Novel Genetic Determinants of High-Density Lipoprotein Cholesterol (HDL-C) in the French Canadian Population

Zari Dastani, MD

Department of Human Genetics

McGill University

Montreal, Canada

May 2009

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

i. Abstract

HDL-C levels, are an independent risk factor for heart disease, and are influenced by multiple genes and environmental factors. Choosing an isolated population could reduce the heterogeneity and increase power to discover susceptibility genes for low HDL-C. To identify novel genetic determinants that contribute to HDL-C levels, I used several genetic approaches on highly informative families of French Canadian descent with probands selected for HDL-C levels < 5^{th}% (age and sex specific). The estimated heritability of HDL-C for this cohort was 0.73. Using a candidate gene approach, I identified a novel mutation of *ApoAI* (apo AI_{E136X}) in three probands. Segregation analysis showed a strong co-segregation of the *ApoAI* allele with the low HDL-C trait in these kindreds.

I also performed whole genome linkage analysis using 485 markers situated at intervals of 6 centiMorgans (cM). In parametric two-point linkage analyses, the highest two-point LOD score of 4.6 was observed with the marker D4S424 on chromosome 4q31.21. I further refined the linked region from 12.2 cM to 2.9 cM (2.37 Mb) by genotyping 15 additional markers in the three families with the highest LOD-scores. None of the genes residing in the significantly restricted 2.37 Mb region has previously been associated with HDL-C metabolism.

In quantitative linkage analysis, I identified significant linkage to a locus on chromosome 16q23-24 that had been previously implicated in linkage scans for HDL-C. I examined this region for association using both family-based and

case-control analyses. Within an 18.1 cM region (7.8 Mb) four families demonstrated segregation. All coding regions and exon-intron boundaries of all genes within this region were sequenced. A missense variant in the *CHST6* gene showed segregation in four families. An association analysis that included all of the French Canadian families as well as other cohorts showed association for one common intronic SNP (rs2548861) in the *WWOX* gene, which is close to the *CHST6* gene. In addition, RT-PCR results from cultured cells demonstrated a significant difference in the expression level of *CHST6* and *KIAA1576* in the region. This study provides significant evidence for genes influencing HDL-C on chromosomes 4q31.21 and 16q23-24.

ii. Résumé

IL est bien établi qu'un niveau bas de HDL-C représente un facteur de risque des maladies coronariennes, le HDL-C est influencé par plusieurs gènes et facteurs environnementaux. Des études génétiques récentes ont montré que le ciblage d'une population bien définie pourrait réduire l'hétérogénéité et contribuerait à découvrir de nouveaux gènes de susceptibilité aux niveaux bas du HDL-C. Dans le but d'identifier de nouveaux gènes qui control la régulation du HDL-C, nous avons utilisé plusieurs approches génétiques sur les familles d'ascendance canadienne-française avec des probands sélectionnés pour les niveaux HDL-C <5 e percentile (âge et sexe). Nos analyses démontrent une augmentation de l 'héritabilité estimée de HDL-C dans cette cohorte (0.73). D'autre part, l'utilisation d'une approche gène candidat, nous a permis l'identification d'une nouvelle mutation de *l'apo AI* (apo AIE136X) dans trois familles de cet échantillon avec une forte co-ségrégation de l'allèle *ApoAI* avec un HDL-C bas.

J'ai également effectué une analyse de liaison du génome en utilisant 485 marqueurs situés à des intervalles d'environ 6 centiMorgans (cM). Nos résultas démontrent un LOD score de 4-6 sur la base de l'utilisation du marqueur D4S424 sur le chromosome 4q31.21.

Une analyse quantitative de liaison génétique a permis l'identification des liens avec un locus sur le chromosome 16q23-24 impliqués dans la régulation HDL-C. Nos résultas démontrent qu'une variante dans le gène *CHST6* (missence) montré une ségrégation dans quatre familles. D'autre part, l'analyse de toutes les

familles canadiennes-françaises, ainsi que d'autres cohortes ont montré une association commune pour un intron, SNP (rs2548861) dans le *WWOX* gène, chromosome 16q23.3 cartes-q24. En outre, les résultats de RT-PCR à partir de cellules en culture ont démontré une différence significative dans le niveau d'expression des *CHST6* et *KIAA1576*, un autre gène de la région. Cette étude mis en évidence les gènes ayant une influence significative sur le HDL-C sur les chromosomes 4q31.21 et 16q23-24. Cette etude permet une meilleure compréhension des facteurs génétiques de prédisposition à la maladie coronarienne.

iii. Table of Contents

iv. List of Figures

v. List of Tables

vi. List of Abbreviations

ABCA1, ATP binding cassette sub-family A1;

ABCG, ATP-binding cassette sub-family G;

ANGPTL4;, angiopoietin-like 4

ApoA, apolipoprotein A;

ApoAI, apolipoprotein A1;

ApoB, apolipoprotein B;

ApoC, apolipoprotein C;

ApoE, apolipoprotein E;

BMI, body mass index;

CAD, coronary artery disease;

CETP, cholesteryl ester transfer protein;

CM, chylomicrons;

CTCF, CCCTC-binding factor (zinc finger protein)-

DM, diabetes mellitus;

EL, endothelial lipase;

eNOS, endothelial NO synthase ;

FCHL, familial combined hyperlipidemia;

FH, familial history;

GALNT2, polypeptide N-acetylgalactosaminyltransferase 2;

GWAS, genome-wide association studies;

HDL, high density lipoproteins;

HDL-C, plasma high-density lipoprotein cholesterol;

HL, hepatic lipase;

HNF4A, hepatocyte nuclear factor 4 alpha;

HTN, hypertension;

IBD, identical by descent;

ICAM-1, intercellular adhesion molecule-1;

IDL, intermediate density lipoproteins

LCAT, lecithin:cholesterol acyltransferase;

LDL, low density lipoproteins

LDL-C, plasma low density lipoprotein cholesterol;

LP(a), lipoprotein a ;

MADD, MAP-kinase activating death domain-containing;

FOLH1, folate hydrolase 1;

MVK-MMAB, mevalonate kinase- cob(I)alamin adenosyltransferase precursor;

NO, nitric oxide;

PAF-AH, platelet-activating factor acetylhydrolase;

PAGGE, polyacrylamide gradient gel electrophoresis;

PBS, phosphate-buffered saline;

PON1, paraoxonase-1;

PON3, paraoxoanse-3;

QTL, quantitative trait locus;

RCT, reverse cholesterol transport ;

SMCs, smooth muscle cells;

SNP, single nucleotide polymorphism;

SR-BI, scavenger receptor B1;

T Chol, plasma total cholesterol;

TG, plasma triglycerides;

TTC39B, tetratricopeptide repeat domain 39B;

VLDL, very low density lipoproteins;

VCAM-1, vascular cell adhesion molecule-1;

vii. Contribution of Authors

As permitted under the "Guidelines for Thesis Preparation" provided by the Graduate and Postdoctoral Studies Office, chapter 3, 4, 5, and 6 of this thesis include manuscripts either published or submitted for publication. The contributions of all authors are listed below:

Chapter 3:

The contents of this chapter have been published as:

Zari Dastani, Carole Dangoisse, Betsie Boucher, Katia Desbiens, Larbi Krimbou, Robert Dufour, Robert A. Hegele, Päivi Pajukanta, James C. Engert, Michel Marcil, Jacques Genest. "*A Novel Mutation of Apolipoprotein A-I (Apo AI_{E136X}) Causes Low HDL cholesterol in French Canadians*." **Atherosclerosis**; 2006 March; 185(1): 127-136.

The contributions made by co-authors for this chapter:

I designed the experiments and carried out all the bench work, performed data analysis, and wrote the draft and tables and figures. Larbi Krimbou performed 2D gel electrophoresis and Michel Marcil carried out the LCAT assay. Katia Desbiens, Betsie Boucher and Carole Dangoisse provided technical support. James Engert supervised the sequencing work. The rest of the authors provided helpful suggestions and critically reviewed the manuscript.

Chapter 4:

The contents of this chapter have been published as:

Zari Dastani, Leigh Quiogue**,** James C. Engert, Michel Marcil, Jacques Genest, Paivi Pajukanta. "*Evidence for a Gene Influencing HDL-C on Chromosome 4q31.21*." **Journal of Atherosclerosis, Thrombosis and Vascular Biology;** 2006 Feb; 26(2): 392-7.

The contributions made by co-authors for this chapter:

I did all the experiments for fine mapping, and sequencing, as well as the data analysis, table and figure preparation and drafted the manuscript. Paivi Pajukanta and Leigh Quiogue performed parametric linkage analysis and made suggestions to revise the work. James Engert helped select the individuals for genotyping, and performed the quality control on the genotypes. The remaining authors provided helpful suggestions and critically reviewed the manuscript.

Chapter 5:

The contents of this chapter have been published as:

Zari Dastani, Isabelle L. Ruel, Carole Dangoisse, James C. Engert, Jacques Genest, Michel Marcil. "S*phingomyelinase Phosphodiesterase-1 (SMPD-1) Gene Polymorphisms and Haplotype analysis in Low HDL cholesterol French Canadian Population*." **BMC Medical Genetics;** 2007 Dec; 8: 79.

The contributions made by co-authors for this chapter:

I performed the genotyping and data analysis. Isabelle Ruel contributed to the genotyping. The remaining authors provided helpful suggestions and reviewed the manuscript.

Chapter 6:

The contents of this chapter have been published as:

Zari Dastani, Päivi Pajukanta, Michel Marcil, Nicholas Rudzicz, Isabelle Ruel, Swneke D. Bailey, Jenny C. Lee , Mathieu Lemire, Janet Faith, Jill Platko, John Rioux, Thomas J. Hudson, Daniel Gaudet, James C. Engert, Jacques Genest "F*ine Mapping and Association Studies of a High-Density Lipoprotein Cholesterol Linkage Region on Chromosome 16 in French-Canadian Subjects*." Resubmitted in revised form to the **European Journal of Human Genetics**; 2009

The contributions made by co-authors for this chapter:

I designed the experiments and carried out all the lab work including genotyping, candidate gene sequencing, and RT-PCR. I also analyzed the data including haplotying and prepared the tables and figures, and drafted the manuscript for this publication. Nicholas Rudzicz helped with the QTL analysis and the running of the SOLAR program. Swneke D. Bailey helped in making the figures. Isabelle Ruel provided technical support. Mathieu Lemire, Janet Faith, Jill Platko, John Rioux, Thomas J. Hudson, and Daniel designed the SLSJ study, generated the SLSJ genotypes and the provided SLSJ genome wide data. The remaining authors provided helpful suggestions and reviewed the manuscript.

viii. Acknowledgments

Fist of all, I would like to thank the **Research Institute of the MUHC**, the **McGill Faculty of Medicine** and the **Heart and stroke foundation of Canada** for financial support throughout my PhD.

I would like to thank my thesis committee members**, Dr. Patricia Tonin** and **Dr. Jacek Majwesky** for their expert advice.

I would like to thank my co-supervisor **Dr. James C. Engert**, who provided a fantastic guidance and encouragement during all stages of my study. Many thanks for all that you have taught me.

I would like to specially thank my supervisor and friend **Dr. Jacques Genest** for considering me as graduate student. It is frankly an honour for me to have you as a supervisor. Many thanks for all that you have taught me about science, academia and life in general.

I look forward to continuing collaboration with both of you.

I would like to say thank to all past and present people in the lab. I had lots of excitement moments in time with you. I would like to state special thanks to **Isabelle Ruel** and **Nick Rudzicz** for their incredible help in revising my thesis draft and **Larbi Krimbou** for his help in translating my abstract to French.

Finally, I would like to convey my thanks to my fantastic **family**, to whom I am in debt everything. To my husband, **Dr. Javad Sadri**, thank you for your unlimited motivation and patience, encouragement and love. To my two incredible sons, **Iman** and **Erfan**, thank you for your patience for all the times that I was busy with my studies and not being beside you. I cannot describe how

important you are to me and how much I love you. To **my parents**, **sisters** and **brothers**, who were patient all over the time that I was studying abroad.

Thank you!

CHAPTER 1
INTRODUCTION & LITERATURE REVIEW

1.1 Cardiovascular diseases and risk factors

Cardiovascular diseases are the leading cause of death in both sexes in developed countries[1]. According to data from the American Heart Association, in the United States in 2008, approximately one person has a heart attack every 26 seconds and among these subsequently one dies about every 60 seconds[2]. Particularly, at least a third of these deaths are premature occurring among people younger than 55 years of age. Most of this premature mortality occurs in a subgroup of the population that is prone to accelerated atherogenesis caused by genetic, lifestyle, environmental factors and the interaction between them.

Atherosclerosis or coronary artery disease (CAD), is characterized by the accumulation of lipids and fibrous elements in large and medium-sized coronary arteries. The initial lesions of atherosclerosis consist of subendothelial accumulations of cholesterol-loaded macrophages called foam cells. Eventually, smooth muscle cells (SMCs) move from the middle layer into the subendothelial layer and proliferate to form fibrous plaques. This accumulation of fat-laden cells, SMCs, and other material forms a patchy deposit called an atheroma or atherosclerotic plaque. With additional growth, the atheroma progressively narrows the arterial lumen or later ruptured, resulting in ischemia and heart attack[3].

Epidemiological studies have revealed numerous risk factors for CAD. These factors can be classified into two groups: the first group includes those factors with an important genetic component, such as high levels of plasma total cholesterol (T Chol), low-density lipoprotein cholesterol (LDL-C), triglyceride (TG), lipoprotein a (Lp(a)), C-reactive protein and homocysteine, low levels of plasma high-density lipoprotein cholesterol (HDL-C), high body mass index, high blood pressure, and type 2 diabetes; And the second group is environmental risk factors, such as smoking, obesity, a diet high in fat and calories, physical inactivity, inflammatory diseases, stress, and air pollution.

Although some risk factors-such as increasing age and male gender-cannot be changed, a number of them can be modified, despite having a genetic component. These include cholesterol levels and hypertension. In view of the fact that disorders of lipoprotein metabolism are associated with atherosclerosis, the next section will discuss the metabolism of lipoproteins in more detail.

1.2 Metabolism of plasma lipoproteins in humans

Cholesterol has several vital functions in the body. It plays an important role in the stability of the cell membrane but also provides a precursor for bile acids, adrenal steroid and sex hormones. Therefore, an adequate quantity of cholesterol is essential for survival. Cholesterol is structurally different from the other lipids in the body, triglycerides and phospholipids. The molecule consists of three main regions: a hydrocarbon tail consisting of a chain of eight carbons, a

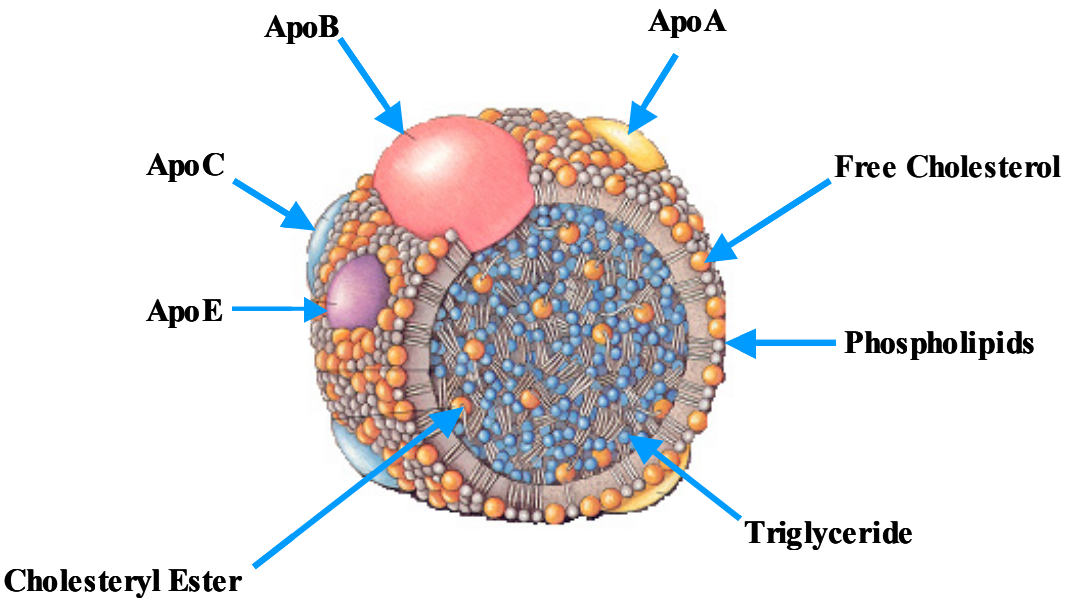
steroid-ring nucleus, and a hydroxyl group. Cholesterol is a lipid and, therefore, travels along with triglyceride in the blood through complexes called lipoproteins.

1.2.1 Lipoprotein structure and function. [4, 5]

Lipoproteins are a heterogeneous population of macromolecules containing a core of neutral lipids consisting of cholesteryl esters and triglycerides surrounded by a monolayer surface of polar lipids (unesterified cholesterol and phospholipids) and proteins called apolipoproteins (apoproteins) **(Figure 1)**. In addition, a number of different phospholipids cover the surface of lipoproteins, the more common of which are phosphatidylcholine, phosphatidylserine, phosphatidylethanolamine, and sphingomyelin.

