important for the lipolytic processing of chylomicrons. *Curr. Opin. Lipidol.* **18**, 389–396 (2007).
44. Talmud, P. J. Rare *APOA5* mutations — clinical consequences, metabolic and functional effects: an ENID review. *Atherosclerosis* **194**, 287–292 (2007).
45. Stroes, E. S. *et al.* Intramuscular administration of AAV1-lipoprotein lipase S447X lowers triglycerides in lipoprotein lipase-deficient patients. *Arterioscler. Thromb. Vasc. Biol.* **28**, 2303–2304 (2008).
46. Whitfield, A. J., Barrett, P. H., van Bockxmeer, F. M. & Burnett, J. R. Lipid disorders and mutations in the *APOB* gene. *Clin. Chem.* **50**, 1725–1732 (2004).
47. Jones, B. *et al.* Mutations in a Sar1 GTPase of COPII vesicles are associated with lipid absorption disorders. *Nature Genet.* **34**, 29–31 (2003).
48. Ahituv, N. *et al.* Medical sequencing at the extremes of human body mass. *Am. J. Hum. Genet.* **80**, 779–791 (2007).
49. Kryukov, G. V., Pennacchio, L. A. & Sunyaev, S. R. Most rare missense alleles are deleterious in humans: implications for complex disease and association studies. *Am. J. Hum. Genet.* **80**, 727–739 (2007).
50. Cohen, J. *et al.* Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in *PCSK9*. *Nature Genet.* **37**, 161–165 (2005).
51. Cohen, J. C. *et al.* Multiple rare variants in *NPC1L1* associated with reduced sterol absorption and plasma low-density lipoprotein levels. *Proc. Natl Acad. Sci. USA* **103**, 1810–1815 (2006).
52. Fahmi, S., Yang, C., Esmail, S., Hobbs, H. H. & Cohen, J. C. Functional characterization of genetic variants in *NPC1L1* supports the sequencing extremes strategy to identify complex trait genes. *Hum. Mol. Genet.* **17**, 2101–7 (2008).
53. Wang, J. *et al.* Resequencing genomic DNA of patients with severe hypertriglyceridemia (MIM 144650). *Arterioscler. Thromb. Vasc. Biol.* **27**, 2450–2455 (2007).
54. deLemos, A. S., Wolfe, M. L., Long, C. J., Sivapackianathan, R. & Rader, D. J. Identification of genetic variants in endothelial lipase in persons with elevated high-density lipoprotein cholesterol. *Circulation* **106**, 1321–1326 (2002).
55. Yoshida, K., Shimizugawa, T., Ono, M. & Furukawa, H. Angiopoietin-like protein 4 is a potent hyperlipidemia-inducing factor in mice and inhibitor of lipoprotein lipase. *J. Lipid Res.* **43**, 1770–1772 (2002).
56. Romeo, S. *et al.* Population-based resequencing of *ANGPTL4* uncovers variations that reduce triglycerides and increase HDL. *Nature Genet.* **39**, 513–516 (2007).
57. Frazer, K. A. *et al.* A second generation human haplotype map of over 3.1 million SNPs. *Nature* **449**, 851–861 (2007).
58. Kathiresan, S. *et al.* A genome-wide association study for blood lipid phenotypes in the Framingham Heart Study. *BMC Med. Genet.* **8** (Suppl. 1), S17 (2007).
59. Hunter, D. J., Altshuler, D. & Rader, D. J. From Darwin's finches to canaries in the coal mine — mining the genome for new biology. *N. Engl. J. Med.* **358**, 2760–2763 (2008).
60. Kathiresan, S., Musunuru, K. & Orho-Melander, M. Defining the spectrum of alleles that contribute to blood lipid concentrations in humans. *Curr. Opin. Lipidol.* **19**, 122–127 (2008).
61. Krauss, R. M. What can the genome tell us about LDL cholesterol? *Lancet* **371**, 450–452 (2008).
62. Lusis, A. J. & Pajukanta, P. A treasure trove for lipoprotein biology. *Nature Genet.* **40**, 129–130 (2008).
63. Uyeda, K. & Repa, J. J. Carbohydrate response element binding protein, ChREBP, a transcription factor coupling hepatic glucose utilization and lipid synthesis. *Cell. Metab.* **4**, 107–110 (2006).
64. Redon, R. *et al.* Global variation in copy number in the human genome. *Nature* **444**, 444–454 (2006).
65. Pollex, R. L. & Hegele, R. A. Genomic copy number variation and its potential role in lipoprotein and metabolic phenotypes. *Curr. Opin. Lipidol.* **18**, 174–180 (2007).
66. Reue, K. & Doolittle, M. H. Naturally occurring mutations in mice affecting lipid transport and metabolism. *J. Lipid Res.* **37**, 1387–1405 (1996).
67. Beigneux, A. P. *et al.* Glycosylphosphatidylinositol-anchored high-density lipoprotein-binding protein 1 plays a critical role in the lipolytic processing of chylomicrons. *Cell. Metab.* **5**, 279–291 (2007).
68. Gin, P. *et al.* Normal binding of lipoprotein lipase, chylomicrons, and apo-AV to GPIHBP1 containing a G56R amino acid substitution. *Biochim. Biophys. Acta* **1771**, 1464–1468 (2007).
69. Wolfrum, C., Poy, M. N. & Stoffel, M. Apolipoprotein M is required for prebeta-HDL formation and cholesterol efflux to HDL and protects against atherosclerosis. *Nature Med.* **11**, 418–422 (2005).
70. Shimizugawa, T. *et al.* ANGPTL3 decreases very low density lipoprotein triglyceride clearance by inhibition of lipoprotein lipase. *J. Biol. Chem.* **277**, 33742–33748 (2002).
71. Koster, A. *et al.* Transgenic angiopoietin-like (angptl)4 overexpression and targeted disruption of angptl4 and angptl3: regulation of triglyceride metabolism. *Endocrinology* **146**, 4943–4950 (2005).
