Genetic determinants of plasma lipoproteins

Rebecca L Pollex and Robert A Hegele*

SUMMARY

The search for common genetic determinants of plasma lipoproteins began in the early 1980s. Despite some exceptions, these efforts have not yet yielded a set of biological markers that can be used in clinical practice. By contrast, successes in defining the molecular basis of rare single-gene disorders, such as familial hypoalphalipoproteinemia, have shown the value of experimental designs that focus on genomic analysis of individuals within the tails of Gaussian distributions of quantitative lipoprotein traits. For example, this strategy showed that a small but relevant proportion of individuals within the <5% tail of plasma HDL-cholesterol distribution have mutations in genes that cause familial hypoalphalipoproteinemia. The value of clinical testing for genomic variants as an adjunct to a biochemical measurement of plasma lipoproteins, however, is at best questionable. A more direct impact of genetic studies is that definitions of 'common' and 'large genetic effects' have become more tempered, reflecting perhaps the biological reality that plasma lipoproteins are probably determined by the aggregate of numerous modest and occasional large genetic effects in addition to environmental factors. Here, we review recent progress on genomic variants and cholesterol metabolism, and discuss the impact these genetic studies will have on clinical cardiology.

KEYWORDS association studies, genetics, lipoprotein, single-nucleotide polymorphisms

REVIEW CRITERIA

We searched PubMed for all available publications between the years 2005 and 2006, using the following terms: "lipoprotein", "cholesterol" and "polymorphism". All relevant articles identified were English-language papers.

RL Pollex is a Research Assistant in the Vascular Biology Group and RA Hegele is Director of the Blackburn Cardiovascular Genetics Laboratory, Robarts Research Institute, London, ON, Canada.

Correspondence
*Blackburn Cardiovascular Genetics Laboratory, Robarts Research Institute, Room 406,
100 Perth Drive, London, ON N6A 5K8, Canada
hegele@robarts.ca

Received 15 January 2007 **Accepted** 28 July 2007
www.nature.com/clinicalpractice
doi:10.1038/ncpcardio1005

INTRODUCTION

The variation in plasma lipoprotein concentrations seen throughout a population is a reflection of the different environmental influences and genetic factors affecting each individual. Genetic variation exerts a sizeable, complex influence on lipoprotein phenotypes, with heritability estimates (i.e. estimates of the proportion of variation that is related to inherited factors) exceeding 50% for total cholesterol, HDL cholesterol and triglycerides.[1]

Over the past 25 years, numerous researchers have attempted to understand the genetic basis of differing lipoprotein phenotypes by studying both common polygenic lipoprotein phenotypes and rare Mendelian disorders of lipoprotein metabolism. The interest in identifying genetic determinants stems from the strong relationships that elevated plasma LDL-cholesterol levels and reduced HDL-cholesterol levels have with the risk of cardiovascular disease. We review recent progress on genomic variants and cholesterol metabolism. We focus on key genes that are determinants of plasma lipoproteins, namely well-characterized genes that have been consistently linked to cholesterol metabolism over the years, such as *APOE*, *CETP* and *APOC3*, known genes whose role in cholesterol metabolism has recently been supported by new evidence—*ABCA1*, *APOA1* and *LCAT*—and new and promising gene candidates such as *PCSK9*, *LRP6*, *ANGPTL4* and *NPC1L1* that influence plasma lipoproteins.

EARLY GENETIC STUDIES AND PLASMA LIPOPROTEINS: A BRIEF REVIEW

Abnormalities in the synthesis, processing and catabolism of lipoprotein particles can act alone or in combination to create variation in lipoproteins. Molecules critical in these processes include the various apolipoproteins, cellular receptors, and enzymes involved in the lipoprotein metabolism pathways. From these biochemically defined determinants, candidate genes are identified for study to determine their association with plasma lipoprotein variations. Hundreds of candidate gene association studies have been performed over the past 25 years, with inconsistent results. Genetic polymorphisms in apolipoprotein (apo) B and apoE (*APOB* and *APOE*) and in the LDL receptor (*LDLR*) have been studied for association with plasma LDL-cholesterol levels, and of these, only *APOE* polymorphisms have shown consistent associations.[2] Similarly, genetic polymorphisms in the major protein of HDL cholesterol, apoA-I (*APOA1*), have been evaluated for association with plasma HDL-cholesterol levels,[2] as have polymorphisms in the gene coding for lipoprotein lipase (*LPL*), the major triglyceride-hydrolyzing enzyme in plasma, with inconsistent results.[3] Since 1982, however, only a small portion of the genetic basis of lipoprotein variation has been actually uncovered by association studies. When you consider that most of the observed effects of these candidate genes on lipoprotein metabolism are small and that replication has often been difficult, it can be understood why the impact of these studies on clinical cardiology to date has been minimal.

GENETIC ASSOCIATION STUDIES: 2005 TO PRESENT

The majority of the more-recent genetic association reports using case–control designs have not substantially expanded the spectrum of valid and replicable genetic determinants of plasma lipoproteins (Tables 1–3). Significant genetic association has been reported in at least two studies for the *APOA1/C3/A4/A5* gene cluster, as well as for *APOB*, *APOE*, the cholesteryl ester transfer protein (*CETP*), and the hepatic lipase (*LIPC*) genes—all previously mentioned in earlier comprehensive reviews.[1,2] Significant genetic associations with cholesterol levels have also been consistently reported for the ATP-binding cassette type A1 (*ABCA1*) and proprotein convertase subtilisin/kexin type 9 (*PCSK9*) genes. These genetic association studies mainly involve single-nucleotide polymorphisms (SNPs) within candidate genes. The most commonly examined genes are *APOE* and *CETP*; recent analyses involving more than 12,000 individuals reported strong associations between the *APOE* ε2/ε3/ε4 alleles and LDL-cholesterol level ($P<0.0001$),[4] and the *CETP* TaqIB polymorphism and HDL-cholesterol level ($P<0.0001$).[5]