There are four major classes of apolipoproteins, each with its own subclasses: apolipoprotein B (apoB), apolipoprotein A (apoA), apolipoprotein C (apoC), and apolipoprotein E (apoE). While the majority of apoproteins are synthesized in the liver, the intestine secretes apoB48 and a portion of apoAs and apoE.

Figure 1. Structure of lipoproteins.
Modified from the book: Cholesterol and Atherosclerosis Daignosis and treatment written by Scott M. Grundy[6].



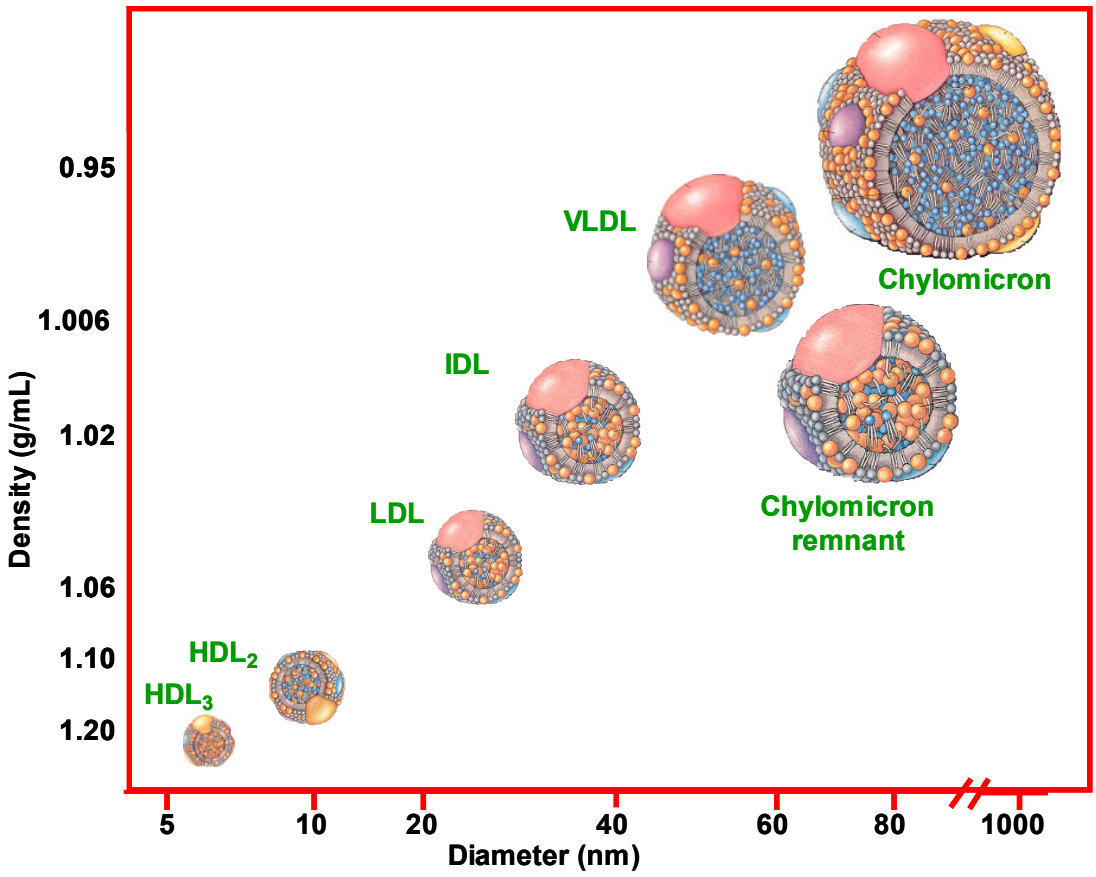
Apoproteins differ between lipoproteins and serve a variety of important functions in the body. Explicitly, they require the synthesis and secretion of specific lipoproteins, stablize the lipoprotein coat and therefore, whole particles, provide structural integrity for the lipoprotein, act as cofactors or inhibitors in the activation of specific enzymes, and interact with specific cell membrane receptors that are involved in the receptor-mediated removal of lipoproteins from circulation or contribute to lipoprotein metabolism.

1.2.2 Classification of lipoproteins

The main functions of the various lipoproteins are determined by their apolipoprotein and lipid components. In general, lipids have lower density than proteins; hence the lipid-to-protein ratio of a lipoprotein determines its class.

Those with a lower amount of lipid relative to protein will have a higher density than lipoproteins with a larger lipid-to-protein ratio. Traditionally, plasma lipoprotein classification depends on the density or mobility on agarose gel-electrophoresis based on their size. As shown in **Figure 2**, they are divided into five main categories: chylomicrons (CM), very low-density lipoproteins (VLDL), intermediate density lipoproteins (IDL), low-density lipoproteins (LDL), and high-density lipoproteins (HDL). The following sections will describe each class of lipoprotein in more detail.

Figure 2. Plasma lipoprotein classification.
Based on their diameter on x-axis and their density on Y-axis.



1.2.2.1 Chylomicron (CM)

CMs have a large core of mainly triglycerides and a small amount of cholesteryl esters. Their diameter ranges from 70-1200 nm with a density of less than 0.95g/ml **(Table 1)**. The role of CM is to transport dietary triglycerides to various tissues. They are synthesized within intestinal mucosal cells from assembly of triacylglycerols, derived from dietary fat absorption, and apoB48. ApoB48 results from the apoB100 gene, modified post-transcriptionally by the apo B editing Complex which demethylates a cysteine on the apo B mRNA to uracyl and thus leads to a stop codon, forming a mRNA and a protein that has

48% of the mature apo B100. Subsequently, CMs are secreted into the mesenteric lymph and enter systemic circulation. Other apoproteins are also secreted with CMs. As CMs enter the plasma, they acquire apoE and apoCs (CI, CII, and CIII). In the capillaries of adipose tissue or muscle cells in the peripheral circulation, lipoprotein lipase (LPL) hydrolyzes the triglycerides of CM and releases fatty acids and the A and C apolipoproteins back into systemic circulation. ApoCII (an activator) and apoCIII (an inhibitor) modulate LPL activity. The fatty acids released during lipolysis can be taken up either by muscle cells for energy utilization or adipose tissue for storage. Following LPL action, CMs contain apoB48, apoE, cholesterol ester, and less triglyceride, and are called remnant CMs. They return to the liver for degradation and reutilization of their core constituents. **(Figure 3A)**

1.2.2.2 Very low-density lipoprotein (VLDL), intermediate density lipoprotein (IDL)

VLDL is the major lipoprotein synthesized by the liver and is responsible for transporting endogenously synthesized triacylglycerols (also called triglycerides) from the liver to peripheral tissues. These triglyceride-rich lipoproteins are smaller than chylomicrons, having diameters ranging from 30-80 nm and densities between 0.95-1.006g/ml. ApoB100 is the main structural apolipoprotein in the VLDL particle but it also contains apoE and apoCs (CI, CII, and CIII) **(Table 1)**. ApoB100 is synthesized in the rough endoplasmic reticulum and is assembled with triglycerides in hepatocytes and secreted as nascent VLDL.

Triglyceride is the major lipid component of the nonpolar core of VLDL but cholesterol ester is also present. As these lipoproteins circulate, they transform into mature VLDL particles via the acquisition of cholesteryl esters and ApoCII, CIII, and E transferred from HDL. VLDL particles follow the same catabolic pathway through lipoprotein lipase as chylomicrons. During hydrolysis of phospholipids, most apoCs and some apoEs leave the surface coat of VLDL and are transferred to HDL. VLDLs also exchange triglycerides for cholesteryl esters from HDL, mediated by cholesteryl ester transfer protein (CETP). After hydrolysis of triglycerides, the VLDL remnants, called intermediate-density lipoprotein (IDL), have two fates: they can be taken up directly by a LDL-receptor via its apoE moiety or they can be further delipidated by hepatic lipase (HL) to form the LDL particle. **(Figure 3A)**

1.2.2.3 Low-density lipoprotein (LDL)

As mentioned previously, LDL is the product of the interaction of IDL with HL. The major cholesterol-carrying lipoprotein of plasma is LDL. In other animal species, such as ruminants and some rodents, most cholesterol is transported in HDL[5]. The lipid core of LDL is composed almost entirely of cholesteryl esters surrounded by a surface coat of apoB100 (its only apolipoprotein) and unesterified cholesterol and phospholipids **(Table 1)**. Normally, triglycerides constitute only 4 - 8 percent of the LDL mass. Circulating LDL is removed mainly by the liver and a smaller proportion by extrahepatic tissues. The removal of LDL occurs mainly by a receptor pathway[7]. Eighty

percent of LDL receptors are localized to specialized regions on the cell surface called "coated pits"[5, 8]. After LDL particles bind to the LDL receptor, cells internalize LDL and its receptor and make an endosome. Inside the endosome the receptors dissociate from the lipoprotein and are recycled back to the coated pits at the cell surface to be used again. Subsequently, LDL is delivered to lysosomes where catalytic enzymes including cholesteryl ester hydrolase and cathepsins release free cholesterol and degrade apoB100[5, 8-11]. The free cholesterol derived from esterified cholesterol can serve as a constituent for cell membranes, be re-esterified for storage, or be removed from the cell and excreted in the bile. The amount of cholesterol entering the cell regulates the following downstream events: 1) cholesterol synthesis, with the rate-limiting step goverened by hydroxymethylglutaryl coenzyme A (HMG-CoA) reductase, 2) cholesterol efflux from the plasma membrane, 3) intracellular cholesterol esterification via acyl-CoA:cholesteryl acyltransferase (ACAT), and 4) the synthesis of new LDL receptors[7, 12]. Therefore, the regulation of serum LDL depends on two factors: the quantity of VLDL produced by the liver and the rate of LDL clearance. Hence, the number of LDL receptors is a key regulator of serum LDL concentration by affecting both the rate of formation and the rate of clearance of LDL. **(Figure 3A)**

1.2.2.4 High-density lipoprotein (HDL)

High-density lipoproteins are the smallest lipoproteins, having a diameter of 5-12 nm with a density of 1.063-1.210 g/ml **(Table 1)**. HDLs are a heterogeneous population of lipoproteins whose main apoprotein is ApoAI. The

surface of HDL particle contains ApoAII, apoCs and apoE **(Table 1)**. HDL particles also carry other proteins, such as lecithin: cholesterol acyltransferase (LCAT)[13], serum amyloid A (SAA) protein[14], and beta-glycoprotein-I[15]. Recently, shotgun proteomics identified 48 proteins in HDL[16]. HDL can be sub-grouped by density or charge properties. In 1951, Lindgren *et al*. first identified two HDL subspecies based on their buoyancy with analytic ultracentrifugation[17]: HDL_2, which has a density range of 1.063–1.125 g/mL, and HDL_3, has a range of 1.125–1.210 g/mL. HDL_2 is about 50% larger than HDL_3 with 3-4 fold more cholesteryl esters and 2-fold more phospholipids than HDL_3[18]. Alternatively, HDL can be classified based on charge. Non-denaturing lipoprotein electrophoresis has been used for decades as a standard laboratory technique to analytically separate lipoproteins. In these systems, HDL migrates as an α-band. Separation of HDL by 2-dimensional electrophoresis (incorporating a size-based separation in addition to separation based on charge) has revealed additional heterogeneity, with pre–β-1, pre–β-2, and pre-α HDL in addition to the α-migrating species[19]. The next section will discuss this specific class of lipoproteins in more detail.

Table 1. Plasma Lipoprotein Components

	Origin	Density (gm/ml)	Size (nm)	Protein (%)	Major apo	Other apo
CM	Intestine	<0.95	100-1000	1-2	B48	AI, C's
VLDL	Liver	<1.006	40-50	10	B100	AI, C's
IDL	VLDL	1.006-1.019	25-30	18	B100	E
LDL	IDL	1.019-1.063	20-25	25	B100	
HDL	Liver, intestine	1.063-1.210	6-10	40-55	AI, AII	AIV, Cs, E, F

Apo:. apolipoproteins; HDL: high-density lipoprotein; IDL: intermediate-density lipoprotein; LDL: low-density lipoprotein; VLDL: very-low-density lipoprotein ; CM: chylomicrons.

1.3 HDL

1.3.1 Epidemiology of HDL

A large body of evidence derived from epidemiological and clinical studies has shown that a low serum level of HDL-C is a powerful predictor of cardiovascular disease. In 1977, the analysis of 2815 men and women aged 49–82 years from the Framingham Study showed that elevated LDL-C and reduced HDL-C levels are independent predictors for developing cardiovascular disease[20]. Likewise, the Prospective Cardiovascular Munster (PROCAM) study of 4559 male participants aged 40–64 years revealed a strong negative linear correlation between HDL-C levels and the incidence of CAD (CHD risk ratio of 4.0 for HDL-C <25 mg/dL versus HDL-C >65 mg/dL, $p < 0.001$)[21]. In addition, several

independent studies estimate that every 0.0259 mmol/L increase in serum HDL-C decreases the risk of cardiovascular disease by 2%[22-24]. Therefore, Canadian and American CAD prevention guidelines define HDL-C as a categorical risk factor and HDL-C is used in a multivariate model to predict cardiovascular risk and determine the need and intensity of preventive therapies[25-27].

Some evidence demonstrates that HDL subclasses may have different functions in terms of cardio protection (reviewed by Joy and Hegele[28]), and there is an inverse correlation between HDL_2 and atherosclerosis, hypertriglyceridaemia, and obesity[29-31]. Furthermore, patients with CAD have lower α-1 and pre-α-1 HDL, but higher α-3 HDL subfractions compared with control groups, suggesting that HDL_2 is more atheroprotective than HDL_3[28, 32]. These protective functions of HDL particles have been attributed mainly to their capacity to facilitate transporting excess cholesterol from peripheral cells to the liver for excretion in a process known as reverse cholesterol transport (RCT). The **RCT** pathway is shown in **Figure 3B**.

1.3.2 Metabolism of HDL

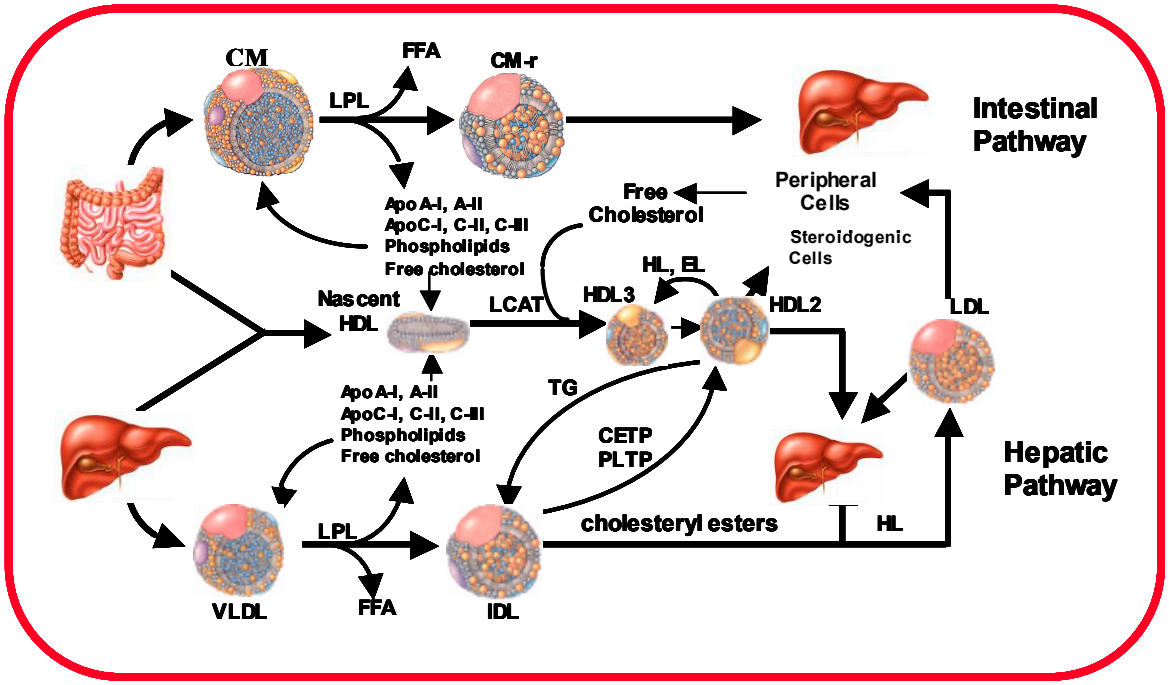
The molecular metabolism of HDL is complex since the components of HDL particles are synthesized from several sources and are also metabolized at different sites. The main source of HDL production is the liver. Approximately 80% of ApoAI, the key apolipoprotein of HDL, is secreted by the liver and 20% by the intestine[33, 34]. Phospholipids, derived from the cell membrane and hydrolysis of triglyceride-rich lipoproteins, are incorporated with lipid-free