72. Lee, M. H. *et al.* Identification of a gene, *ABCG5*, important in the regulation of dietary cholesterol absorption. *Nature Genet.* **27**, 79–83 (2001).
73. Berge, K. E. *et al.* Accumulation of dietary cholesterol in sitosterolemia caused by mutations in adjacent ABC transporters. *Science* **290**, 1771–1775 (2000).
74. Yang, C. *et al.* Disruption of cholesterol homeostasis by plant sterols. *J. Clin. Invest.* **114**, 813–822 (2004).
75. Altmann, S. W. *et al.* Niemann-Pick C1 Like 1 protein is critical for intestinal cholesterol absorption. *Science* **303**, 1201–1204 (2004).
76. Simon, J. S. *et al.* Sequence variation in NPC1L1 and association with improved LDL-cholesterol lowering in response to ezetimibe treatment. *Genomics* **86**, 648–56 (2005).
77. Haemmerle, G. *et al.* Defective lipolysis and altered energy metabolism in mice lacking adipose triglyceride lipase. *Science* **312**, 734–737 (2006).
78. Lass, A. *et al.* Adipose triglyceride lipase-mediated lipolysis of cellular fat stores is activated by CGI-58 and defective in Chanarin–Dorfman Syndrome. *Cell. Metab.* **3**, 309–319 (2006).
79. Reue, K. The role of lipin 1 in adipogenesis and lipid metabolism. *Novartis Found. Symp.* **286**, 58–68; discussion 68–71, 162–163, 196–203 (2007).
80. Brasaemle, D. L. Thematic review series: adipocyte biology. The perilipin family of structural lipid droplet proteins: stabilization of lipid droplets and control of lipolysis. *J. Lipid Res.* **48**, 2547–2559 (2007).
81. Katan, M. B. Apolipoprotein E isoforms, serum cholesterol, and cancer. *Lancet* **1**, 507–508 (1986).
82. Ebrahim, S. & Davey Smith, G. Mendelian randomization: can genetic epidemiology help redress the failures of observational epidemiology? *Hum. Genet.* **123**, 15–33 (2008).
83. Cohen, J. C., Boerwinkle, E., Mosley, T. H. Jr & Hobbs, H. H. Sequence variations in *PCSK9*, low LDL, and protection against coronary heart disease. *N. Engl. J. Med.* **354**, 1264–1272 (2006). **An important study showing how Mendelian randomization can statistically connect genetic variants to the clinical end points that are mediated through plasma lipoproteins.**
84. Baigent, C. *et al.* Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90,056 participants in 14 randomised trials of statins. *Lancet* **366**, 1267–1278 (2005).
85. Brown, M. S. & Goldstein, J. L. Biomedicine. Lowering LDL — not only how low, but how long? *Science* **311**, 1721–1723 (2006).
86. Frikke-Schmidt, R. *et al.* Association of loss-of-function mutations in the *ABCA1* gene with high-density lipoprotein cholesterol levels and risk of ischemic heart disease. *JAMA* **299**, 2524–2532 (2008).
87. Ahnstrom, J. *et al.* Levels of apolipoprotein M are not associated with the risk of coronary heart disease in two independent case–control studies. *J. Lipid Res.* **49**, 1912–1917 (2008).
88. Lusis, A. J. A thematic review series: systems biology approaches to metabolic and cardiovascular disorders. *J. Lipid Res.* **47**, 1887–1890 (2006).
89. Goring, H. H. *et al.* Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes. *Nature Genet.* **39**, 1208–1216 (2007).
90. Chen, Y. *et al.* Variations in DNA elucidate molecular networks that cause disease. *Nature* **452**, 429–435 (2008).
91. Reich, D. E. & Lander, E. S. On the allelic spectrum of human disease. *Trends Genet.* **17**, 502–510 (2001).
92. Wang, W. Y., Barratt, B. J., Clayton, D. G. & Todd, J. A. Genome-wide association studies: theoretical and practical concerns. *Nature Rev. Genet.* **6**, 109–118 (2005).
93. Wang, J. *et al.* Polygenic determinants of severe hypertriglyceridemia. *Hum. Mol. Genet.* **17**, 2894–2899 (2008).
94. Steinberg, D., Glass, C. K. & Witztum, J. L. Evidence mandating earlier and more aggressive treatment of hypercholesterolemia. *Circulation* **118**, 672–677 (2008).
95. Kathiresan, S. *et al.* Polymorphisms associated with cholesterol and risk of cardiovascular events. *N. Engl. J. Med.* **358**, 1240–1249 (2008).
96. Humphries, S. E., Yiannakouris, N. & Talmud, P. J. Cardiovascular disease risk prediction using genetic information (gene scores): is it really informative? *Curr. Opin. Lipidol.* **19**, 128–132 (2008).
97. Medina, M. W., Gao, F., Ruan, W., Rotter, J. I. & Krauss, R. M. Alternative splicing of 3-hydroxy-3-methylglutaryl coenzyme A reductase is associated with plasma low-density lipoprotein cholesterol response to simvastatin. *Circulation* **118**, 355–362 (2008).
98. Frank-Kamenetsky, M. *et al.* Therapeutic RNAi targeting PCSK9 acutely lowers plasma cholesterol in rodents and LDL cholesterol in nonhuman primates. *Proc. Natl Acad. Sci. USA* **105**, 11915–11920 (2008).
99. Akdim, F., Stroes, E. S. & Kastelein, J. J. Antisense apolipoprotein B therapy: where do we stand? *Curr. Opin. Lipidol.* **18**, 397–400 (2007).
100. Holleboom, A. G., Vergeer, M., Hovingh, G. K., Kastelein, J. J. & Kuivenhoven, J. A. The value of HDL genetics. *Curr. Opin. Lipidol.* **19**, 385–394 (2008).