Although the biological validity of the *APOE* ε4 association with higher concentrations of LDL cholesterol and apoB-containing lipoproteins is not in doubt, the potential clinical advantage of using *APOE* genotyping above simple plasma LDL-cholesterol measurement remains uncertain. Furthermore, despite the genetic association between *CETP* and plasma HDL cholesterol[5]—which was part of the underlying rationale for targeting CETP in order to raise HDL-cholesterol levels pharmacologically[6]—clinical trials of torcetrapib, an inhibitor of CETP activity, showed that targeting this mechanism did not reduce the number of deaths due to coronary heart disease.[7]

Common alleles with moderate-to-large effects on plasma lipoproteins

The *APOE* genetic polymorphism is an example of a common allele that has been consistently associated with variation in plasma lipoprotein levels. The *APOE* Cys112Arg/Arg158Cys polymorphism underlies the differences in the electrophoretic protein isoforms that were designated E4, E3 and E2. Over the past 25 years, apoE isoforms have consistently been shown to be associated with variation in plasma LDL-cholesterol and apoB levels, with E4 having a greater influence than E3, and in turn, E3 having a greater influence than E2, across a 10–15% range.[8] This rather substantial effect is important clinically, because plasma LDL-cholesterol level is an unequivocal risk factor for cardiovascular disease, especially CHD. The genetically determined 5–7% difference in LDL-cholesterol level from the reference (wild type) ε3/ε3 genotype in carriers of either the ε4 (higher LDL-cholesterol levels) or ε2 alleles (lower LDL-cholesterol levels) becomes even more important in light of the fact that only approximately 50% of individuals in most populations have the ε3/ε3 genotype, with the remainder carrying at least one ε4 or ε2 allele. Investigators have subsequently searched—albeit relatively fruitlessly—for comparable genetic effects on plasma lipoproteins.

In the 1990s, investigations of candidate genes in families, clinical cohorts and epidemiological study subpopulations frequently claimed to find large effects on plasma lipoprotein concentrations. Often-cited examples include the *APOA1/C3/A4* gene cluster and *LIPC* gene, each initially reported to account for 20–25% of inter-individual variation in plasma HDL-cholesterol level,[9] and the cholesterol 7α-hydroxylase (*CYP7A*) gene, originally reported to account for 15% of the variation in LDL-cholesterol concentration.[10] On a molecular level, the associations

Table 1 Single-nucleotide polymorphisms showing association with HDL cholesterol (2005–2006).

Gene	Polymorphism	*P* value	Sample size (nationality)
ABCA1	Arg219Lys (1051G>A)[42]	0.04	969 (American, multi-ethnic)
	rs2515602[43]	<0.05 (African American)	3,999 CARDIA study (African American, Caucasian)
	Gly596Ala[44]	0.02	3,302 CAD patients (American)
	Val771Met (–14C>T)[45]	<0.05	>2,300 (Turkish)
APOA1/C3/A4/A5	*APOA4*: 347T>A[46]	0.017	902 diabetic women (American)
APOA5	Val153Met[47]	<0.01 (females)	2,559 (Czech)
	–1131T>C[48]	0.013 (controls)	211 CAD patients; 677 controls (Chinese)
APOE	[a]E2/E3/E4[49]	0.006 (girls)	439 children (Finnish)
	E2/E3/E4[4]	≤0.02	12,491 ARIC study (African American, Caucasian)
	E2/3/4**Taq*IB(*CETP*)[50]	0.042	550 (Spanish)
	E2/E3/E4[51]	0.005	605 (Greek)
CETP	[a]*Taq*IB/–629C>A/Ile405Val[52]	<0.04	171 (South Indian)
	*Taq*IB[52]	<0.05 (men)	171 (South Indian)
	rs708272[43]	<0.02	3,999 CARDIA study (African American, Caucasian)
	–629C>A, Ile405Val, *Taq*IB[53]	<0.001	8,289 (Dutch)
	*Taq*IB[54]	0.049 (men), <0.001 (women)	1,729 (Japanese)
	TaqIB[44]	<0.001	3,219 CAD patients (North American)
	–1337C>T[55]	0.0001	534 (French)
	TaqIB[56]	<0.05 (diabetic); <0.001 (controls)	187 diabetic; 2,011 controls (Turkish)
	–629C>A[57]	<0.001	7,083 PREVEND study (North American)
	Ala373Pro[58]	0.0007	2,816 (Caucasian, North American)
	Ile405Val[58]	<0.0001	2,816 (Caucasian, North American)
	Arg451Glu[58]	0.003	2,816 (Caucasian, North American)
	*Taq*IB[5]	<0.0001	13,677 (meta-analysis)
	*Taq*IB[59]	0.003 (patients); 0.05 (controls)	173 CAD patients; 111 controls (Turkish)
LIPC	–480C>T[60]	<0.001	6,489 (Australian)
	–514C>T[61]	0.014 (African American)	302 (North American, multi-ethnic)

[a]Both LDL and HDL cholesterol levels affected. Abbreviations: *ABCA1,* ATP-binding cassette, type A1; *APOA1/C3/A4/A5,* apolipoprotein A1/C3/A4/A5 gene cluster; *APOA5,* apolipoprotein A5; *APOE,* apolipoprotein E; ARIC, Atherosclerosis Risk in Communities; CAD, coronary artery disease; CARDIA, Coronary Artery Risk Development in Young Adults; *CETP,* cholesteryl ester transfer protein; *LIPC,* hepatic lipase; PREVEND, Prevention of Renal and Vascular End Stage Disease.

between these candidate genes and plasma lipoproteins were attributed to functional SNPs in the promoter sequences of the genes.[10,11] In retrospect, however, these studies of a few common SNPs markedly overestimated the effect of these polymorphisms on plasma lipoproteins. Many replication studies failed to demonstrate genetic effects of comparable size.[12,13] Past overestimation of genetic effects could have arisen partly from the use of estimates of heritability within family units derived using standard deviation, rather than by determining the percentage difference between mean values in different communities or populations.