ApoAI. Lipid-poor ApoAI can interact with the ATP-binding cassette (ABC) transporter A1 (ABCA1) in peripheral cells promoting efflux of free cholesterol and phospholipids, resulting in the formation of nascent HDL particles. These particles, containing ApoAI and phospholipids (and little cholesterol) form a structure resembling a flattened disk in which the phospholipids form a bi-layer surrounded by two molecules of ApoAI arranged in a circular fashion at the periphery of the disk. These nascent HDL particles mediate further cellular cholesterol efflux and are further modified by LCAT, an enzyme activated by ApoAI. LCAT transfers an acyl chain (a fatty acid) from the R2 position of a phospholipid to the 3'-OH residue of cholesterol, generating cholesteryl esters. The hydrophobic cholesteryl esters move to the core of the lipoprotein and the HDL particle subsequently assumes a spherical configuration (the HDL_3 particle). With further cholesterol esterification, the HDL particle increases in size to become the more buoyant HDL_{2a}. Mature HDL particles can serve as acceptors of cellular cholesterol for the ATP-binding cassette sub-family G member 1 (ABCG1) mediated cholesterol efflux[35, 36]. There is also cholesterol efflux to mature HDL by passive efflux[37] and possibly other transporters as well. In humans and some other species, cholesteryl esters in HDL can be transferred to triglyceride-rich apoB-containing lipoproteins by CETP in exchange for triglycerides. These triglyceride-enriched HDL are denoted HDL_{2b}. Triglycerides and phospholipids of HDL_{2b} are hydrolysed by HL and endothelial lipase (EL), respectively, converting them back to HDL_3 particles. Finally, HDL-C can be taken up by the liver and subsequently secreted into the bile in a process that may

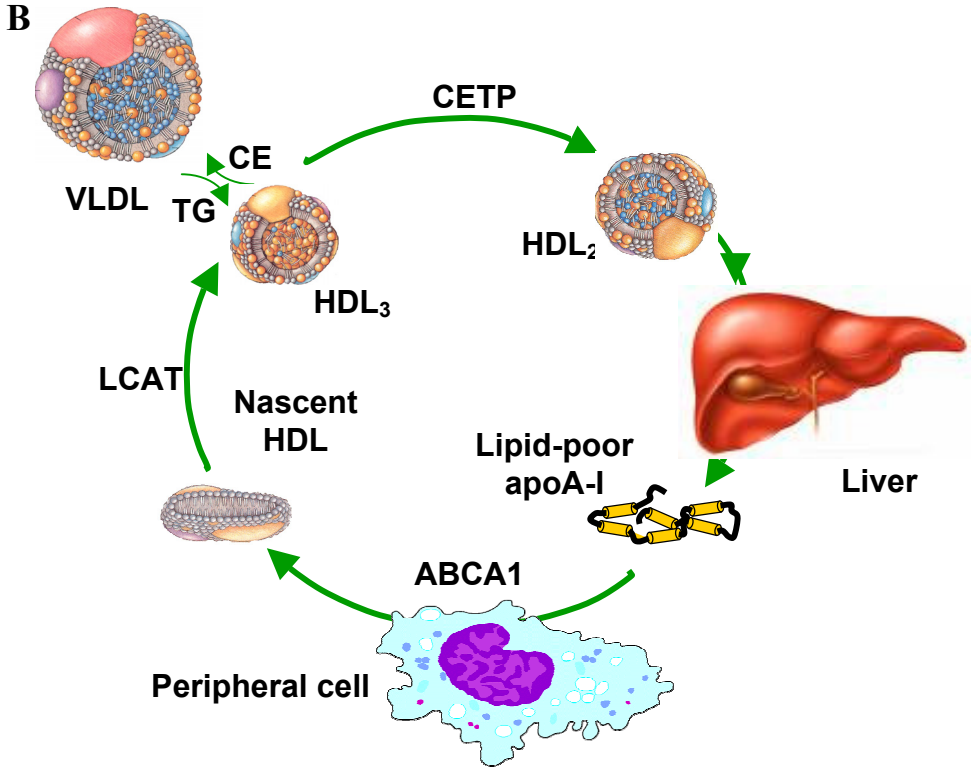
involve selective lipid uptake by scavenger receptor BI (SR-BI). The liver catabolizes ApoAI from mature HDL particle, while the kidneys catabolise lipid-poor ApoAI.

Figure 3. Metabolism of Lipoproteins (A) & Reverse cholesterol transport (RCT) (B)

A



B



1.3.3 Other properties of HDL

Besides RCT, HDL may afford protection from vascular disease by exerting additional effects that include antioxidant, anti-inflammatory, antiapoptotic, antithrombotic, and vasodilatory functions[38].

1.3.3.1 Antioxidant properties of HDL

Typically, the accumulation of LDL in the subendothelial matrix initiates the process of atherosclerosis. The trapped LDL in contact with the oxidative waste of vascular cells undergoes lipid oxidation. Such modifications generate minimally oxidized LDL species that have pro-inflammatory action. Enzymes associated with HDL particles play potential roles in minimizing LDL oxidation and promoting HDL antioxidant properties. These enzymes include paraoxonase-1 (PON1), paraoxoanse-3 (PON3), platelet activating factor acetylhydrolase (PAF-AH) and possibly glutathione phospholipid peroxidase[39-42]. PON1, an esterase carried on HDL, can inhibit oxidation of LDL and degrade certain biologically active oxidized phospholipids that play key roles in several aspects of atherogenesis. PAF-AH can hydrolyse and inactivate PAF, the potent lipid mediator involved in inflammatory diseases as well as in atherogenesis[43]. PAF-AH can also effectively hydrolyze oxidized phospholipids.

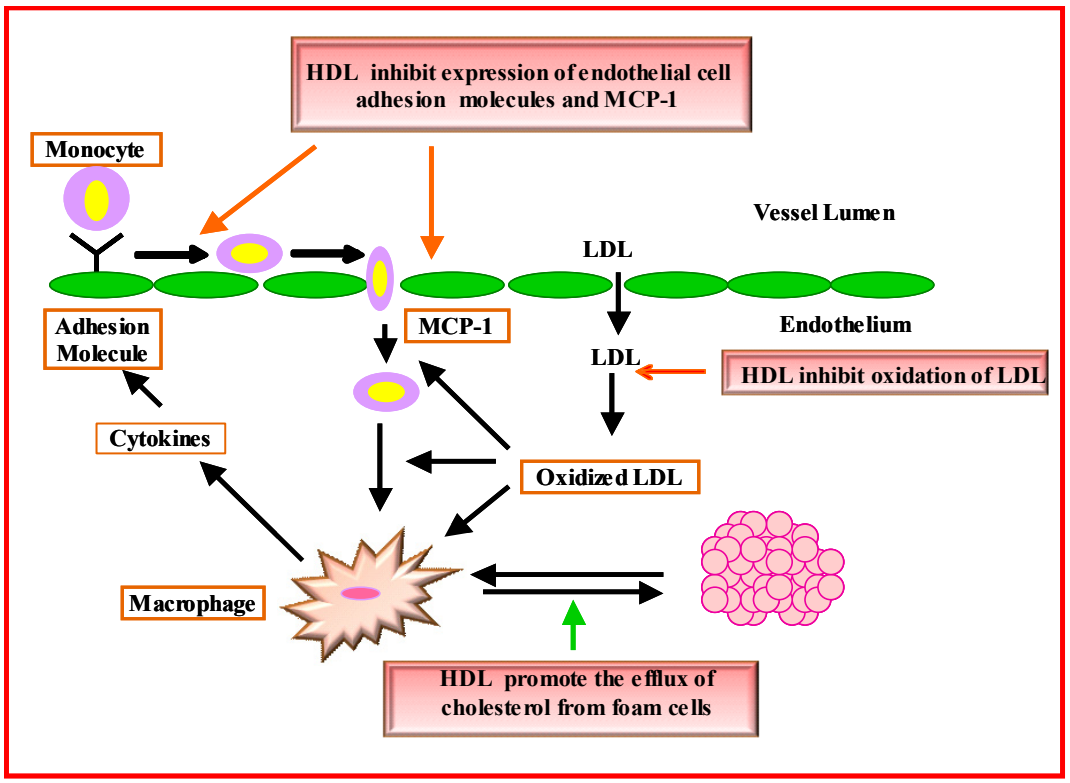
1.3.3.2 Anti-inflammatory properties of HDL

Inflammation is a major element in the process of atherosclerosis and is characterized by an accumulation of macrophages and T lymphocytes in the

arterial intima[44, 45]. In the early stages of this process, the accumulation of minimally oxidized LDL is a trigger for endothelial cells to express a number of adhesion proteins including vascular cell adhesion molecule-1 (VCAM-1), intercellular adhesion molecule-1 (ICAM-1), and E-selectin[46, 47] and subsequently cause the adhesion of monocytes to endothelial cells[48-50].

Once monocytes bind to adhesion proteins on the surface of endothelial cells, they are recruited into the subendothelial space by chemokines such as monocyte chemotactic protein-1 (MCP-1) **(Figure 4)**. The anti-inflammatory activity of HDL is manifested by reducing cytokine-mediated expression of adhesion molecules, diminishing neutrophil infiltration within the arterial wall, and reducing the generation of reactive oxygen species[51-53]. *In vitro* studies have also shown that HDL inhibits monocyte transmigration in response to oxidized LDL[54]. Accordingly, antioxidant and anti-inflammatory properties of HDL contribute to the antiatherogenic potential of these lipoproteins.

Figure 4. Anti-inflammatory Properties of HDL



1.3.3.3 Enhancing endothelial function

HDL particle attenuates endothelial cell apoptosis and enhances endothelial cell proliferation and migration in order to promote endothelial monolayer integrity. For antiapoptotic actions, HDL prevents the sustained rise in intracellular calcium induced by proapoptotic agents such as OxLDL. Additionally, HDL inhibits endothelial apoptosis induced by TNF-α in a dose-dependent manner by inhibiting caspase 3 activity[55]. HDL also promotes endothelial cell proliferation in a calcium-dependent manner by multiple kinase cascades involving PI3-kinase (PI3K), p38 and p42/44 MAP kinase (MAPK), and Rho kinase and by the small GTPase Rac[56, 57].

Nitric oxide (NO) is an athero-oprotective signalling molecule that can be modified by HDL. Endothelial NO synthase (eNOS) is located in cholesterol-enriched plasma membrane caveolae. Oxidized LDL, as a cholesterol acceptor, can disrupt caveolae and eNOS function. HDL regulates the lipid environment in caveolae by providing cholesterol esters that subsequently regulate eNOS subcellular distribution. HDL causes membrane-initiated signalling, which stimulates eNOS activity and production[58, 59] It also prevents the effect of LDL on eNOS that favours the production of O_2 over NO.

1.3.3.4 Antithrombotic properties of HDL

HDL particle has direct actions on the vascular endothelium and on mechanisms mediating thrombosis. HDL has multiple antithrombotic actions such as the promotion of blood flow, reduction of thrombin production, and the alteration of endothelial and platelet activation. HDL increases blood flow by increasing NO and prostacyclin production[60]. Besides up-regulation of NO and prostacyclin synthesis, HDL reduces platelet activating factor release and thromboxane A2 (TxA2) synthesis that contributes to the antithrombotic effect of HDL[61, 62]. HDL also enhances the activity of protein C and protein S, which subsequently inactivate the coagulation factor Va[63].

Summing up, multidimensional cardioprotective properties of HDL and the unambiguous inverse correlation of HDL-C concentration with CAD have inspired many studies to identify factors including genetic and environmental factors that modulate HDL in man.

1.4 Tools to study genetic varability of HDL-C

The majority of common diseases and most quantitative traits (e.g. blood pressure, serum cholesterol or glucose levels) that can be measured in human populations are complex traits, with multiple genetic and environmental factors interacting to determine the ultimate phenotype. Genetic complexity arises from genetic heterogeneity, epistasis (gene-gene interaction), and pleiotropy (multiple phenotype expression by one gene), with contributions from environmental factors that are often unknown or immeasurable. In addition, the true genetic models that underlie common phenotypes are not completely known and may be different for each variant.

A systematic approach to search for genetic determinants of a disease includes the following steps: first, determine genetic heritability by performing families, twins, or adoption studies; second, illustrate the type and frequency of susceptibility alleles and the mode of inheritance by using segregation analysis; third, perform linkage analysis to identify susceptible loci; fourth, narrow down the candidate region and prioritize the genes at the locus, and fifth, ultimately identify variation in candidate gene(s) and define the biochemical action and source of variation in the trait.

1.4.1 Familial aggregation and heritability

The initial step in searching for susceptible genes for any disease is to show that genetics plays a role in determining susceptibility. There are several approaches to estimate the genetic component of a trait or disease: many diseases

cluster in families in a pattern suggesting a strong genetic component to disease susceptibility. The relative risk factor (lambda) is one measure of the influence of genetics. It compares the frequency of the disease in close relatives of affected individuals with its prevalence in the general population. However, familial aggregation does not separate genetic and non-genetic factors. Heritability is used to estimate the extent to which familial aggregation is caused by the contribution of inherited factors; it quantifies the portion of the population variation that can be explained by genetic factors. Heritability can be approximated either from family, twin, or adoption studies.

1.4.2 Segregation studies

Segregation analysis is a general method to evaluate the transmission of a trait within pedigrees. It is used to assess the mode of inheritance and can also provide evidence for a major gene effect in a population. Due to the participation of multiple genes as well as environmental factors in complex diseases, it is often difficult to determine the true models that give rise to diseases with a complex mode of inheritance.

1.4.3 Approaches for identifying evidence for susceptible genes

In the past two and a half decades, techniques for the identification of genes associated with biological traits have evolved dramatically as the complexity of the genome became partially unravelled and as the tools for

genomic research increased the analytical throughput by several orders of magnitude.

Several approaches have been used to find disease susceptibility genes: the candidate gene approach, genome-wide searches by either linkage or association, and gene expression array analysis.

1.4.3.1 Candidate gene approach

With the identification of metabolic pathways involved in a disease or trait, the genes involved in these processes can be cloned. These genes can then be targeted for sequencing and association studies. Some rare mutations in candidate genes have been identified through sequencing that are responsible for the extreme level of quantitative traits. Many association studies of candidate genes have shown that both rare and common variants are implicated in determining the level of these traits in the general population[64]. However, for most complex traits there are many candidate genes whose functional effects are unknown; the candidate gene approach is often used to follow-up genome-wide efforts to discover regions of the genome with evidence of association with the trait under study.

1.4.3.2 Genome-wide linkage studies

As the first step in dissecting the genetics of complex phenotypes, focusing on familial forms of the disease can facilitate the detection of the genes involved in a complex or quantitative trait. This approach can take the form of

linkage analysis that enables specific genetic intervals to be linked with disease susceptibility.

Roughly speaking, there are three possible ascertainment strategies for analyzing quantitative traits. First, under the affected ascertainment scheme, a proband is chosen based on a threshold value for a measured quantitative trait. The second strategy is the selection of sib pairs exhibiting extreme discordance: families are chosen where one sib is in the lower tail of the distribution of the quantitative trait and another sib is in the upper end of the distribution of the same trait. The last scheme is a random selection of families from a population regardless of the quantitative trait.

Having collected families and genotyped them for genetic markers (usually microsatellite markers), linkage analysis can locate a locus that shows segregation between the disease and linked markers inside the families. The degree of linkage is measured by the logarithm of odd ratio (LOD) score. This score was first proposed by Morton[65] and it statistically measures the linkage between trait or disease and genetic markers.

Positive LOD scores give evidence in favour of linkage and a LOD score of 3.3 has been proposed as the threshold of significant linkage[66].

Several approaches to linkage analysis have been developed for complex and quantitative traits. Some investigators use parametric or model-based linkage analysis. In this approach the model of disease (disease allele frequency, mode of inheritance and marker allele frequency) must be pre-specified. However, for complex traits this approach may reduce the power to detect genes, especially if

the model is misspecified. In contrast, model-free (non-parametric) linkage approaches have been developed that are based on the idea that in the regions of susceptibility genes the affected relatives have share alleles that are identical by descent (IBD), more than expected by chance. Recently, genome–wide scanning techniques have been applied to the identification of the genes of quantitative traits. It assumes that if a marker is linked to a gene influencing the trait, two siblings who have more similar trait values share IBD more than expected. Therefore, quantitative trait locus (QTL) analysis has also been used to identify chromosomal regions that are associated with variations in the trait such as HDL-C levels.

1.4.3.3 Association studies

In the past few years, there has been a remarkable increase in genomic discoveries for complex and quantitative traits using genome-wide association studies (GWAS). A GWAS is a method to identify genes involved in human disease. In this approach, thousands of samples are genotyped with single nucleotide polymorphisms (SNPs). The frequencies of single SNP alleles, genotypes, or haplotypes between case and control groups (or in families) are examined to identify loci with statistically significant differences.

This method has several advantages compared to other gene discovery approaches. In comparison to candidate gene studies, which select genes for study based on known or assumed metabolic pathways, GWAS allow a comprehensive scan of the genome in an unbiased way and therefore have the potential to

identify entirely novel susceptibility genes. In contrast to family linkage-based methods, association studies are able to uncover the meiotic recombination events in a population, rather than only those in the families studied. Therefore, association signals are confined to smaller regions containing only one to a few genes. In addition, GWAS can detect SNPs with only modest effect on disease. However, the power of association studies is a function of several factors, such as the frequency of the risk allele, sample size, disease frequency, the linkage disequilibrium between the genotyped marker and the true causative allele, and genetic homogeneity of the sample population.

Considerable successes have been reported using this new strategy. For example, genetic variations that influence the risk of type 2 diabetes, coronary heart disease, obesity, Crohn's disease, and many other diseases have been reported[67-69].

1.5 Genetics of HDL-C

1.5.1 HDL & heritability

To examine the genetic contribution to the determination of HDL-C levels, there have been at least nine published studies in twins[70-78] and fourteen family studies[77, 79-90]. Estimates for the heritability of plasma HDL cholesterol levels vary between 0.24 to 0.83[84, 91, 92].