101. Wang, J. *et al. APOA5* genetic variants are markers for classic hyperlipoproteinemia phenotypes and hypertriglyceridemia. *Nature Clin. Pract. Cardiovasc. Med.* **5**, 730–737 (2008).
102. Pullinger, C. R. *et al.* An apolipoprotein A-V gene SNP is associated with marked hypertriglyceridemia among Asian-American patients. *J. Lipid Res.* **49**, 1846–1854 (2008).
103. Grundy, S. M. & Denke, M. A. Dietary influences on serum lipids and lipoproteins. *J. Lipid Res.* **31**, 1149–1172 (1990).
104. Ordovas, J. M. Pharmacogenetics of lipid diseases. *Hum. Genomics* **1**, 111–125 (2004).
105. Pollin, T. I. *et al.* A null mutation in human *APOC3* confers a favorable lipid profile and apparent cardioprotection. *Science* **322**, 1702–1705 (2008).

Acknowledgements
The author wishes to acknowledge the contributions to this field of numerous colleagues whose work could not be cited owing to space constraints. M. Peterfy, P. Connelly, M. Lanktree, T. Joy, P. Lahiry, J. Robinson, M. Ban, J. Wang and H. Cao provided helpful comments. The author is supported by the Jacob J. Wolfe Distinguished Medical Research Chair; the Edith Schulich Vinet Canada Research Chair (Tier I) in Human Genetics; the Jean Davignon Award for Cardiovascular Research (Pfizer, Canada); and by operating grants from the Canadian Institutes for Health Research (MOP-13430, MOP-39533, MOP-39833), the Heart and Stroke Foundation of Ontario (PRG-5967, NA-6059, T-6018), the Ontario Research Fund, and by Genome Canada through the Ontario Genomics Institute.

DATABASES

Entrez Gene: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene
ABCA1 | *APOE* | *CETP* | *CHREBP* | *GALNT2* | *GCKR* | *MTTP* | *PCSK9*
UniProtKB: http://www.uniprot.org
apolipoprotein A-I | apolipoprotein B | HMGCR

FURTHER INFORMATION

Hegele laboratory homepage: http://www.robarts.ca/hegele

SUPPLEMENTARY INFORMATION

See online article: S1 (table)

ALL LINKS ARE ACTIVE IN THE ONLINE PDF