Common alleles with small genetic effects on plasma lipoproteins

Within certain families and isolated communities, the effect of a single candidate gene on lipoprotein metabolism can be profound, as exemplified by mutations in the *LDLR* gene that cause familial hypercholesterolemia (FH) and premature atherosclerosis.[14] At a population level, however, the effects of variants in most candidate genes are small to undetectable.

Small genetic effects of common variants might represent a more realistic genetic model for plasma lipoprotein variation.[15] An example is a common variation in the apoC-III (*APOC3*)

Table 2 Single-nucleotide polymorphisms showing association with LDL cholesterol (2005–2006).

Gene	Polymorphism	*P* value	Sample size (nationality)
APOA5	553G>T[62]	0.044 (diabetics); 0.022 (controls)	71 diabetic; 152 controls (Chinese)
APOB	*Eco*RI[63]	<0.001 (CAD patients)	150 CAD patients; 100 controls (Turkish)
APOE	[b]E2/E3/E4[49]	0.006 (girls))	439 children (Finnish)
	E2/E3/E4[49]	0.007 (boys)	439 children (Finnish)
	E2/E3/E4[64]	0.008 (males)	351 (Slovak)
	E2/E3/E4[4]	<0.0001	12,491 ARIC study (African American, Caucasian)
	E2/E3/E4/–219G>T/[65]	0.007 (males)	525 (Finnish)
	[a]E2/E3/E4[66]	<0.05 (hypertensive)	185 hypertensive; 200 controls (Indian)
	[a]E2/E3/E4[67]	<0.05	1,900 (Korean)
	E2/E3/E4[68]	<0.0001	348 girls (Vietnamese)
CETP	[b]*Taq*IB/–629C>A/Ile405Val[52]	<0.04	171 (South Indian)
	[a]*Taq*IB[69]	<0.05	29 renal transplant patients; 29 controls (Turkish)
PCSK9	Arg46Leu[70]	<0.0001 (Caucasian)	3,543 Dallas Heart Study (African American, Caucasian)
	Leu253Phe, Ala443Thr, His553Arg[70]	<0.05 (African American)	3,543 Dallas Heart Study (African American, Caucasian)
	Glu670Gly[71]	0.005	372 LCAS study (American)

[a]Both total cholesterol and LDL cholesterol levels affected. [b]Both LDL and HDL cholesterol levels affected. Abbreviations: *APOA5,* apolipoprotein A5; *APOB,* apolipoprotein B; *APOE,* apolipoprotein E; ARIC,Atherosclerosis Risk in Communities; CAD, coronary artery disease; *CETP,* cholesteryl ester transfer protein; LCAS, the Lipoprotein and Coronary Atherosclerosis Study; *PCSK9,* proprotein convertase subtilisin/kexin type 9 serine protease.

Table 3 Single-nucleotide polymorphisms showing association with total cholesterol (2005–2006).

Gene	Polymorphism	*P* value	Sample size (nationality)
APOA1/C3/A4/A5	13 SNPs[72]	NA (<0.05)	3,831(African American, Caucasian)
APOB	*Eco*RI[63]	<0.05 (CAD patients)	150 CAD patients; 100 controls (Turkish)
	Glu4154Lys[73]	<0.01 (youth)	413 youth, young adults (African American, Caucasian)
APOE	113G>C[65]	0.012 (males)	525 (Finnish)
	[a]E2/E3/E4[66]	<0.05 (hypertensive)	185 hypertensive; 200 controls (Indian)
	[a]E2/E3/E4[67]	<0.05	1,900 (Korean)
CETP	[a]*Taq*IB[69]	<0.05	29 renal transplant patients; 29 controls (Turkish)

[a]Both total cholesterol and LDL cholesterol levels affected. Abbreviations: *APOA1/C3/A4/A5,* apolipoprotein A1/C3/A4/A5 gene cluster; *APOB,* apolipoprotein B; *APOE,* apolipoprotein E; CAD, coronary artery disease; *CETP,* cholesteryl ester transfer protein; NA, not available; SNP, single-nucleotide polymorphism.

promoter that causes slightly elevated plasma triglyceride concentrations,[16] which are often within the physiologic range and are sometimes too small to be detected in association analyses. Such sequence variants might lead to modestly altered expression of qualitatively normal gene products that have a cumulative effect on plasma lipoprotein concentrations over the long term. For example, when secondary factors are present, altered expression of *APOC3* predisposes individuals to hypertriglyceridemia.[17] Genetic variants that produce small effects on promoter function might, thus, be one component of the predisposition to complex diseases. Many small effects in aggregate could contribute to a background of susceptibility that, under appropriate conditions, leads to the development of dyslipidemia.

Rare alleles with large effects at extremes of quantitative traits

Contemporary technological platforms and study designs have helped define a revised model of the impact of genomic variants in quantitative lipoprotein traits. Now, it seems that a proportion of individuals in the general population, at the extremes of the Gaussian distribution of a continuous trait, carry dysfunctional variants of genes, which were originally examined because