1.5.2 Segregation of HDL-C

Different modes of inheritance have been observed in different segregation studies, likely due to the complexity of HDL-C metabolism and population-specific differences. A recessive mode of inheritance has been identified for low HDL-C in a genetic study of hypoalphalipoproteinemia[93]. Another study identified the mixed model, including major gene and multifactorial transmissible components, in a sample of nuclear families from Israel[94]. Some evidence for a recessive major gene for high HDL-C has also been reported in NHLBI Familiy Heart Study and in Utah pedigrees[95]. Considering HDL-C as a continuously distributed trait in genetic analysis of the Lipid Research Clinic's collaborative study, which consists of first degree relatives of Stanford University employees, is more consistent with some degree of multifactorial transmission or major gene model with additive allelic effects and similar allele frequencies[96]. In addition, some studies showed evidence of a major gene for HDL-C but were not able to discriminate between a dominant and an additive model[97, 98]. However, there have also been reports of no segregation of HDL-C, leading some investigators to suggest that considering other factors such as TG to improve the likelihood of their genetic models significantly[99]. Thus, by including other parameters in data from 526 Mexican Americans, Mahaney and colleague found evidence for a major locus with a co-dominant mixture model[91]. These inconsistent results are likely due to the etiological heterogeneity of HDL-C. Differences in the specific protocols for ascertaining probands and differences in methodology could also lead to discrepant results.

1.5.3 Candidate genes regulating HDL metabolism

Genes encoding proteins associated with HDL particles or involved in its metabolism have been studied in great detail. Wang and Paigent reported 54 genes involved in HDL metabolism and function[100]. Of these, *ABCA1*, *ApoAI*, *CETP*, *LCAT*, *LPL*, *HL*, *EL*, and *ABCG5/8* have been shown to regulate HDL in humans. Thus, they will be explained in more detail in the following section.

1.5.3.1 ABCA1

The *ABCA1* gene encodes a 240-kDa protein, a membrane protein with 2 transmembrane domains that facilitates phospholipid and cholesterol transport. ABCA1 belongs to a superfamily of ATP-binding cassette transporters, which contains 50 exons and spans 149 kb on chromosome 9 (9q31.1). The main sites of ABCA1 expression are the liver, macrophages, and steroidogenic tissues. Mutations in *ABCA1* result in Tangier disease[101-103] and familial HDL deficiency[104] in the homozygous and heterozygous form, respectively. At least three studies have examined the prevalence of *ABCA1* mutations in patients with low HDL-C. Cohen *et al.*[64] examined patients from the United States and Canada and identified an *ABCA1* mutation or gene variant in 10-16% of patients with low HDL-C. A similar finding was reported in Germany with approximately 10% of low HDL-C being associated with *ABCA1* mutations[105]. Alrassadi K. *et al*. reported a 20% prevalence of *ABCA1* mutations in subjects of French-Canadian descent[106], possibly reflecting a founder effect. Several association studies have

been performed on this candidate gene with HDL-C levels. A distinct example was recently shown by the Copenhagen City Heart Study [107],which demonstrated in more than 9000 participants a pronounced effect for four rare mutations in *ABCA1* on HDL-C levels (a reduction of 17 mg/dL for heterozygotes vs noncarriers, $p < 0.001$). However, 19 SNPs of *ABCA1* investigated in other association studies with a large number of individuals did not demonstrate significant effects on HDL-C levels (see review [108]). Nevertheless, the occurrence of many rare variants in *ABCA1* each with a strong effect, could sum to an important influence on HDL metabolism. These rare mutations would not be identified by association studies using common variants because of the small number of mutation carriers even in large studies.

1.5.3.2 ApoAI

The *ApoAI* gene contains four exons, covers 2.21 kb, and is closely linked with two other apolipoprotein genes (*ApoCIII* and *ApoAIV*) on chromosome 11q23.1. *ApoAI* constitutes the major apolipoprotein in HDL, accounting for approximately 70% of the protein mass within HDL. ApoAI plays the key role in the formation and metabolism of HDL particles. It is synthesized in the intestine and liver as a 267-amino acid (aa) preproapolipoprotein and is secreted as a 249-aa proprotein that is rapidly cleaved by a plasma protease to generate the mature single polypeptide with 243-aa residues and a molecular weight of 28 kDa[109]. This protein circulates in plasma with a concentration of 1.0-1.5 mg/mL. Recent characterization of the ApoAI crystal structure[110] and molecular modeling of the

lipidated ApoAI peptide[111] have shed considerable light on the structure-function of ApoAI domains critical for nascent HDL particle assembly. ApoAI promotes cholesterol efflux from tissues to the liver for excretion through binding to ABCA1 and is the most potent in vivo activator of LCAT[112, 113]. ApoAI is composed of 10 amphiphilic α- helices, with 11-aa and 22-aa repeats, which begin at residue 44 and are thought to have a crucial role in lipid binding[114, 115]. Central amphipatic α-helical regions are critical in LCAT activation[116]. The N-terminal (1-43) is important for stabilization of lipid-free ApoAI in solution[117, 118], while the C-terminal domain has been shown to play a specific role in ABCA1-mediated efflux[119]. Mutations within the *ApoAI* gene can lead to altered ApoAI and HDL-C levels, but not all *ApoAI* mutations lead to decreased ApoAI concentration[120]. Other mutations lead to amyloidosis, a disease that has multi-systemic effects but is not associated with atherosclerosis. So far, 60 mutations have been identified in this gene, including four nonsense mutations, Gln-2 X, W8X, Q32X, Q84X[121-124]; five frameshift mutations[125-129] and 51 missense mutations. Sorci et al have studied at least 47 of the mutations within the *ApoAI* gene. About 30 of these mutations are associated with low HDL-C and some but not all of them are associated with CAD[120]. In this study, most of the 18 reported mutations that affect ApoAI and HDL-C levels appear to cluster between residues ~100 and 200, i.e. within amphipathic helices 5, 6 and 7, which are critical for HDL assembly. Mutations clustering in the amino terminus are associated with amyloidosis. Other mutations, scattered throughout the ApoAI gene, are not associated with low HDL-C or amyloidosis; their physiological significance is

unknown. Paradoxically, two mutations, ApoAI $_{Milano}$ (apo A-I_{R173C}) and ApoAI $_{Paris}$ (apo A-I_{R151C}), are associated with low HDL-C with no increase in the incidence of heart disease[130]. Most cases of ApoAI mutations have been described in single families and account for a minority of HDL deficiency cases seen in humans[120]. Only one mutation at promoter -75G>A is associated with high levels of HDL-C[131]. Despite numerous identified rare mutations in the *ApoAI* gene that related to low HDL-C, only three common SNPs (rs670, rs5069 and rs5070) have been investigated and showed for an association with HDL-C levels (see review[108]).

1.5.3.3 LCAT

The *LCAT* gene contains 6 exons and spans approximately 4.2 kb on chromosome 16q22.1. It is mainly synthesized in the liver and circulates in the plasma with a concentration of 6mg/L, but is also synthesized by the brain. The LCAT glycoprotein consists of 416 amino acids, and displays two activities: a phospholipase A_2 activity which hydrolyses the fatty acid from phosphatidylcholine, and an acyltransferase activity which transfers the fatty acid to free cholesterol and forms cholesterol ester.

Mutations in this gene cause two forms of disease: complete LCAT deficiency, also called Norum disease, which causes a typical triad of diffuse corneal opacities, target cell hemolytic anemia, and proteinuria with renal failure. The partial LCAT deficiency causes Fish-eye disease, also known as dyslipoproteinemic corneal dystrophy, so called because of the pathogomonic

corneal infiltrations seen in these patients due to accumulation of cholesterol deposits in the cornea[132-134].

At least 30 mutations have been identified[135], which may have a dominant effect on HDL-C[136, 137]. While patients with LCAT deficiency have a marked decrease in HDL-C, accumulating evidence suggests that most familial LCAT deficiency are strongly associated with premature CAD[138, 139].

Although *LCAT* is a well-known gene involved in HDL metabolism, only a few large studies examined the possible role of this gene on HDL-C levels. Five common polymorphisms of *LCAT* have been examined for an association with HDL-C levels; however, the findings are not completely consistent (see review[108]). At least two SNPs (rs2292318, rs5922) have been shown to be associated with increases in HDL-C[140, 141]. Moreover, the influence of *LCAT* on HDL-C levels is supported by genome-wide association studies (see below).

1.5.3.4 CETP

The *CETP* gene contains 15 exons and spans approximately 22 kb on chromosome 16q13. CETP has a central role in the metabolism of HDL mediating the equimolar transfer of cholesteryl esters from HDL particles to triglyceride-rich (VLDL and IDL) particles in exchange for triglycerides; the effect of this transfer is an increase in cholesterol of apoB-containing particles and a decrease of HDL-C (see review[142]). Numerous polymorphisms have been detected in the promoter, introns, and exons of the *CETP* gene. Mutations in the *CETP* gene with loss of function are associated with increases in HDL-C.

Deficiency of CETP increases the cholesteryl ester content of HDL and the formation of large, buoyant HDL_2 particles. It has recently been found that these particles maintain their functional abilities, in terms of promoting cellular cholesterol efflux in an apoE and ABCG1-dependent pathway[143]. However, two SNPs, rs1800777 (R451Q), rs5880 (A373P), have been associated with a decreased level of HDL-C[108, 144].

Boes *et al*. have summarized several studies that investigated the association of *CETP* polymorphisms with HDL-C concentrations. Common polymorphisms mostly located in the promoter region such as rs1800775 (–629C>A), rs4783961 (–971G>A), and rs708272 (*Taq*1B) in addition to rs5882 (Ile405Val) have been clearly associated with elevated HDL-C levels[141, 142, 145-148]. A meta-analysis on *CETP* with 39,581–68,134 participants and 6 SNPs underlines the significant influence of *CETP* gene variation on HDLC levels[149].

1.5.3.5 LPL

The *LPL* gene contains 10 exons spanning about 30 kb, and is located on chromosome 8p21.3. The encoded protein, consisting of 475 aa, hydrolyzes triglycerides within chylomicrons and VLDL through activation by apoCII. LPL is therefore indirectly involved in HDL metabolism[150]. Deficiency of LPL or its activator apoCII causes severe hypertriglyceridemia and hyperchylomicronemia with a secondary HDL-C deficiency [151]. The inability to hydrolyze triglycerides within chylomicrons and VLDL leads to the accumulation of these particles in plasma. Nearly 100 structural variations in the *LPL* gene have been identified.

There are 61 missense mutations, most of which are located on exons 5 and 6; 12 nonsense mutations; 10 frameshift mutations or small insertions/ deletions; 3 gross mutations; 8 splicing mutations; and 4 promotor variants[152]. Several variants in the *LPL* gene are associated with either decreased [–93T>G, –53G>C, rs1801177 (Asp9Asn), Gly188Glu (no rs#), and rs268 (Asn291Ser)] or increased [rs328 (Ser447Ter)] LPL expression or activity (reviewed in [152]). A meta-analysis of four common polymorphisms confirmed considerable associations with HDL-C: the minor allele of rs328 was associated with higher levels of HDL-C, whereas the minor alleles of rs268, Gly188Glu and rs1801177 were associated with decreased HDL-C levels[153]. In addition, in a multi-ethnic cohort (the SHARE study) two SNPs, rs331 and rs328, associated with higher HDL-C levels in all ethnicities [154].

1.5.3.6 HL (LIPC, HTGL)

The *HL* gene contains 9 exons and spans approximately 30 kb on chromosome 15q22.1. The encoded protein consisting of 449 aa, is secreted from the liver and bound to hepatocytes, where it hydrolyzes triglycerides and phospholipids from lipoprotein. It is also involved in the selective uptake of cholesterol ester from HDL[155]. Two SNPs, rs28933094 (Ter383Met) and Ser267Phe, can contribute to HL deficiency and increased HDL-C concentrations[156]. Seven common SNPs of the *HL* gene have been investigated for association with HDL-C concentration. Several studies consistently showed an association of two of these SNPs, rs1800588 (-514C>T) and rs2070895 (-

250G>A), with increased plasma HDL-C concentration [108]. Isaacs A. *et al*., in a meta-analysis including more than 24,000 individuals, verified that one and two copies of the minor allele of rs1800588 increases HDL-C levels by 1.5 and 3.5 mg/dL, respectively[157].

1.5.3.7 EL (LIPG)

The *EL* gene contains 11 exons and spans about 71.4 kb on chromosome 18q21.1. The encoded 354-aa protein has substantially more phospholipase than triglyceride lipase activity[158]. EL, by enhancing the turnover of HDL components and promoting the catabolism of ApoAI, plays an important role in HDL metabolism. Several independent studies revealed five SNPs of *LIPG* that were significantly associated with HDL-C levels. The minor allele of three of those SNPs, rs2276269 (C+42T/ln5), rs6507931 (T+2864C/ln8), and rs3744841 (2237G > A), were associated with decreased HDL-C levels[159, 160]. Other studies revealed the minor allele of two other SNPs, rs3813082 (−384A > C) and rs2000813 (584 C/T), associated with lower HDL-C levels.[160-162]

1.5.3.8 ABCG5/8

The *ABCG5* and *ABCG8* genes, which are in a head-to-head orientation on chromosome 2p21, consist of 9 and 13 exons, respectively. Both genes are expressed in the liver and intestine [163]. Mutations in these ABC monomers were observed in patients with sitosterolemia, and may have contributed to the accumulation of sterol and subsequently atherosclerosis[164]. While ABCG5 is

known to mediate cellular cholesterol efflux to high-density lipoproteins and plays an important role in HDL metabolism[36], no studies have so far revealed any association with HDL-C levels.

1.5.3.9 PLTP

The *PLTP* gene contains 14 exons and spans approximately 14 kb on chromosome 20q13.12. The encoded 446-aa protein transfers phospholipids and cholesterol from TG-rich lipoproteins to HDL during lipolysis, remodelling the size and composition of HDL in the process[165, 166]. The overexpression of PLTP in transgenic mice is associated with an increase in the influx of phospholipids and cholesterol into HDL, resulting in increased pre–β-HDL and decreased HDL levels[167, 168] A study of 3,306 participants in the Dallas Heart Studyrevealed that a common sequence variant of the *PLTP* (rs6065904*)* gene acts in concert with other SNPs in the *CETP* (rs183130, rs5880) and *LPL* (rs2197089) genes to affect plasma levels of HDL-C[169]. Additionally, linear regression models predict that possession of the rs2294213 minor allele increases HDL-C independent of triglycerides [170]. Moreover, the influence of *PLTP* gene on HDL-C levels is supported by genome-wide association studies, in which the minor allele of rs7679 was associated with lower levels of HDL-C[171].

1.5.3.10 SMPD1

Niemann-Pick disease type A and B is caused by a deficiency of the enzyme acid sphingomyelinase, which is coded for by the *SMPD1* gene. This

gene is located on chromosome 11p15.4, is comprised of 6 exons, and spans about 5 Kb. The encoded protein consists of 631 aa and exists in at least three protein isoforms. *SMPD1* gene defects are reported to be associated with a severe reduction in plasma HDL-C [172, 172]. Our group has previously reported that compound heterozygosity at the *SMPD1* gene is associated with decreased activity of acid sphingomyelinase and low HDL-C in French Canadian families [173]. Furthermore, a decreased activity of lysosomal and secreted acid sphingomyelinase is believed to cause low HDL-C in part by decreased activation of LCAT and impaired formation of cholesteryl ester–enriched HDL particles [174]. Although rare mutations in the *SMPD1* gene can impair the function of acid sphingomyelinase and result in both Niemann-Pick disease type A or B and low HDL-C, it is not known whether common amino acid changing variants in *SMPD1* can modulate HDL-C levels within a population.

1.5.3.11 Other candidate genes

Association studies of several other genes involved in HDL metabolism such as *SR-B1*, *Apo CIII*, *Apo AIV*, *PON1*, and *ApoE* have shown evidence of association with HDL-C levels (see review[108]).

1.5.4 Whole genome linkage studies of HDL-C

Numerous whole-genome linkage scans have been performed for various human diseases and traits. This data was mined for associations with many parameters, including HDL-C levels. Parametric linkage studies resulted in the

identification of several chromosomal regions that might harbor genes that influence HDL-C levels. Pajukanta *et al*. performed linkage analysis on the combined data of Finnish and Dutch familial combined hyperlipidemia (FCHL) families. In addition to the familial combined hyperlipidemia trait (FCHL is frequently associated with a low HDL-C), they analyzed four component traits of the disorder TG, TChol, ApoB, and HDL-C. Subjects were coded as affected or unaffected based on the 10^{th} age and sex specific percentiles for HDL-C. The two-point maximum LOD score of the parametric linkage analysis using a recessive mode of inheritance detected evidence of linkage to 16q24.1, with LOD scores of 3.6 (P= 0.00002) for the low HDL-C trait. In their study, the parametric multipoint analysis also revealed a location score of 3.2 (P= 0.0001) when allowing for heterogeneity (α value = 0.75) in the same region on chromosome 16[175]. Based on this approach, Soro *et al*., using the same HDL-C percentile as affection status and again a recessive mode of inheritance, found strongest statistical evidence of linkage with a two-point LOD score of 4.7 and 1.6 between the low HDL-C and loci on chromosomes 8q23 and 6q24.1-24.2, respectively in the Finnish population[176].

In a FCHL genome scan from a Dutch population, the nonparametric multipoint LOD score for the same region on chromosome 16 was 1.4, the peak occurring within the same interval close to marker D16S402 for FCHL[177].

Numerous quantitative trait loci (QTLs) for plasma HDL-C also have been reported and these have been reviewed by Wang and Paigen (2005). They reported that 30 QTLs with LOD scores more than 1.4 for HDL-C have been

identified in humans (Table 3 from their review) [100]. Most of these QTLs (> 90%) are concordant with mice QTLs in syntenic regions. The high degree of concordance between mouse and human QTLs suggests a strong likelihood that the underlying genes will be the same. These authors therefore argue that an examination of HDL-C QTLs in mice will be a more efficient way to identify potential chromosomal regions of interest in the human genome.

The identified loci linked to HDL-C are scattered on different chromosomes and this could be explained by different underlying segregating variants conferring susceptibility to low levels of plasma HDL-C between the families and indeed between populations. To date, the only reproducible region reported in different independent studies is located on chromosome 16. A total of 10 genome scans have implicated this region of chromosome 16 with a LOD score >1; these are summarized in **Table 2**.

Table 2. Published genome-wide linkage analysis for HDL-C that implicates chromosome 16

Genetic Distance (cM)	Cytogenic Location	Nearest Marker	LOD Score	Population	Authors
45.2	16p12.1	D16S403-D16S769	1.25	Hong Kong Chinese Study	Ng MC. et al.[178]
62.11	16p12.1	D16S313	2.06	Bogalusa Heart Study	Bu X. et al.[179]
75	16q21	D16S3057	1.1*	Framingham Heart Study	Shearman AM. et al.[180]
82	16q22.3	D16S514	1.7	Bogalusa Heart Study	Amos CI.[97]
92	16q22.3-23.1	D16S2624- D16S518	4.33	Mexican American Study	Mahaney,M.C et al.[181]
99	16q22.1-24.3	D16S3096	3.4	Dutch & Finnish Study	Pajukanta P. et al.[175]
106	16q23.3	D16S402	1.4	Dutch families Study	Aouizerat B.E et al.[177]
111.2	16q24.1-24.2	D16S3091	2.2	Finnish families Study	Soro A et al.[176]
113.52	16q23.3	D16S402	IR†=6	FCHL Caucasian(Canada)	Gagnon F, et al.[182]
113.52	16q23.3	AFM031XA5(D16S402)	1.95	Framingham study	Yip Agustin G, et al.[183]

*LOD score for log (TG/HDL-C), †Intensity rate

The initial stage of a whole genome scan typically identifies a large (10-20 cM) genomic region containing hundreds of genes. The first step in identifying the causal gene is to narrow the linked region. This step has proven to be more arduous than originally anticipated. The results of the HapMap project[184] should make the laborious work of gene identification easier by enabling the analysis of a more efficient set of single nucleotide polymorphism (SNPs) and haplotypes in humans and allowing for a more fine-scale comparison of mouse and human homologous QTL's.

1.5.5 Whole-genome association studies and HDL-C

Several genome-wide scans and meta-analyses of lipids have been conducted in independent population samples. Common variants at 30 loci were reported to be associated with the concentration of various lipoproteins[171]. Most of these loci were previously implicated in lipid metabolism; however, at least eight new loci near *MVK-MMAB, GALNT2*[185, 186], *CTCF-PRMT8, MADD-FOLH1*[187], *ANGPTL4, FADS1-FADS2-FADS3, HNF4A*, and *TTC39B*[171] have been found to influence HDL-C levels. However, the associated SNPs at the new loci are non-coding and explain only a minor fraction of the variance in HDL-C levels. Therefore, the causal variants, or even the causal genes, are not yet clear. Of particular note, in these studies the previously-known loci *ABCA1, ApoA1, LCAT* and *CETP*, which cause mendelian syndromes or harbor multiple rare alleles that contribute to HDL-C variation[64], also contained common SNPs that were asssociated with HDL-C. Therefore, sequencing of all newly identified loci,

while considering all genes in the linkage disequilibrium pattern as high-priority targets, will be necessary to discover all common and rare variants and verify the full impact of each locus.

CHAPTER 2
STUDY DESCRIPTION

2.1 Main ineterst or problem

A reduced level of HDL-C is known to be directly associated with an increased risk of coronary artery disease. HDL-C is involved in the removal of excess cholesterol from cells to the liver for metabolism and excretion. It has been clearly demonstrated that variation at several genes has a significant effect over the spectrum of HDL-C levels in the general population. The identification of low HDL-C that is the result of the variation in a single gene may lead to a better understanding of the complex pathways of HDL metabolism and pave the way for novel therapeutic approaches. We postulate that novel genes related to HDL metabolism may contain variants that underlie low HDL-C in a proportion of the Quebec population. In addition, some genes may contain more common variants that contribute to the overall variability of HDL-C.

2.2 Proband and family collection

Patients with a severe decrease in HDL-C (less than 5% of age and sex adjusted percentiles based on the Lipid Research Clinics Population Studies Data Book, as previously described[188]) have been targeted, and members of their family have been recruited for the past 15 years in Dr. Jacques Genest's clinic. In total, 40 extended families with over 900 subjects have been recruited. Family members were sampled after a 12 h fast and the discontinuation of lipid-

modifying medications for at least 4 weeks. Demographic and clinical information, medications, blood pressure, serum glucose, and lipoprotein profiles were determined for all participating subjects. Consent was obtained from the probands and family members for plasma sampling, DNA isolation and skin biopsy. The research protocol was reviewed and approved by the Research Ethics Board of the McGill University Health Center.

In thirteen of these families, various mutations in ABCA1 were identified as the cause of low HDL-C[101, 106]. One family had a mutation in the SMPD1 gene (the cause for Niemann-Pick disease type I, sub-groups A and B) that segregated with low HDL-C in the family[189]. For this study I focused on families with no known cause of low HDL-C (severe hypertriglyceridemia– defined as plasma triglycerides >10 mmol/L, cellular cholesterol or phospholipid efflux defect, or previously known mutations in genes associated with HDL-C deficiency). In total, 25 informative families with 711 individuals and 20 singleton probands were investigated in this project.

2.3 Experimental approach

Using families of probands with HDL-C <5th percentile for age and sex-matched subjects, two approaches were used to identify the genes related to HDL; **Candidate gene approach**: Candidate genes were screened either by direct sequencing or by haplotyping a collection of well-characterized and informative microsatellite markers flanking the genes of interest.

Genome-wide scan: In families with sufficient meiosis and affected individuals, identified chromosomal regions that segregated with the low HDL-C by either binary or quantitative trait loci (QTL) analysis were densely mapped with microsatellites and SNPs. Genetic databases (NCBI, UCSC) were queried to identify all genes in the LOD-1 region.

2.4 General objective

The main objective of this project is to identify novel genes and their variants related to HDL metabolism through the genetic analysis of highly informative families of French Canadian descent.

2.5 Specific objectives

The research project is composed of four specific objectives that have been the subject of four different publications in scientific journals.

2.5.1 Specific objective 1

The first part of the project was the direct sequencing of *ApoAI* and *LCAT*, known genes involved in HDL metabolism and thus possible causes of low HDL-C. The objective was to identify non-synonymous variants in these genes that segregate with low HDL-C in the carrier families. Using this approach, I sequenced the *ApoAI* and *LCAT* genes in all 54 probands. Direct sequencing of genomic DNA revealed a novel mutation of the *ApoAI* gene, $ApoAI_{E136X}$, which was identified in 3/54 patients (6%) with low HDL-C. In two families, the E136X

mutation showed a strong co-segregation with low HDL-C levels. In addition, we examined a group of 105 CAD patients with a low HDL-C for the presence of the $ApoAI_{E136X}$ mutation by restriction endonuclease digestion. The result of this work has been published in journal of atherosclerosis: "*A novel nonsense apolipoprotein A-I mutation (ApoAI (E136X)) causes low HDL cholesterol in French Canadians. Atherosclerosis. 2006 Mar;185(1):127-36. Epub 2005 Jul 14*." and will be presented in detail in **CHAPTER 3.**

2.5.2 Specific objective 2

The disorder Niemann-Pick disease, types A and B have identified as a cause of low HDL-C [189]. The *SMPD1* that codes for acid sphingomyelinase is the cause of Niemann-Pick diseases A and B. In order to investigate the possible association of common variation in *SMPD1* and low HDL-C in humans, two common coding polymorphisms in the *SMPD1* gene, G1522A (G508R) and a hexanucleotide repeat sequence within the signal peptide region, were investigated in 118 unrelated subjects of French Canadian descent with low plasma levels of HDL-C (< 5th percentile for age and gender-matched subjects). Control subjects (n=230) had an HDL-C level greater than the 25th percentile. The results suggested that the two common coding variants at the *SMPD1* gene locus are not associated with low HDL-C levels in the French Canadian population, and are published in the journal BMC Medical Genetics: "*Sphingomyelin phosphodiesterase-1 (SMPD1) coding variants do not contribute to low levels of high-density lipoprotein cholesterol. BMC Med Genet. 2007 Dec 18; 8:79."* and will be presented in **CHAPTER 4.**

2.5.3 Specific objective 3

We performed a genome scan on 25 families, including 13 mutli-generational families with an average of 12 affected individuals per family. The genome scan was executed using 485 autosomal microsatellite markers covering the human genome with an average marker density of about 6 cM. Using parametric linkage analysis, a locus was identified on chromosome 4q31.21, which was subsequently fine-mapped. The genes in the LOD-1 region were sequenced from probands of two linked families. This finding has been published in the journal of Arteriosclerosis, Thrombosis, and Vascular Biology: " *Evidence for a gene influencing high-density lipoprotein cholesterol on chromosome 4q31.2 Arterioscler Thromb Vasc Biol. 2006 Feb; 26(2):392-7.*" and will be presented in **CHAPTER 5.**

2.5.4 Specific objective 4

In addition, we have performed a genome-wide QTL analysis. Variance components methods implemented in the software SOLAR identified several regions of linkage to the HDL-C trait, including a locus on chromosome 16q that had a LOD score of 2.69 at 95 cM. I fine-mapped this region on chromosome 16q23-24 and sequenced all coding regions of the 39 genes located in this region. The R162G variant of the *CHST6* gene consistently segregated with HDL-C affection in four families. However, an association study in 136 cases and 218 controls did not confirm this finding. RT-PCR results demonstrated a significant difference in the expression level of *CHST6* and *KIAA1576* in this region. The

results of this work has been resubmitted in revised form to the European Journal of Human Genetics: "*Fine Mapping and Association Studies of a High-Density Lipoprotein Cholesterol Linkage Region on Chromosome 16 in French-Canadian Subjects"* and will be presented in **CHAPTER 6.**

Preface to chapter 3

This chapter is based on candidate gene approach using direct sequencing. The objective was to identify non-synonymous variants in candidate genes that segregated with low HDL-C in the carrier families. Using this approach, I sequenced the *ApoAI* and *LCAT* genes in 54 probands. The result of this work has been published: "*A novel nonsense apolipoprotein A-I mutation (ApoAI (E136X)) causes low HDL cholesterol in French Canadians. Atherosclerosis. 2006 Mar;185(1):127-36. Epub 2005 Jul 14*."

CHAPTER 3

A NOVEL NONSENSE APOLIPOPROTEIN A-I MUTATION (APOA-I_{E136X}) CAUSES LOW HDL CHOLESTEROL IN FRENCH CANADIANS

Zari Dastani[1], Carole Dangoisse[1], Betsie Boucher[1], Katia Desbiens[1], Larbi Krimbou[1], Robert Dufour[2], Robert A. Hegele[3], Päivi Pajukanta[4], James C. Engert[1], Jacques Genest[1], Michel Marcil[1]

From the [1]Division of Cardiology, McGill University Health Center, Royal Victoria Hospital Montréal, Québec and the [2]Clinical Research Institute of Montreal, Montréal, Québec, Canada; the [3]Robarts Research Institute, London, Ontario, Canada; the [4]Department of Human Genetics, David Geffen School of Medicine at UCLA, University of California, Los Angeles, California, USA.

Key words: Hypoalphalipoproteinemia; High-density lipoprotein deficiency; Apolipoprotein AI; Gene mutation; Coronary artery disease; French Canadian population.

Atherosclerosis **2006, 185(1):27-136**

3.1 Abstract

The molecular causes of severe high-density lipoprotein cholesterol (HDL-C) deficiency was examined in a group of 54 unrelated French Canadian subjects. The lecithin:cholesterol acyl transferase (LCAT) and apolipoprotein (apo) AI gene were analyzed in all probands by direct DNA sequencing. While no LCAT mutation was detected, a novel nonsense apoAI mutation (E136X) was found in 3/54 probands. Genetic analysis of two kindreds showed a strong co-segregation of the apoAI locus with the low HDL-C trait. The E136X mutation was detected in families by *Mae*I restriction digestion. E136X carriers (n=17) had marked HDL-C deficiency; among the nine carriers ≥35 years old, five men had developed premature coronary artery disease (CAD). A peptide of apparent molecular weight of 14 KDa was identified in fresh plasma, the HDL fractions, and lipoprotein deficient plasma from the three probands but not in normal controls (n=3), suggesting that the mutant apoAI peptide is secreted and binds lipids. The mutation was not observed in an additional 210 chromosomes from unrelated subjects of French Canadian descent, <60 years of age, with CAD, and low HDL-C levels. We conclude that apoAI (E136X) is a cause of HDL-C deficiency in the French Canadian population and is associated with premature CAD.

3.2 Introduction

Genetic lipoprotein disorders are frequently seen in patients with premature coronary artery disease (CAD)[190]. One of the major, categorical cardiovascular risk factors is a decreased plasma level of high-density lipoprotein cholesterol (HDL-C). In many cases, a low HDL-C level reflects increased hepatic secretion of apolipoprotein (apo) B–containing lipoproteins and is a feature of the metabolic syndrome [25, 191]. In patients with very low levels of HDL-C, mutations of genes involved in the metabolism of HDL particles have been identified. The most commonly reported gene defects in patients with a low HDL-C involve the ATP binding cassette A1 (ABCA1), apolipoprotein (apo) AI and lecithin:cholesterol acyltransferase (LCAT) genes[101, 104, 134]. Many other genetic defects can cause a low HDL-C level but are relatively rare; conversely, mutations in the cholesteryl ester transfer protein (CETP) gene are associated with elevated HDL-C[192].

HDL particles are involved in various metabolic processes, including cholesterol transport[120], vascular endothelial function, protection against LDL particle oxidation and anti-inflammatory properties[38, 193-196]. The best studied role of HDL particles and presently considered to be the most important against the development of atherosclerotic cardiovascular disease – is reverse cholesterol transport. This pathway involves the secretion of apoAI (the major protein moiety in HDL) from hepatic and intestinal cells, the binding of apoAI to the cellular ABCA1 transporter and formation of nascent HDL particles, the esterification of free cholesterol by LCAT and the formation of spherical HDL_3 and HDL_2

particles, the exchange of core cholesteryl esters for triglycerides to apo B–containing lipoproteins via cholesteryl ester transfer protein (CETP), the selective uptake of cholesteryl esters via the scavenger receptor B1 (SR-B1), and the hydrolysis of core triglycerides by hepatic lipase and endothelial lipase[197]. At least 49 mutations of the apoAI gene have been identified in the past 2 decades and are associated with HDL deficiency states, amyloidosis and premature CAD[120]. Of these, four are nonsense mutations, leading to a premature termination of the protein (Q-2X, W8X, Q32X, and Q84X)[198-201]. In none of these mutations, could the truncated peptide be identified in plasma.

In the present study, we identified a novel apoAI mutation, $apoAI_{E136X}$, in 3 out of 54 unrelated probands of French Canadian descent selected for low levels of plasma HDL-C.

3.3 Methods

3.3.1 Subject selection

All study subjects were examined and sampled in the Preventive Cardiology / Lipid Clinic of the McGill University Health Centre. The inclusion criteria include a HDL-C < 5th percentile (age and gender matched based on the Lipid Research Clinics Population Studies Data Book, as previously described[1]) and no known cause of low HDL-C (severe hypertriglyceridemia – defined as plasma triglycerides >10 mmol/L, cellular phospholipid efflux defect, or previously known mutations in the genes associated with HDL-C deficiency). When possible, all available living relatives were asked to participate in the study.

Family members were sampled under strict conditions (12 h fast, discontinuation of lipid modifying medications for at least 4 weeks). Demographic and clinical information, medications, blood pressure, serum glucose, and lipoprotein profiles were determined on all participating subjects. Separate consent forms were obtained from the probands and family members for plasma sampling, DNA isolation and skin biopsy. The research protocol was reviewed and approved by the Research Ethics Board of the McGill University Health Center. A second group consisted of 105 unrelated patients of French Canadian origin with premature CAD (mean age, 48±9 years) selected for a HDL-C < 5th percentile[202].

3.3.2 Biochemical analysis

Plasma was isolated in all study subjects, after a 12 h fast, in EDTA–containing tubes. The buffy coat was kept for DNA isolation. Lipids and HDL-C were measured using standard techniques and the LDL-C was calculated according to the Friedewald formula (LDL-C = total cholesterol – triglycerides/2.2 - HDL-C, all concentration in mmol/L), unless the triglycerides were > 4.5 mmol/L. In this case, ultracentrifugation of plasma was used and the lipoprotein cholesterol concentration was measured directly.