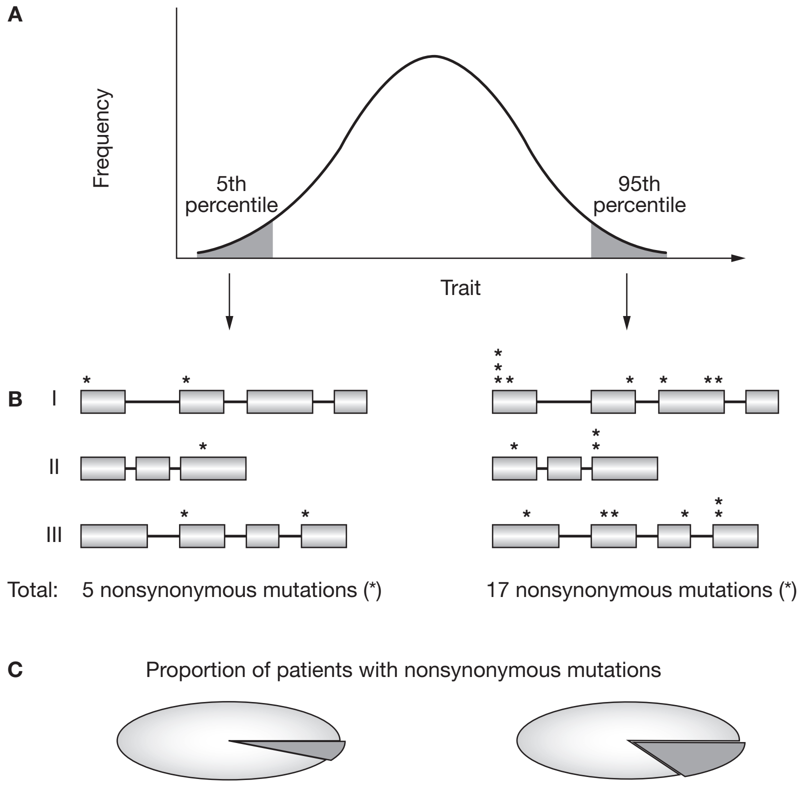

Figure 1 'Missense-accumulation' analysis. (**A**) The frequency distribution of a quantitative lipoprotein trait is shown at the top of the figure. Individuals at the extremes (in this case, the 5th and 95th percentiles) of the population distribution are selected for further study and their genomic DNA is analyzed using high-throughput automated DNA sequencing, focusing on the coding regions of specific candidate genes. (**B**) These genes (I, II and III) are selected for study for a variety of reasons, including clues from animal models, biochemical studies, genome-wide studies, or in this case, because homozygous mutations cause extremely rare diseases, perhaps at the extreme >99.9% tail of the trait's distribution. Sequence analysis focuses on coding regions (exons represented by rectangles). DNA sequence analysis identifies both neutral variants and potential functional variants for which one criterion is whether the amino acid coding sequence is altered (i.e. 'missense' or 'nonsynonymous' variants). (**C**) The cumulative total or, more precisely, proportion of individuals in each tail of the trait carrying missense mutations is compared and a statistical difference suggests an association between the candidate gene(s) and the trait. Some argue that most statistically associated mutations detected by this strategy are very likely to be dysfunctional,[41] although studies of function can be performed.

they underlie very rare monogenic disorders. This model probably reflects the biological reality of genetic effects on complex traits in the general population. For example, DNA sequence analysis of individuals with atypical type 2 diabetes showed that up to 10% actually carry gene mutations that cause a rare monogenic subtype of diabetes called maturity onset diabetes of the young (MODY).[18] A similar strategy has been used to examine patient subgroups defined by phenotypic thresholds, such as 'low HDL', 'high HDL', 'low LDL' or 'high LDL'. As a result of this study design, the original concept that a few common SNPs explain a large proportion of variation in plasma lipoprotein traits has been developed into a newer model in which multiple rare missense mutations are carried in a substantial minority of individuals at the extremes of the distribution of plasma lipoprotein levels in the general population (Figure 1).

Candidate genes studied were those for highly penetrant, but rare, monogenic disorders defined by extreme variation of the trait of interest. Rare homozygous mutations in *APOA1* and *LCAT*, the gene for lecithin-cholesterol acyltransferase, for example, underlie extremely rare states of complete HDL-cholesterol deficiency,[19] and homozygous mutations in *ABCA1* (the gene for ATP-binding cassette, sub-family A, member 1) cause Tangier disease,[20] an extremely rare disorder characterized by undetectable levels of plasma HDL cholesterol and some systemic manifestations. Mutations in these three genes also underlie some cases of familial hypoalphalipoproteinemia, a rare phenotype characterized by extremely low plasma HDL-cholesterol levels.[20] Cohen *et al.* used automated genomic DNA sequence analysis to test whether rare DNA sequence variants in *APOA1, LCAT* and *ABCA1* collectively contribute to low plasma HDL-cholesterol levels in individuals from a population-based study.[21] The study design identified cases and controls on the basis of HDL-cholesterol levels in or below the 5th and above the 95th percentile for age and sex. Nonsynonymous coding sequence variants in these three genes were significantly more common among individuals with HDL-cholesterol levels below the 5th percentile than in those with HDL-cholesterol levels above the 95th percentile in two independent population samples (16% versus 2%).[21] Furthermore, most variants in the low HDL-cholesterol group were imputed to have a functional consequence. These findings established the paradigm that multiple rare alleles with major phenotypic effects underlie a substantial minority of cases of low plasma HDL-cholesterol in the general population.

Mutations in *PCSK9* can cause both very low and very high plasma LDL-cholesterol levels. Studies in families with FH not caused by an *LDLR* mutation were the first to show that heterozygous gain-of-function mutations in *PCSK9* were causative.[22] Subsequently, *PCSK9* loss-of-function mutations were identified by sequencing the gene's coding region in individuals with low plasma LDL-cholesterol levels.[23] Two *PCSK9* nonsense mutations, namely Tyr142X and Cys679X, had a combined frequency of approximately 2% in individuals of African ancestry and less than 0.1% in those with European ancestry.[23] Being heterozygous for the mutations was associated with a 40% reduction in plasma LDL-cholesterol concentration. A subsequent population survey showed that 2.6% of individuals of African ancestry had *PCSK9* nonsense mutations that were associated with reductions of 28% and 88% in mean LDL-cholesterol level and CHD risk, respectively.[24] Somewhat less dramatic findings were observed in those of European descent.

These results defined both a new pathway of cholesterol regulation and, therefore, a new drug target.[25] *PCSK9* encodes a protease that most probably degrades the LDL receptor, limiting excessive cellular uptake of LDL cholesterol. As statins inhibit cholesterol biosynthesis, activate sterol regulatory element binding proteins and increase the synthesis of both the LDL receptor and PCSK9, pharmacologic inhibition of PCSK9 might lower LDL-cholesterol in a manner synergistic to that of statins. Some researchers have suggested that lifelong low LDL-cholesterol levels in carriers of *PCSK9* mutations together with the low incidence of CHD observed in these individuals provide support for the pharmacological lowering of LDL-cholesterol from a very young age in those at risk of CHD.[24,25] The design of the study that brought *PCSK9* and CHD risk to light,[24] however, was not truly prospective, and vascular outcomes could have been affected by unmeasured confounding environmental and lifestyle factors, such as diet and exercise.

The rare allele–large effect concept has also been demonstrated by the recent discovery of a LDL receptor-related protein 6 (*LRP6*) mutation that is associated with high LDL-cholesterol and triglyceride levels and other metabolic syndrome phenotypes in a single family, although it has not yet been demonstrated in the general population.[26] Another study involving the resequencing of the angiopoietin-like 4 (*ANGPTL4*) gene in the general population reported rare variants that were cumulatively associated with low triglyceride and high HDL-cholesterol levels.[27] Together, such reports, which both involve novel genes in the lipoprotein field, reinforce the view that the genetics underlying the variation in lipoprotein levels are complex and comprise numerous rare variants.

NPC1L1: a new determinant of plasma LDL cholesterol

Niemann-Pick C1-like 1 protein (NPC1L1), is a fascinating example of how elucidation of the target of a pharmacologic agent—the cholesterol absorption inhibitor ezetimibe—served to expose a key physiological pathway and genetic influences on plasma lipoproteins. Since 2003, evidence garnered from human, animal, and cell studies has clearly indicated that NPC1L1 has a key role in the intestinal absorption of sterols, including plant sterols and cholesterol.[28] Human genetic studies showed that NPC1L1 had a role in cholesterol absorption and was the target of ezetimibe. Studies showed, for example, that a nonresponder to ezetimibe was a compound heterozygote for two rare nonsynonymous *NPC1L1* mutations.[29] Further research has revealed that common SNPs in *NPC1L1* also underlie more subtle interindividual differences in plasma LDL-cholesterol levels in response to ezetimibe,[30,31] strengthening the case for NPC1L1 as the target for ezetimibe. Furthermore, a population-based study that examined multiple, rare *NPC1L1* sequence variants found interindividual differences in both intestinal sterol absorption and baseline plasma concentrations of LDL cholesterol.[32] Until direct functional assessment of these common or rare variants has been performed, no definitive conclusions about the gene mutations and NPC1L1 function can be made. Evidence, however, supports NPC1L1 as a physiological determinant of intestinal sterol absorption in humans and of plasma lipoprotein levels, both in the fasting state and in response to medications.

THE IMPORTANCE OF TECHNOLOGICAL ADVANCES

Automated genomic DNA sequencing was the key technological advance in the discovery of rare dysfunctional variants among individuals with extreme lipoprotein phenotypes. Further refinements to this method have since increased its throughput and reduced the associated costs. These advances permit comprehensive analysis of nucleotides within a specified region of genomic DNA, rather than interrogation of a just few selected SNPs. Although whole genome sequencing is not yet used as a general screening method, this technique will soon become cost-effective (e.g. the much touted ‘$1,000 genome’, which some researchers believe will be available by 2010[33]). The availability of this technology will allow genome-wide detection of interindividual variation at the nucleotide level. A vast quantity of nucleotide data will be generated for each person tested, and this data will require processing using bioinformatic approaches that compare each nucleotide in the genome to a 'consensus' reference sequence. This approach elucidates the total genome-wide individual burden of variation as compared with a 'consensus' genome. The vast majority of sequence (i.e. sequence that is invariant to the 'consensus' control genome) can then be excluded from further consideration. The variant sequences identified can then be further classified according to predicted phenotypic effects, ranging from 'neutral' to 'likely to be functional', on the basis of similarity to archived disease-causing mutations or through inference using models to predict the biological consequences of such sequence variants. The prioritized genomic-sequence variants from an individual could then be included in various types of studies to detect genotype–phenotype relationships. In the near future, whole-genome sequencing will serve, in effect, the role that genotyping serves today. Newly identified forms of variation, such as copy number variation,[34] will also need to be itemized in each individual genome.

GENETIC STUDIES: WHAT IS THEIR BENEFIT TO CLINICAL CARDIOLOGY?

Geneticists have promoted the idea that findings from human research will bear fruit through a number of scenarios. First, findings will provide new understanding of biological and biochemical mechanisms, enabling identification of new targets for interventional strategies. Second, findings will facilitate the development of new diagnostic tools and, third, new prognostic tools. The potential clinical benefit from all three scenarios is large, but is the suggestion that genetic studies can produce clinically useful results justified? The issue of whether genomics can help discover new pathways is distinct from whether genomic information is useful in a clinical setting. Genomics is clearly valuable as a means to achieve a more complete understanding of biology, biochemistry and physiology. The definition of new pathways and targets is the first step towards drug design, which can result in clinical trials and evidence-based changes in practice. The notion that routine screening with a panel of genomic-based markers can add prognostic or diagnostic value, however, is less convincing.