3.3.3 Cellular analysis

ApoAI–mediated cellular phospholipid efflux assays were performed as previously described[20] on all probands to examine possible deficiencies of

cellular lipidation of HDL most likely associated with functional variants of the ABCA1 transporter gene[203].

3.3.4 Haplotyping

Genotypes (deCODE genetics, Reykjavik, Iceland) at markers D11S4206, D11S908 and D11S4089, flanking the apoAI gene locus on chromosome 11q22-23 were used to construct haplotypes.

3.3.5 DNA sequencing

Direct sequencing of the apoAI and LCAT genes was performed on all probands. Exon-specific oligonucleotides were synthesized (Operon Biotechnologies, Huntsville, AL), designed in such a way that each intron-exon boundary included at least 10 bp intronic. Sequencing was performed using the dye termination method on a ABI PRISM® 3100-*Avant* Genetic Analyzer (Applied Biosystems, Foster City, CA). To rule out sequencing artifacts, all variants identified were sequenced bi-directionally. The PHRED/ PHRAP/ CONSED programs were used for sequence alignment and analysis.

3.3.6 Detection of mutation

Genotyping for the E136X mutation was performed by PCR amplification in exon 4 (412-bp) followed by restriction digestion with *Fsp*BI (*Mae*I). The forward primer in intron 3 is 5'-AGCCCTCAACCCTTCTGTCT-3'; and the reverse primer in exon 4 is 5'-CAGCTCGTCGCTGTAGGG-3'; annealing

temperature is 63°C. PCR product (15μL) was incubated with *Fsp*BI (Fermentas Canada, Burlington, ON) (5U) in a total volume of 25 μL overnight at 37°C, and the resulting fragments were separated on a 1.5% agarose gel. In the presence of the mutation, a cleavage site is created, resulting in 308- and 104-bp fragments.

3.3.7 Protein analysis

Plasma was obtained from the three probands and three normal controls, and HDL was isolated by differential ultracentrifugation at density (*d*) $1.063 < d < 1.210$ g/mL, using KBr to adjust the density of plasma. In 250 μL of the total plasma, the HDL fractions and the lipoprotein deficient fractions, apoAI was immunoprecipitated by a polyclonal anti-human apoAI antibody (Biodesign International, Saco, ME) and 100 μg of protein were separated on 8-27% polyacrylamide gradient gel (PAGGE; 125 V, 24 h, 4°C), Electrophoretically separated samples were electrotransferred (30 V, 24 h, 4°C) on to nitrocellulose membranes (Hybond ECL, Amersham Bioscience). ApoAI was detected by incubating the membranes with immunopurified polyclonal anti-apoAI antibody (Biodesign International, Saco, ME).

3.3.8 Separation of lipoproteins by two-dimensional non-denaturating gradient gel electrophoresis

Plasma apoAI-containing particles from a normal subject and the three probands were analyzed by two-dimensional non-denaturating gradient gel electrophoresis, as previously described[204]. Briefly, Lipoproteins from plasma samples (200 μL) were separated in the first dimension (according to particle

charge) by 0.75% agarose gel electrophoresis (100 V, 8 h, 4°C) and in the second dimension (according to particle size) by 2–16% polyacrylamide concave gradient gel electrophoresis (80 V, 20 h, 4°C). Iodinated high molecular weight protein mixture (7.1–17.0 nm, Amersham Biosciences) was run as a standard on each gel. Electrophoretically separated lipoproteins and proteins were electrotransferred (30 V, 20 h, 4°C) onto nitrocellulose membranes (Hybond ECL; Amersham Biosciences). Membranes were then incubated for 30 min in phosphate-buffered saline (PBS) containing 5% non-fat milk powder. ApoAI-containing lipoproteins were detected by incubating membranes with affinity-purified goat polyclonal anti-human apoAI antibody (Biodesign International, Saco, ME.) labeled with ^{125}I. After incubation with anti-apoAI, membranes were washed three times (30 min) with PBS containing 0.05% (v/v) Tween 20, and the presence of labeled antibody was revealed by autoradiography on XAR-2 Kodak film.

3.3.9 LCAT assay

Ability of apoAI to esterify cholesterol via an exogenous LCAT was analyzed on HDL from the three probands and a normal control subject. The HDL fraction was isolated from probands and a normolipidemic subject by differential ultracentrifugation, as mentioned above. Twelve-well plates were coated with free ^{3}H-cholesterol (2.5 μCi Perkin Elmer Life Sciences Canada) in 500 μL methanol per well. In the HDL fractions, endogenous LCAT activity was initially inactivated by heating at 56ºC for 45 min and triplicate HDL samples

(420 μg protein/well) were labeled overnight at 4ºC in ^{3}H-cholesterol coated wells. Ten μg protein of labeled HDL (≈10,000 cpm) were added to 0.1 M Tris-HCl, pH 8.3 containing 5 μL of 30% (v/v) bovine serum albumin, 0.5% (v/v) β-mercaptoethanol, and 600 ng (0.005 U/ng) of exogenous LCAT (C-terminal histidine-tagged human recombinant lecithin: cholesterol acyltransferase (hrLCATH6), generously provided by Dr. John S. Parks, Wake Forest University School of Medicine, Winston-Salem, NC) in a total volume of 200 μL. The reaction mixture was incubated at 37°C for 1 h and then extracted in Folch reagent (chloroform-methanol 2:1 v/v) at 37°C for 2 h. Free and esterified ^{3}H-cholesterol were separated on thin layer chromatography plates (Silica Gel Linear K; Whatman, Clifton, NJ), developed in heptane-ether-methanol-acetic acid (80:30:3:1.5 v/v/v/v), and counted by scintillation. Results are expressed as the percent of esterified over the total (free + esterified) ^{3}H-cholesterol.

3.4 Results

As part of a research project on the genetics of HDL, data on 59 probands with severely reduced levels of plasma HDL-C have been collected. Also, 40 families of probands with low HDL-C, including more than 900 family members, have been recruited. At least one first degree relative was affected by a low HDL-C in the recruited families. The size of families varies between 3 and 109, with a median of 14 relatives per family. We have previously reported in four of these subjects the ABCA1 gene as the cause of Tangier disease and familial HDL

deficiency[101, 104] and in one subject, the sphingomyelin phosphodiesterase-1 gene, coding for acid sphingomyelinase as another genetic cause of a low HDL-C[189]. These 5 probands were eliminated from this study. In the present study, we examined 54 unrelated probands of French Canadian descent, selected for a HDL-C < 5th percentile for age and gender. None of the probands had diabetes, severe hypertriglyceridemia (fasting plasma triglycerides > 10 mmol/L) or known mutations at genes modulating HDL-C levels. A defective cellular lipidation of HDL, commonly associated with ABCA1 gene defects, was tested by performing cellular phospholipid efflux on skin fibroblasts on all probands. A reduced cellular phospholipid efflux was found in 7 probands. In the remaining subjects, cellular phospholipid efflux was within normal limits, according to our previously established criteria [203]. Using a candidate gene approach, we sequenced the apoAI and LCAT genes in all 54 probands.

Direct sequencing from genomic DNA revealed, in 3 unrelated probands, a heterozygous G→T substitution at position 478 in exon 4 of the apoAI gene (data not shown), predicting a substitution of glutamic acid (*G*ag) for a stop codon (*T*ag) at residue 136 (E136X). LCAT gene sequencing, however, showed no functional variant on all 54 probands.

The presence of the E136X mutation was also detected by endonuclease digestion (as illustrated in **Fig. 1**), where a *Mae*I site is present in a 412-bp fragment amplified from the 5' end of apoAI–exon 4. As shown in **Fig. 1**, the heterozygous mutation was confirmed on the three affected probands, compared

to DNA from a normal control amplified and submitted to *Mae*I digestion in the same conditions.

Haplotype encompassing the apoAI gene has been constructed in a region spanning 11.4 Mb surrounding the gene, using the three genetic markers D11S4206, D11S908, and D11S4089 localized respectively at 7.2 and 1.5 Mb from the gene in the centromeric region and 4.2 Mb in the telomeric region. Haplotyping data, including the E136X mutation detection in the apoAI gene, were examined in two available kindred from affected probands (**Fig. 2A and 2B**). There was a strong co-segregation of the mutated allele with the phenotype of low HDL-C levels in both kindred.

In the first pedigree (**Fig. 2A**), 7 subjects (5M, 2F) are found heterozygous carriers of the mutation with a HDL-C < 5^{th} percentile. All the three male carriers, 35 years old or older, had developed premature CAD. Only one female carrier, older than 35 years, had not developed premature CAD. The proband underwent coronary by-pass grafting (CABG) at age 53; one of his brothers, subject 307, also has premature CAD, and underwent CABG at age 48; and subject 403 underwent percutaneous coronary intervention (angioplasty) at age 30. All three subjects have a low HDL-C and carry the $apoAI_{E136X}$ mutation. Interestingly, individual 406, who is the spouse of a non-carrier subject in this kindred, shows a decreased HDL-C below the 5^{th} percentile without carrying the $apoAI_{E136X}$ mutation. Also, the son (individual 509) of this subject shows a severe decrease in HDL-C, suggesting that another molecular cause for a HDL-C deficiency might

be present in this pedigree. These subjects have not been investigated in the present study.

The second pedigree (**Fig. 2B**) shows 9 subjects (5M, 4F) who carry the heterozygous mutation and all have a decreased HDL-C < 5th percentile. In carriers older than 35 years, 2/3 males had developed premature CAD, whereas the only female had not developed CAD at 58 years of age. The proband had a diagnosis of angiographically documented CAD and underwent CABG at age 39; one brother (subject 307) had CAD diagnosed at age 56 (CABG), has a very low HDL-C and also carries the apoAI$_{E136X}$ mutation. Three females (individuals 302, 313, and 404), non-carrier of the mutation, have an unexplained decrease of HDL-C level, suggesting the presence of other genetic, metabolic or environmental causes of a decreased HDL-C level in this kindred.

Haplotyping in the kindred from the third affected proband has not been carried out due to the unavailability of the family members.

We examined 62 family members from probands 1 and 2; of those, 14 were affected with apoAI$_{E136X}$. All affected subjects had an HDL-C level < 5th percentile for age- and gender-matched subjects. **Table 1** shows the lipid and lipoprotein lipid levels in the three affected probands and affected family members. Proband 1 had a HDL-C of 0.67 mmol/L with a normal triglyceride level of 1.36 mmol/L. Proband 2 had moderate hypertriglyceridemia (8.23 mmol/L) and a HDL-C of 0.37 mmol/L. Similarly, proband 3 had moderate hypertriglyceridemia (7.45 mmol/L) with a HDL-C of 0.17 mmol/L. ApoB levels were normal in proband 1 and 2 and elevated in proband 3, while apoAI was

decreased in the three probands. **Table 2** shows the mean values for the group of 54 probands with HDL-C deficiency, as well as a second group of patients with premature CAD (n=105) and markedly reduced HDL-C levels. A reference group[202], consisting of normolipidemic subjects (152 men and 61 women) of French Canadian origin, matched for age is shown for comparison purposes.

The three probands were carefully examined by one of us (J.G.). There were no cutaneous or tendinous xanthomas, no xanthelasmas, no corneal opacifications seen on slit lamp examination and no enlarged lymph nodes, liver or spleen. There is no family history of amyloidosis although rectal biopsies or echocardiographic examinations were not performed on any of the subjects, for the diagnosis of systemic or cardiac amyloidosis.

We also examined a group of 105 CAD patients with a low HDL- C for the presence of the $apoAI_{E136X}$ mutation, by endonuclease digestion. We did not find the mutation in any of the 210 chromosomes analyzed in this second group of CAD patients with low HDL-C.

ApoAI protein from the three probands and three normal controls was examined by electrophoretic separation on polyacrylamide gradient gel and immunodetection with a polyclonal anti-apoAI antibody. **Fig. 3** shows the apoAI found in HDL fractions from the three probands (left panel) and the three normal subjects (right panel). A band below the molecular weight 30 kDa (standards not shown), observed in both probands and normals, corresponds to the major form of apoAI protein with a normal molecular weight of 28.3 kDa. A distinct band, very likely the truncated form of apoAI, close to the molecular weight 14 kDa,

was identified only in subjects with the apoAI$_{E136X}$ and was not present in normals. This band was also detected in apoAI from total plasma and lipoprotein deficient plasma fractions from the three probands but not in the three normals (data not shown). This observation suggests that the apoAI$_{E136X}$ encodes a truncated form of apoAI protein, which is secreted in the plasma and is partially lipidated.

Plasma apoAI-containing lipoproteins were examined by two-dimensional gel electrophoresis in the three probands, compared to a normal control subject (**Fig. 4**). There was a marked reduction in the larger mature HDL particles (HDL$_{2a}$ and HDL$_{2b}$) from the three probands. The lack of these particles seems to correlate respectively with the low levels of HDL-C measured in the total plasma of the three probands.

The ability of the HDL fraction to activate LCAT was assessed using HDL fractions from the three probands and a normal control subject. Equal amounts of HDL proteins were used in the *in vitro* LCAT assay, performed after inactivating native LCAT. The results of the LCAT activity assay, determined by the percentage of esterified cholesterol from the total radiolabeled cholesterol in each HDL fraction, is shown in **Fig. 5**. HDL from heterozygous apoAI$_{E136X}$ maintain the ability to esterify cholesterol, due to the functioning apoAI allele. As negative control, we also performed the *in vitro* LCAT assay from reconstituted LpAI (Δ122-165), which has been previously shown to prevent the LCAT reaction[120], and found no esterification reaction in this system (data not shown).

3.5 Discussion

A low HDL-C is considered a major risk factor for the development of CAD[25]. The majority of cases of low HDL-C are associated with the male gender, obesity, high blood pressure, hypertriglyceridemia, a sedentary or an unhealthy lifestyle, including cigarette smoking. A low HDL-C is a diagnostic criterion for the metabolic syndrome[2]. Increased hepatic secretion of apoB–containing lipoproteins and hypertriglyceridemia are thought to cause the secondary decrease in HDL-C.

There are rare genetic mutations in the genes associated with HDL metabolism. These include structural proteins (apoAI), modifying enzymes (LCAT), and the ABCA1 transporter, the genetic defect in Tangier disease and familial HDL deficiency. Here, we report a novel mutation of the apoAI gene, $apoAI_{E136X}$, identified in 3/54 patients (6%) with severe HDL-C deficiency, defined as a HDL-C < 5^{th} percentile for age and gender. This mutation was not identified in another 105 patients from same ethnic origin, with CAD and a low HDL-C. In the two kindred examined, including a total of 62 subjects, in addition to the probands, the $apoAI_{E136X}$ was identified in 14 individuals. All had a HDL-C level < 5^{th} percentile for age- and gender-matched subjects and mild to moderate hypertriglyceridemia (**Fig. 2A, 2B and Table 1**). Premature CAD was documented in probands 1 and 2, and in three additional family members. We did not perform non-invasive testing for atherosclerosis in other family members.

This novel mutation is located within the 5th predicted repeat unit of the mature apoAI peptide, and the predicted mutant peptide would be expected to be slightly more that half the secreted mature apoAI.

The $apoAI_{E136X}$ mutant protein appears to be secreted in plasma, as shown in **Fig. 3**. A peptide with an apparent molecular weight of ≈14 KDa, was found on PAGGE in the HDL fraction and in the lipoprotein deficient fraction after preparative ultracentrifugation. This truncated apoAI peptide nevertheless appears to bind lipids (hence its presence in the HDL fraction) and is not identified in three normolipidemic subjects. The plasma concentration of the peptide appears to be markedly less than the wild type apoAI. This suggests that the peptide might be unstable in plasma and rapidly cleared, likely by the kidneys. Of note, the patients examined so far have no clinical features of amyloidosis or nephropathy.

At least 49 mutations of the apoAI gene have been reported[120, 205]. Some of these mutations are associated with hereditary amyloidosis and appear to cluster in the amino terminus of the mature peptide; all consist of amino acid changes. Another cluster of mutations near the apoAI repeat units 6, 7 and 8 (residues143-208)[120], consisting of amphipatic α-helices, are associated with a low HDL-C. We postulate that $apoAI_{E136X}$ would be unable to activate LCAT, thereby contributing to the low HDL-C in our patients. To determine whether HDL from affected patients (heterozygous for $apoAI_{E136X}$) could still activate LCAT with the normal apoAI, we examined the formation of cholesteryl esters using an *in vitro* assay. We used equal amounts of HDL protein from control and patients. We have shown in **Fig. 3** that the majority of apoAI in the plasma of

affected subjects contained the normal apoAI. The results shown on **Fig. 5** confirm that the remaining apoAI allele can activate LCAT normally.

Genetic analysis in the two affected kindred suggest that this mutation is autosomal dominant. All heterozygous carriers of the mutation have an HDL-C level < 5^{th} percentile, and plasma HDL-C levels are decreased by approximately 50%, suggesting that the mutant apoAI protein is unable to carry cholesterol in plasma.

The $apoAI_{E136X}$ mutation was identified in three unrelated probands of French Canadian descent with severe HDL-C deficiency. This mutation is associated with premature CAD in our family studies. The same mutation was not discovered when 128 individuals and 155 individuals were sequenced in the US and Canada, respectively[64]. In addition, we were unable to identify this mutation in a second group of patients with premature CAD, selected for a low HDL-C (**Table 1**). The possibility of a founder effect within the French Canadian population remains to be examined. While the prevalence of this mutation in our initial group of patients is 6%, the prevalence is likely much lower in the general population.