Another claim that supports the continued discovery of common gene variants that contribute to variation in levels of putative atherosclerotic risk factors is Mendelian epidemiology—that is, an epidemiologically associated factor with a genetic basis is less likely to be confounded in its association with CHD than acquired variations.[35] This relatively novel concept suggests that knowledge of genotype is a better predictor of disease than an intermediate biochemical phenotype. Nevertheless, this concept requires validation through research and clinical application. *PCSK9* variants, for example, reduce LDL-cholesterol levels and also seem to reduce CHD risk.[36] By contrast, *APOE* ε4 is consistently associated with raised LDL-cholesterol levels, yet its relationship with CHD has been less consistent.[37]

Will genotyping *APOE* isoforms or screening for *PCSK9* mutations add clinically useful information above simple biochemical determination of LDL-cholesterol levels? Some have proposed that genomic markers added to a panel of diagnostic tests will predict future risk of atherosclerosis.[38,39] Markers that cannot be measured in any way other than by genomic technologies might one day be added to diagnostic algorithms, but established biomarkers, such as LDL-cholesterol concentration, represent the integration of many genetic and environmental effects—not all of them known and measurable. Even with 'complete' genomic information, confounders such as genetic pleiotropy (i.e. causation by different genes), epistasis (i.e. interactions between genes underlying a trait), influence of genetic background, or gene–environment interactions, could complicate a hypothesized direct relationship between a genotype and phenotype. Any attempt to account for all inputs into an individual's plasma LDL-cholesterol concentration will, therefore, be incomplete at best. Biochemically measured LDL-cholesterol levels have already proven to be a valid predictor of risk and a target for therapy. The possible clinical value of an adjunctive genetic test for LDL-cholesterol level is, at best, unclear.

Though the clinical utility of general genetic screening remains unresolved at present, there are some situations in which genetic research might prove beneficial in a clinical setting. The early diagnosis of FH by testing for mutations in *LDLR*, for example, or the identification of individuals who are homozygous or compound heterozygous for FH, or have familial lipoprotein lipase deficiency, and are therefore candidates for potential gene therapy. Although identification of SNPs influencing LDL-cholesterol levels might not be more predictive than measuring

LDL-cholesterol concentrations, SNP genotyping could have an important role for assessing other atherosclerotic risk factors that cannot be measured as biochemical surrogates. Complete risk prediction cannot be achieved though genetics alone, but evidence indicates that genetic data could provide earlier or enhanced prediction even when traditional risk factors are considered.[40]

CONCLUSIONS

Progress in elucidating the genetic influence on plasma lipoprotein concentrations has included definition of both the underlying genes and the nature of the genetic contribution. Over the past decade, the potential effects a single genetic variable can have on lipoprotein metabolism and our expectations of the size of the effects have been down-sized. This more realistic understanding of the genetic determinants of plasma lipoproteins is reflected in more strategic experimental designs, larger sample sizes, required replicability and a more pragmatic long-term expectation of the utility of genetic markers. Classification of individuals who are susceptible to atherosclerosis using genetic information might still be possible, especially if newer strategies such as high-throughput genome-wide association analysis find as-yet undiscovered candidate genes that are important determinants of disease. Prediction of the precise evolution of dyslipidemia in any individual might be impossible, however, because of the confounding influence of varying interindividual environments, nonlinear interactions between genes and environment, and even because of biological chaos influencing pathogenesis.

KEY POINTS

- Differences in plasma lipoprotein concentrations throughout a population reflect the different environmental and genetic factors affecting each individual
- The contribution that a single genetic variable might be expected to have on lipoprotein concentration has been down-sized over the past decade
- A recent strategy used to identify genes associated with plasma lipoproteins is ‘missense-accumulation’ analysis—the identification of rare alleles with large effects at extremes of quantitative traits
- Although genetic research has proven useful in some settings, the clinical utility of ‘across the board’ genetic screening remains somewhat unsettled at present

References

1 Lusis AJ *et al.* (2004) Genetics of atherosclerosis. *Annu Rev Genomics Hum Genet* **5:** 189–218
2 Dammerman M and Breslow JL (1995) Genetic basis of lipoprotein disorders. *Circulation* **91:** 505–512
3 Talmud PJ and Humphries SE (2001) Genetic polymorphisms, lipoproteins and coronary artery disease risk. *Curr Opin Lipidol* **12:** 405–409
4 Volcik KA *et al.* (2006) Apolipoprotein E polymorphisms predict low density lipoprotein cholesterol levels and carotid artery wall thickness but not incident coronary heart disease in 12,491 ARIC study participants. *Am J Epidemiol* **164:** 342–348
5 Boekholdt SM *et al.* (2005) Cholesteryl ester transfer protein TaqIB variant, high-density lipoprotein cholesterol levels, cardiovascular risk, and efficacy of pravastatin treatment: individual patient meta-analysis of 13,677 subjects. *Circulation* **111:** 278–287
6 Brousseau ME *et al.* (2004) Effects of an inhibitor of cholesteryl ester transfer protein on HDL cholesterol. *N Engl J Med* **350:** 1505–1515
7 Nissen SE *et al.* (2007) Effect of torcetrapib on the progression of coronary atherosclerosis. *N Engl J Med* **356:** 1304–1316
8 Walden CC and Hegele RA (1994) Apolipoprotein E in hyperlipidemia. *Ann Intern Med* **120:** 1026–1036
9 Cohen JC *et al.* (1994) Variation at the hepatic lipase and apolipoprotein AI/CIII/AIV loci is a major cause of genetically determined variation in plasma HDL cholesterol levels. *J Clin Invest* **94:** 2377–2384
10 Wang J *et al.* (1998) Linkage between cholesterol 7α-hydroxylase and high plasma low-density lipoprotein cholesterol concentrations. *J Clin Invest* **101:** 1283–1291
11 Guerra R *et al.* (1997) A hepatic lipase (LIPC) allele associated with high plasma concentrations of high density lipoprotein cholesterol. *Proc Natl Acad Sci USA* **94:** 4532–4537
12 Hegele RA *et al.* (1999) Absence of association between genetic variation in the LIPC gene promoter and plasma lipoproteins in three Canadian populations. *Atherosclerosis* **146:** 153–160
13 Hegele RA *et al.* (2001) Variable association between genetic variation in the CYP7 gene promoter and plasma lipoproteins in three Canadian populations. *Atherosclerosis* **154:** 579–587
14 Yuan G *et al.* (2006) Heterozygous familial hypercholesterolemia: an underrecognized cause of early cardiovascular disease. *CMAJ* **174:** 1124–1129
15 Hegele RA (1997) Small genetic effects in complex diseases: a review of regulatory sequence variants in dyslipoproteinemia and atherosclerosis. *Clin Biochem* **30:** 183–188
16 Li WW *et al.* (1995) Common genetic variation in the promoter of the human apo CIII gene abolishes regulation by insulin and may contribute to hypertriglyceridemia. *J Clin Invest* **96:** 2601–2605
17 Pollex RL *et al.* (2006) Metabolic syndrome in aboriginal Canadians: prevalence and genetic associations. *Atherosclerosis* **184:** 121–129
18 McKinney J *et al.* (2003) Maturity-onset diabetes of the young (MODY) mutation in type 2 diabetes and latent autoimmune diabetes of the adult. *Diabetes Care* **26:** 3358–3359
19 Funke H (1997) Genetic determinants of high density lipoprotein levels. *Curr Opin Lipidol* **8:** 189–196
20 Brooks-Wilson A *et al.* (1999) Mutations in ABC1 in Tangier disease and familial high-density lipoprotein deficiency. *Nat Genet* **22:** 336–345
21 Cohen JC *et al.* (2004) Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science* **305:** 869–872

22 Abifadel M *et al.* (2003) Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nat Genet* **34:** 154–156
23 Cohen J *et al.* (2005) Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. *Nat Genet* **37:** 161–165
24 Cohen JC *et al.* (2006) Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. *N Engl J Med* **354:** 1264–1272
25 Tall AR (2006) Protease variants, LDL, and coronary heart disease. *N Engl J Med* **354:** 1310–1312
26 Mani A *et al.* (2007) LRP6 mutation in a family with early coronary disease and metabolic risk factors. *Science* **315:** 1278–1282
27 Romeo S *et al.* (2007) Population-based resequencing of ANGPTL4 uncovers variations that reduce triglycerides and increase HDL. *Nat Genet* **39:** 513–516
28 Huff MW *et al.* (2006) NPC1L1: evolution from pharmacological target to physiological sterol transporter. *Arterioscler Thromb Vasc Biol* **26:** 2433–2438
29 Wang J *et al.* (2005) Compound heterozygosity for two non-synonymous polymorphisms in NPC1L1 in a non-responder to ezetimibe. *Clin Genet* **67:** 175–177
30 Hegele RA *et al.* (2005) NPC1L1 haplotype is associated with inter-individual variation in plasma low-density lipoprotein response to ezetimibe. *Lipids Health Dis* **4:** 16
31 Simon JS *et al.* (2005) Sequence variation in NPC1L1 and association with improved LDL-cholesterol lowering in response to ezetimibe treatment. *Genomics* **86:** 648–656
32 Cohen JC *et al.* (2006) Multiple rare variants in NPC1L1 associated with reduced sterol absorption and plasma low-density lipoprotein levels. *Proc Natl Acad Sci USA* **103:** 1810–1815
33 Service RF (2006) Gene sequencing: the race for the $1,000 genome. *Science* **311:** 1544–1546
34 Redon R *et al.* (2006) Global variation in copy number in the human genome. *Nature* **444:** 444–454
35 Davey Smith G and Ebrahim S (2003) 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* **32:** 1–22
36 Brown MS and Goldstein JL (2006) Biomedicine: lowering LDL—not only how low, but how long? *Science* **311:** 1721–1723
37 Letonja M *et al.* (2004) Apolipoprotein E gene polymorphism effects triglycerides but not CAD risk in Caucasian women younger than 65 years. *Ann Genet* **47:** 147–153
38 Hegele RA (1997) Candidate genes, small effects, and the prediction of atherosclerosis. *Crit Rev Clin Lab Sci* **34:** 343–367
39 Topol EJ (2005) Simon Dack Lecture: the genomic basis of myocardial infarction. *J Am Coll Cardiol* **46:** 1456–1465
40 Humphries SE *et al.* (2007) Candidate gene genotypes, along with conventional risk factor assessment, improve estimation of coronary heart disease risk in healthy UK men. *Clin Chem* **53:** 8–16
41 Kryukov GV *et al.* (2007) Most rare missense alleles are deleterious in humans: implications for complex disease and association studies. *Am J Hum Genet* **80:** 727–739
42 Benton JL *et al.* (2006) Associations between two common polymorphisms in the ABCA1 gene and subclinical atherosclerosis: Multi-Ethnic Study of Atherosclerosis (MESA). *Atherosclerosis* **193:** 352–360
43 Klos KL *et al.* (2006) Consistent effects of genes involved in reverse cholesterol transport on plasma lipid and apolipoprotein levels in CARDIA participants. *Arterioscler Thromb Vasc Biol* **26:** 1828–1836
44 Whiting BM *et al.* (2005) Candidate gene susceptibility variants predict intermediate end points but not angiographic coronary artery disease. *Am Heart J* **150:** 243–250
45 Hodoglugil U *et al.* (2005) Common polymorphisms of ATP binding cassette transporter A1, including a functional promoter polymorphism, associated with plasma high density lipoprotein cholesterol levels in Turks. *Atherosclerosis* **183:** 199–212
46 Qi L *et al.* (2006) Associations of the apolipoprotein A1/C3/A4/A5 gene cluster with triglyceride and HDL cholesterol levels in women with type 2 diabetes. *Atherosclerosis* **192:** 204–210
47 Hubacek JA *et al.* (2005) Sex-specific effect of APOAV variant (Val153>Met) on plasma levels of high-density lipoprotein cholesterol. *Metabolism* **54:** 1632–1635
48 Hsu LA *et al.* (2006) Genetic variations of apolipoprotein A5 gene is associated with the risk of coronary artery disease among Chinese in Taiwan. *Atherosclerosis* **185:** 143–149
49 Salminen M *et al.* (2006) Apolipoprotein E polymorphism and changes in serum lipids during a family-based counselling intervention. *Public Health Nutr* **9:** 859–865
50 Sorli JV *et al.* (2006) The effect of the APOE polymorphism on HDL-C concentrations depends on the cholesterol ester transfer protein gene variation in a Southern European population. *Clin Chim Acta* **366:** 196–203
51 Kolovou GD *et al.* (2005) Association of apolipoprotein E genotype with early onset of coronary heart disease in Greek men. *Angiology* **56:** 663–670
52 Padmaja N *et al.* (2007) Common variants of Cholesteryl ester transfer protein gene and their association with lipid parameters in healthy volunteers of Tamilian population. *Clin Chim Acta* **375:** 140–146
53 Borggreve SE *et al.* (2006) An increased coronary risk is paradoxically associated with common cholesteryl ester transfer protein gene variations that relate to higher high-density lipoprotein cholesterol: a population-based study. *J Clin Endocrinol Metab* **91:** 3382–3388
54 Tsujita Y *et al.* (2006) The association between high-density lipoprotein cholesterol level and cholesteryl ester transfer protein TaqIB gene polymorphism is influenced by alcohol drinking in a population-based sample. *Atherosclerosis* **191:** 199–205
55 Frisdal E *et al.* (2005) Functional interaction between -629C/A, -971G/A and -1337C/T polymorphisms in the CETP gene is a major determinant of promoter activity and plasma CETP concentration in the REGRESS Study. *Hum Mol Genet* **14:** 2607–2618
56 Hodoglugil U *et al.* (2005) An interaction between the TaqIB polymorphism of cholesterol ester transfer protein and smoking is associated with changes in plasma high-density lipoprotein cholesterol levels in Turks. *Clin Genet* **68:** 118–127
57 Borggreve SE *et al.* (2005) The effect of cholesteryl ester transfer protein -629C->A promoter polymorphism on high-density lipoprotein cholesterol is dependent on serum triglycerides. *J Clin Endocrinol Metab* **90:** 4198–4204
58 Lloyd DB *et al.* (2005) Cholesteryl ester transfer protein variants have differential stability but uniform inhibition by torcetrapib. *J Biol Chem* **280:** 14918–14922
59 Yilmaz H *et al.* (2005) Effects of cholesterol ester transfer protein Taq1B gene polymorphism on serum lipoprotein levels in Turkish coronary artery disease patients. *Cell Biochem Funct* **23:** 23–28
60 McCaskie PA *et al.* (2006) The C-480T hepatic lipase polymorphism is associated with HDL-C but not with risk of coronary heart disease. *Clin Genet* **70:** 114–121
61 Miljkovic-Gacic I *et al.* (2006) Lipoprotein subclass and particle size differences in Afro-Caribbeans,