3.6 Acknowledgments

We thank Dr. John S. Parks for the hLCATH6 generously provided. This work was supported by grants MOP 15042 and 62834 from the Canadian Institutes of Health Research (CIHR). J. Genest holds the McGill University-Novartis Chair in Cardiology.

3.7 Figures

Figure 1. Mutation detection by endonuclease digestion.

Fragment of 412 bp in apoAI-exon 4 was amplified by PCR under conditions described in "Section 3.3.6". As shown in the schematic representation of apoAI gene (*lower panel*), the presence of the E136X mutation creates a *Mae*I restriction site cleaving the fragment into 308 and 104 bp. *Upper panel*, after restriction digestion, PCR fragments from the three probands (PD1, PD2, PD3) and a normal control subject (NL) was separated on 1.5% agarose gel and revealed the heterozygous mutation (+/−) only in the probands

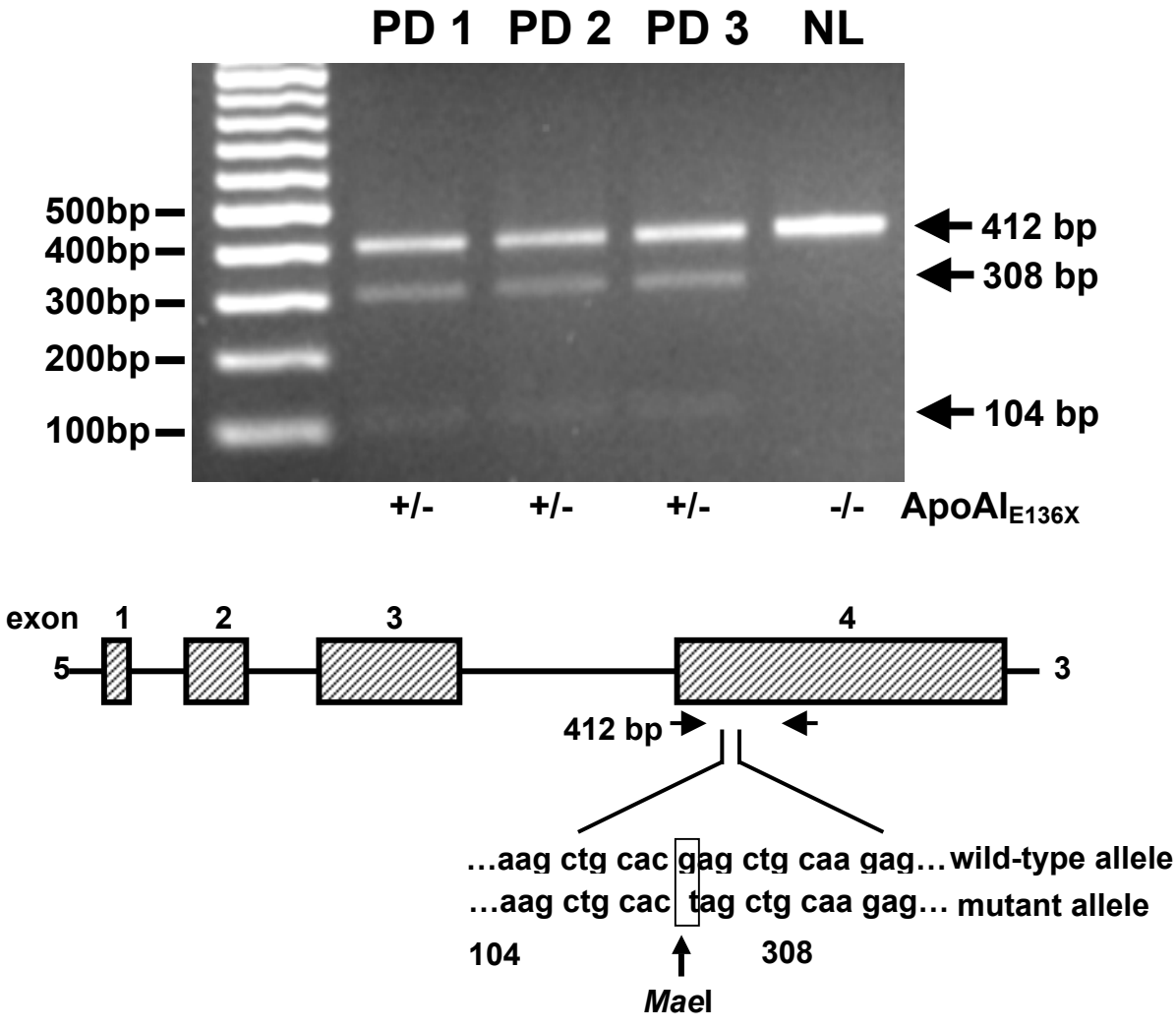


Figure 2. Genetic analysis of the families of proband 1 and 2.

(A) pedigree of proband 1; (B) pedigree of proband 2. Squares represent males and circles represent females. Diagonal lines and crosses through the symbols indicate deceased inviduals or unavailable data, respectively. Haplotypes at the apoAI gene locus (region 11q22–23) include three genetic markers flanking the gene and the E136X mutation (0, absence; 1, presence). Filled bars indicate alleles bearing the mutation. Horizontal line in bars or partly filled bars represent meiotic recombinations. Individual number (as referred in Table 3.2), age, HDL-C value followed by percentile ranking for age and gender are listed below each symbol. Black symbols indicate individuals with a HDL-C value below the fifth percentile, consistent, for most of the cases, with the presence of the E136X mutation in one allele. *, individual with documented premature coronary artery disease; parentheses, inferred marker genotypes; large arrow, proband.

Figure 2A.

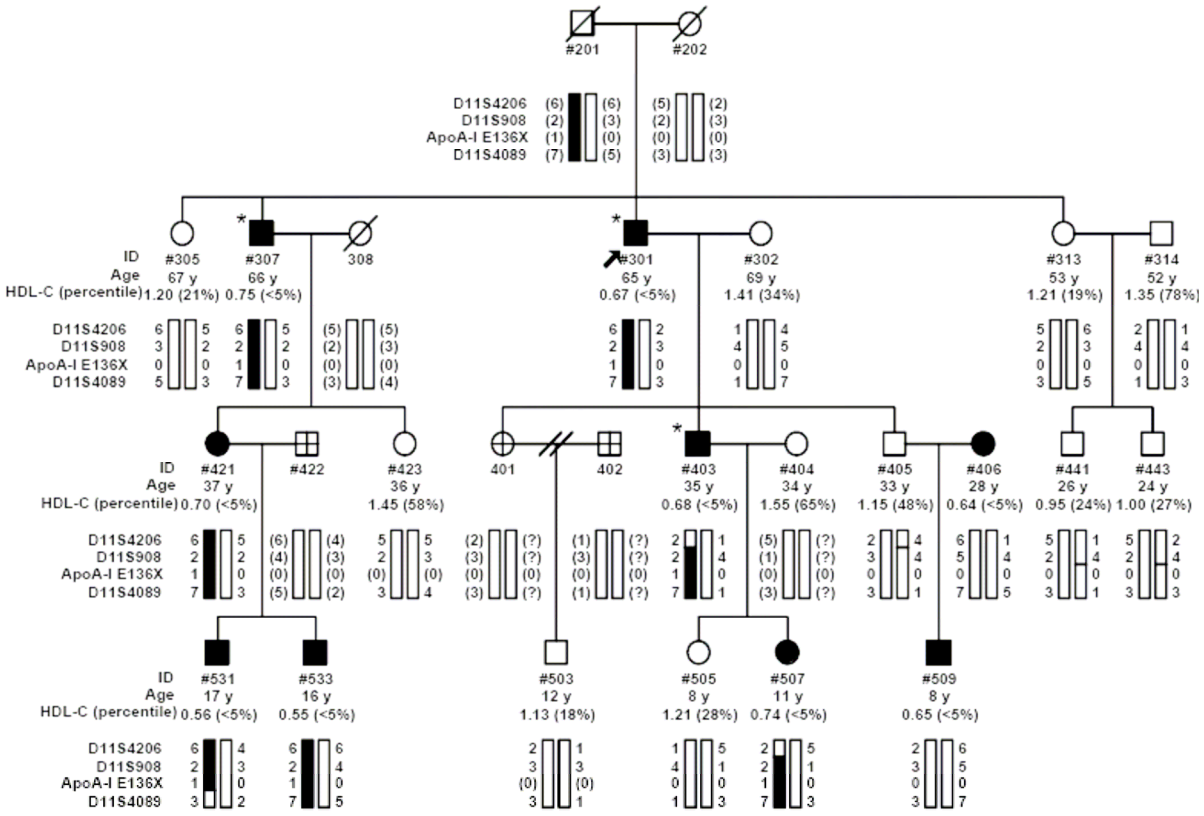

#201
#202
D11S4206 (6) (6) (5) (2)
D11S908 (2) (3) (2) (3)
ApoA-I E136X (1) (0) (0) (0)
D11S4089 (7) (5) (3) (3)
ID #305 #307 308 #301 #302 #313 #314
Age 67 y 66 y 65 y 69 y 53 y 52 y
HDL-C (percentile) 1.20 (21%) 0.75 (<5%) 0.67 (<5%) 1.41 (34%) 1.21 (19%) 1.35 (78%)
D11S4206 6 5 6 5 (5) (5) 6 2 1 4 5 6 2 1
D11S908 3 2 2 2 (2) (3) 2 3 4 5 2 3 4 4
ApoA-I E136X 0 0 1 0 (0) (0) 1 0 0 0 0 0 0 0
D11S4089 5 3 7 3 (3) (4) 7 3 1 7 3 5 1 3
ID #421 #422 #423 401 402 #403 #404 #405 #406 #441 #443
Age 37 y 36 y 35 y 34 y 33 y 28 y 26 y 24 y
HDL-C (percentile) 0.70 (<5%) 1.45 (58%) 0.68 (<5%) 1.55 (65%) 1.15 (48%) 0.64 (<5%) 0.95 (24%) 1.00 (27%)
D11S4206 6 5 (6) (4) 5 5 (2) (?) (1) (?) 2 1 (5) (?) 2 4 6 1 5 1 5 2
D11S908 2 2 (4) (3) 2 3 (3) (?) (3) (?) 2 4 (1) (?) 3 4 5 4 2 4
ApoA-I E136X 1 0 (0) (0) (0) (0) (0) (?) (0) (?) 1 0 (?) 0 0 1 0 0 0 0
D11S4089 7 3 (5) (2) 3 4 (3) (?) (1) (?) 7 1 (3) (?) 3 1 7 5 1 3 3 3
ID #531 #533 #503 #505 #507 #509
Age 17 y 16 y 12 y 8 y 11 y 8 y
HDL-C (percentile) 0.56 (<5%) 0.55 (<5%) 1.13 (18%) 1.21 (28%) 0.74 (<5%) 0.65 (<5%)
D11S4206 6 4 6 6 2 1 1 5 2 5 2 6
D11S908 2 3 2 4 3 3 4 1 2 1 3 5
ApoA-I E136X 1 0 1 0 (0) (0) 0 1 1 0 0 5
D11S4089 3 2 7 5 3 1 1 3 7 3 3 7

Figure 2B:

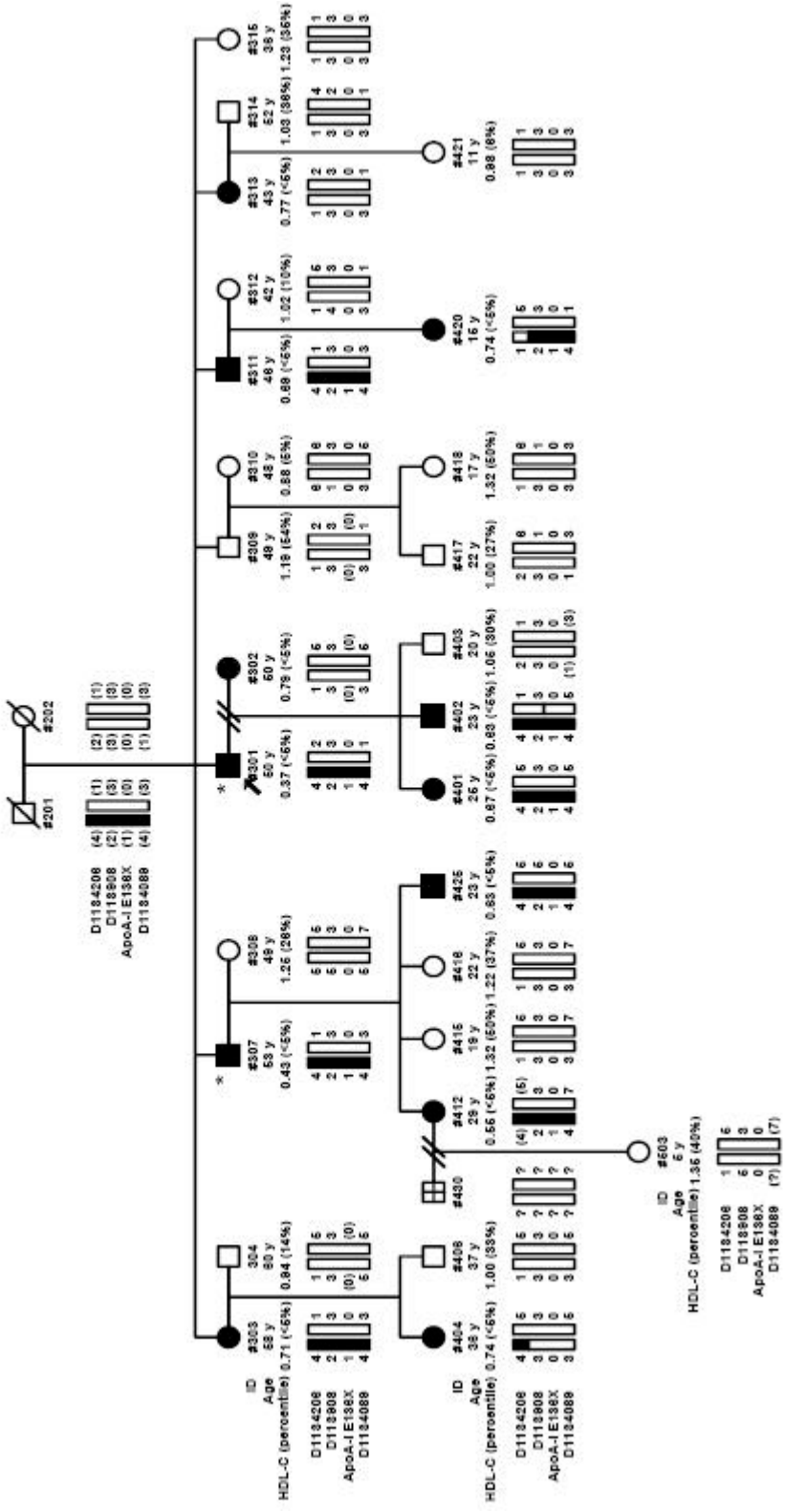


Figure 3. Polyacrylamide gradient gel electrophoresis separation of HDL apoAI from the three probands and three normal subjects.

HDL fraction was isolated from plasma as described in "Section 3.3.7". ApoAI protein was immunoprecipitated from HDL fractions and separated on 8–27% polyacrylamide gradient gel. After electrotansfer onto nitrocellulose membrane, apoAI was immunodetected with a polyclonal anti-apoAI antibody. *Left panel*, HDL apoAI (E136X, +/−) from the three probands (PD1, PD2, PD3) shows a band below 30 kDa corresponding to the normal molecular weight of apoAI (28.3 kDA) produced by the normal allele; an additional band is observed below 14 kDa, most likely corresponding to the truncated form of apoAI produced by the allele bearing the E136X mutation. *Right panel*, in contrast, HDL apoAI (E136X, −/−) from the three normal controls (NL1, NL2, NL3) tested shows only one band corresponding to the normal apoAI molecular weight.

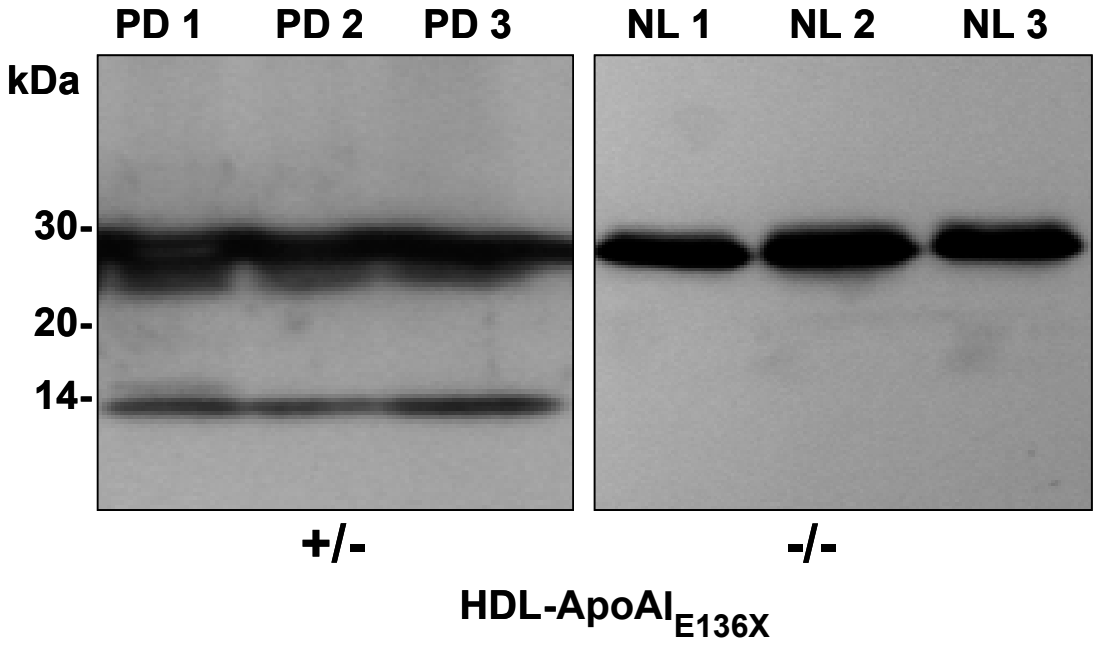


Figure 4. Two-dimensional gradient gel electrophoresis separation of apoAI-containing lipoproteins.