African Americans, and white Americans: associations with hepatic lipase gene variation. *Metabolism* **55:** 96–102

62 Zhai G *et al.* (2006) Association of APOA5 c.553G>T polymorphism with type 2 diabetes mellitus in a Chinese population. *Clin Chem Lab Med* **44:** 1313–1316

63 Duman BS *et al.* (2005) Genetic variations of the apolipoprotein B gene in Turkish patients with coronary artery disease. *Ann Hum Biol* **32:** 620–629

64 Sivakova D *et al.* (2006) Apolipoprotein E polymorphism in relation to plasma lipid levels and other risk factors of atherosclerosis in two ethnic groups from Slovakia. *Coll Antropol* **30:** 387–394

65 Viiri LE *et al.* (2006) Relations of APOE promoter polymorphisms to LDL cholesterol and markers of subclinical atherosclerosis in young adults. *J Lipid Res* **47:** 1298–1306

66 Bhavani AB *et al.* (2005) Lipid profile and apolipoprotein E polymorphism in essential hypertension. *Indian Heart J* **57:** 151–157

67 Shin MH *et al.* (2005) The effect of apolipoprotein E polymorphism on lipid levels in Korean adults. *J Korean Med Sci* **20:** 361–366

68 Nghiem NT *et al.* (2004) Apolipoprotein E polymorphism in Vietnamese children and its relationship to plasma lipid and lipoprotein levels. *Metabolism* **53:** 1517–1521

69 Zeybek U *et al.* (2006) Effects of cholesteryl ester transfer protein TAQ1B polymorphism in renal transplant patients. *Transplant Proc* **38:** 1382–1384

70 Kotowski IK *et al.* (2006) A spectrum of PCSK9 alleles contributes to plasma levels of low-density lipoprotein cholesterol. *Am J Hum Genet* **78:** 410–422

71 Chen SN *et al.* (2005) A common PCSK9 haplotype, encompassing the E670G coding single nucleotide polymorphism, is a novel genetic marker for plasma low-density lipoprotein cholesterol levels and severity of coronary atherosclerosis. *J Am Coll Cardiol* **45:** 1611–1619

72 Hamon SC *et al.* (2006) Evidence for consistent intragenic and intergenic interactions between SNP effects in the APOA1/C3/A4/A5 gene cluster. *Hum Hered* **61:** 87–96

73 Davis CL *et al.* (2005) Genetic and environmental determinants of lipid profile in black and white youth: a study of four candidate genes. *Ethn Dis* **15:** 568–577

Acknowledgments
Support has been provided by the Jacob J Wolfe Distinguished Medical Research Chair, the Edith Schulich Vinet Canada Research Chair (Tier I) in Human Genetics, a Career Investigator award from the Heart and Stroke Foundation of Ontario, and operating grants from the Canadian Institutes for Health Research, the Heart and Stroke Foundation of Ontario, the Ontario Research Fund and by Genome Canada through the Ontario Genomics Institute.

Competing interests
The authors declared no competing interests.