Plasma from a normal subject (NL) and the three probands (PD1, PD2, PD3) were first separated according to the particle charge (on agarose gel, 0.75%), then according to the particle size (on non-denaturating polyacrylamide gradient gel, 2–16%). Lipoproteins were electrotransfered onto nitrocellulose membrane and apoAI was detected with ^{125}I-labeled anti-apoAI antibody. Molecular size markers are indicated on the left-hand panel. ApoAI-containing HDL subpopulations are indicated with arrows on this panel. In the three probands, there is a marked reduction of HDL_{2a} and HDL_{2b} particles compared to the normal subject. In contrast, amount of nascent pre-β_1-LpAI is not affected and HDL_3 particles appears clearly less reduced.

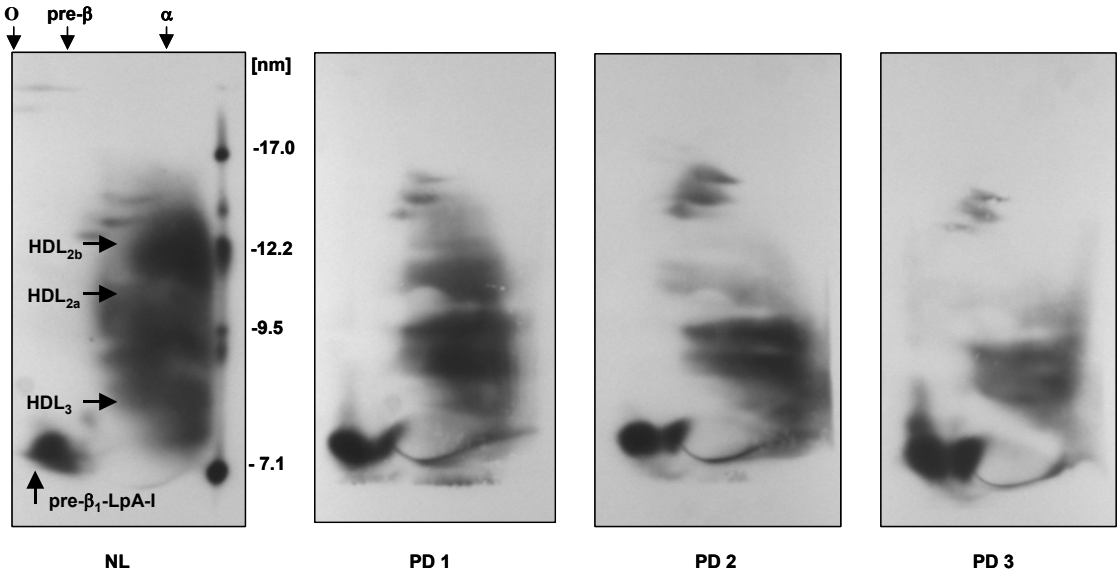
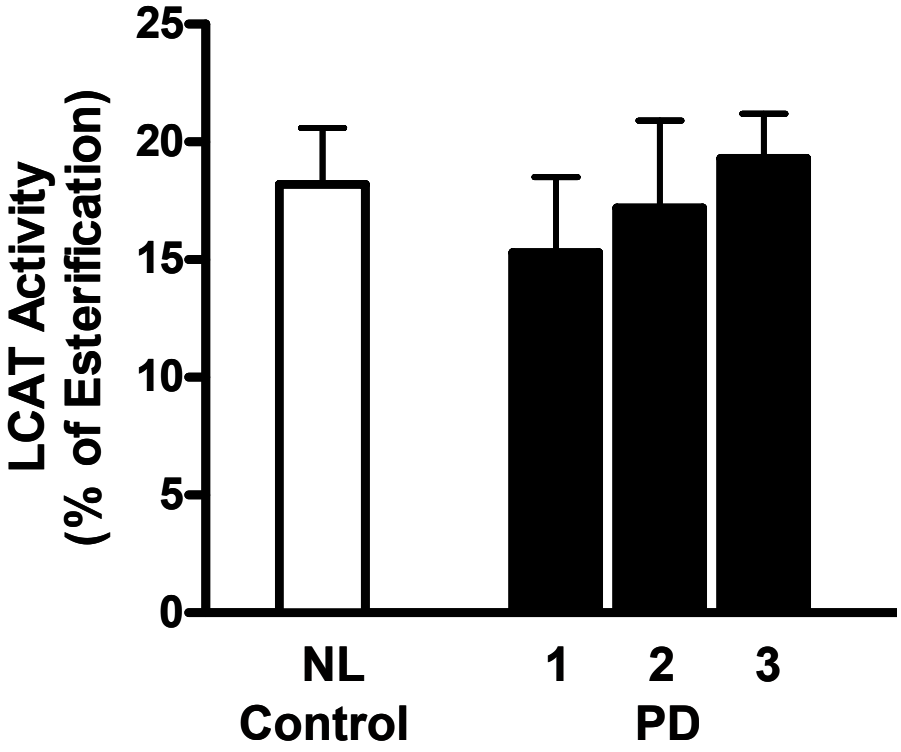


Figure 5. Ability of apoAI to activate LCAT in HDL from the three probands compared to a normal subject.

LCAT activity assay was performed in HDL fraction from the three probands (PD1, 2, 3) and a normal subject (NL control) according to the procedure described in "Section 3.2". From a same amount of HDL proteins, the percentage of cholesterol esterified in HDL during 1 h by an exogenous LCAT was determined. The ability of apoAI to esterify cholesterol in HDL fractions is not significantly different between the three probands and the normal control. Non-mutated apoAI, produced by the normal allele, seems the major form of apoAI present in HDL particles in probands.



3.8 Tables

Table 1. Characteristics of the probands and their family members heterozygous for ApoAI$_{E136X}$.

Subjects	Family no.	Age	Gender	BMI	Total Chol	Triglycerides	HDL Chol	LDL Chol	ApoA-I	ApoB	CAD
		y		*kg/m²*	*mmol/L*	*mmol/L*	*mmol/L*	*mmol/L*	*g/L*	*g/L*	
Proband 1	**301**	**64**	**M**	**28.4**	**4.56**	**1.36**	**0.67**	**3.26**	**0.81**	**0.66**	**+**
Kindred †	307	66	M	33.4	3.87	3.16	0.53	1.68	0.73	0.75	+
	403	35	M	29.4	5.08	1.56	0.68	3.62	0.81	1.36	+
	421	37	F	30.1	4.43	0.78	0.49	3.38	0.51	0.70	–
	507	11	F	22.4	5.38	0.83	0.74	4.00	0.59	0.89	–
	531	17	M	25.9	4.86	0.82	0.56	3.68	0.59	0.79	–
	533	16	M	21.8	3.32	1.04	0.32	2.30	0.50	0.71	–
Proband 2	**301**	**50**	**M**	**32.3**	**7.53**	**8.23**	**0.37**	**2.87**	**0.51**	**1.06**	**+**
Kindred †	303	57	F	25.6	5.76	4.63	0.71	2.77	0.81	1.35	–
	307	53	M	34.3	2.20	2.16	0.36	0.79	0.43	0.54	+
	311	46	M	31.6	6.40	2.28	0.50	4.67	0.59	1.54	–
	401	25	F	33.4	5.00	2.97	0.44	2.98	0.52	1.13	–
	402	23	M	23.0	4.57	1.68	0.52	3.18	0.59	0.95	–
	412	29	F	31.2	4.46	1.50	0.42	3.23	0.46	1.08	–
	420	15	F	19.3	3.37	0.65	0.69	2.33	0.50	0.55	–
	425	23	M	30.1	5.05	1.88	0.49	3.57	0.49	1.04	–
Proband 3	**-**	**52**	**M**	**29.3**	**7.45**	**6.50**	**0.17**	**n.d.**	**< 0.25**	**1.89**	**–**
Mean				*28.3*	*4.90*	*2.47*	*0.51*	*3.02*	*0.56*	*1.00*	
SD				*4.6*	*1.38*	*2.13*	*0.16*	*0.93*	*0.17*	*0.37*	

† Pedigree presented in Fig. 2A, Proband 1; and Fig. 2B, Proband 2.
BMI, body mass index; CAD, coronary artery disease; n.d., not determined.

Table 2. Characteristics of the two study groups of subjects with low HDL-C and a reference group of same origin.

Subjects	Age	Sex	Total Chol	Triglycerides	HDL Chol	LDL Chol	CAD
	y †	*n*	*mmol/L* †	*mmol/L* †	*mmol/L* †	*mmol/L* †	*n*
Probands							
M	54±12	46	4.8±1.5	3.2±2.3	0.6±0.2	2.7±0.9	18
F	52±11	8	4.9±1.0	3.9±1.9	0.6±0.1	2.6±0.9	3
CAD patients							
M	49±10	76	6.5±2.1	4.9±4.0	0.6±0.1	3.7±1.7	76
F	49±6	29	7.8±3.4	9.0±13	0.7±0.2	4.2±1.9	29
Normal Control ‡							
M	47±8	152	5.6±0.9	1.8±1.0	1.2±0.3	3.5±0.9	0
F	51±6	61	5.6±0.9	1.5±0.8	1.5±0.4	3.4±0.8	0

† Mean ± S.D.

‡ As reference of a healthy French Canadian population, described elsewhere[19].

CAD, coronary artery disease.

Preface to chapter 4

In part of candidate gene approach of my project, I investigated the possible association of common variations in *SMPD1* gene and low HDL-C in French Canadian. The results are published in the journal *BMC Medical Genetics*: ”*Sphingomyelin phosphodiesterase-1 (SMPD1) coding variants do not contribute to low levels of high-density lipoprotein cholesterol. BMC Med Genet. 2007 Dec 18;8:79.”* and will be presented in this chapter.

CHAPTER 4

SPHINGOMYELIN PHOSPHODIESTERASE-1 (SMPD1) CODING VARIANTS DO NOT CONTRIBUTE TO LOW LEVELS OF HIGH-DENSITY LIPOPROTEIN CHOLESTEROL

Zari Dastani[1], Isabelle L. Ruel[1], James C. Engert[1], Jacques Genest[1], Michel Marcil[1]

From the [1]Division of Cardiology, McGill University Health Center, Royal Victoria Hospital Montréal, Québec.

Key words: Association study, haplotypes, hypoalphalipoproteinemia, high-density lipoprotein, Niemann-Pick disease, polymorphisms, *SMPD1* gene.

BMC Med Genet. **2007, 18(8):79**

4.1 Abstract

Niemann-Pick disease type A and B is caused by a deficiency of acid sphingomyelinase due to mutations in the sphingomyelin phosphodiesterase-1 (*SMPD1*) gene. In Niemann-Pick patients, *SMPD1* gene defects are reported to be associated with a severe reduction in plasma high-density lipoprotein (HDL) cholesterol. Two common coding polymorphisms in the *SMPD1* gene, the G1522A (G508R) and a hexanucleotide repeat sequence within the signal peptide region, were investigated in 118 unrelated subjects of French Canadian descent with low plasma levels of HDL-cholesterol (< 5th percentile for age and gender-matched subjects). Control subjects (n=230) had an HDL-cholesterol level > the 25th percentile. For G1522A the frequency of the G and A alleles were 75.2% and 24.8% respectively in controls, compared to 78.6% and 21.4% in subjects with low HDL-cholesterol (p=0.317). The frequency of 6 and 7 hexanucleotide repeats was 46.2% and 46.6% respectively in controls, compared to 45.6% and 49.1% in subjects with low HDL-cholesterol (p=0.619). Ten different haplotypes were observed in cases and controls. Overall haplotype frequencies in cases and controls were not significantly different. These results suggest that the two common coding variants at the *SMPD1* gene locus are not associated with low HDL-cholesterol levels in the French Canadian population.

4.2 Introduction

A low plasma level of high-density lipoprotein (HDL) cholesterol is defined as a cardiovascular risk factor and is part of the assessment of global cardiovascular risk stratification[206]. Therapeutic goals set for the prevention of cardiovascular disease include targets for low density lipoprotein (LDL) cholesterol, non-HDL-cholesterol[25], and the total cholesterol to HDL-cholesterol ratio[207]. However, a goal for an absolute HDL-cholesterol value is still a matter of controversy as the current therapeutic approaches are limited in their ability to raise HDL-cholesterol[194, 208]. In the majority of cases, a low HDL-cholesterol is secondary to increased hepatic secretion of apolipoprotein B–containing lipoproteins and triglycerides[208]. Some cases of low HDL-cholesterol are due to genetic defects in HDL–associated protein apolipoprotein AI, modifying enzymes (hepatic lipase, lipoprotein lipase, cholesteryl ester transfer protein, lecithin:cholesterol acyl transferase) and the ATP binding cassette A1 transporter [209].

Niemann-Pick disease type A and B is caused by a deficiency of the enzyme acid sphingomyelinase coded by *SMPD1* gene. *SMPD1* gene defects are reported to be associated with a severe reduction in plasma HDL-cholesterol[210]. In the search for genes causing disorders of HDL-cholesterol, we examined extended (3 or more generations) kindred of French Canadian descent to identify Mendelian traits. Using this approach, we have previously reported that compound heterozygosity in the *SMPD1* gene is associated with decreased activity of acid sphingomyelinase and low HDL-cholesterol[189]. Furthermore, a

decreased activity of lysosomal and secreted acid sphingomyelinase is believed to cause low HDL-cholesterol in part by decreased activation of lecithin:cholesterol acyltransferase (LCAT) and impaired formation of cholesteryl ester–enriched HDL particles[211]. The *SMPD1* gene is located on chromosome 11p15.1-p15.4, is comprised of 6 exons, and encodes a cDNA of 2.5 kb. The acid sphingomyelinase protein (http://genome.ucsc.edu/ ; Genome Browser: NP_000534[212]) consists of 631 amino acids and exists in at least three protein isoforms. The protein also contains a signal peptide, a saposin domain that is common in lysosome–targeted proteins, and a large metallophosphoesterase domain. Although rare mutations in the *SMPD1* gene can impair the function of acid sphingomyelinase and result in Niemann-Pick disease type A or B, as well as in low HDL-cholesterol, it is not known whether common amino acid change variants in *SMPD1* can modulate HDL-cholesterol levels within a population. The objective of this study was to investigate, in a population selected for HDL-cholesterol levels, associations between known common amino acid variants of the *SMPD1* gene and low plasma levels of HDL-cholesterol in subjects of French Canadian descent. We focused on frequent (> 5% in the general population) polymorphisms that affect the coding sequence of SMPD1.

4.3 Methods

4.3.1 Subject characteristics

A total of 348 unrelated subjects of French Canadian origin (118 with low HDL-cholesterol levels and 230 control subjects) were examined at the McGill University Health Centre. Low HDL-cholesterol levels were defined as those less than the 5th percentile (age and gender-matched), based on the Lipid Research Clinics Population Studies Data Book[213]. Subjects with low HDL-cholesterol had no known cause of HDL deficiency (severe hypertriglyceridemia defined as plasma triglycerides > 10 mmol/L, cellular phospholipid or cholesterol efflux defect or previously known mutations in genes associated with HDL deficiency). The control group was of same origin and chosen based on HDL-cholesterol levels > 25th percentile, matched for age and gender. Demographic and clinical information, medications, blood pressure, and lipoprotein profiles were determined on all participating subjects. Hypertension was defined as a blood pressure ≥130/85 mmHg. Coronary artery disease (CAD) was present when angiographically documented or patient had a past history of acute myocardial infarction. Consent was obtained for the plasma sampling and DNA isolation. The research protocol was reviewed and approved by the Research Ethics Board of the McGill University Health Centre (REB No. BMA 05-006).

4.3.2 Measurement of plasma lipids and lipoprotein

The lipid lowering agents were withdrawn in all study subjects for at least four weeks before measurement of the lipid profile. Insulin and oral

hypoglycemic agents were maintained in diabetic patients. Plasma was isolated in all study subjects, after a 12-hour fast, in EDTA-containing tubes. Lipids and lipoproteins were measured using standardized techniques and the LDL-cholesterol was calculated according to the Friedewald formula, unless triglyceride levels were >4.5 mmol/L[214, 215].

4.3.3 DNA analysis

DNA was isolated from the buffy coat obtained after centrifugation of whole blood. Two previously reported common polymorphisms of the *SMPD1* gene in Niemann-Pick disease type A and B[216, 217] were examined. The G→A substitution at position c.1522 located in exon 6 of the *SMPD1* gene, predicting a substitution of arginine (R) for a glycine (G) at residue 508 (G508R) (rs1050239) was detected by polymerase chain reaction followed by digestion with the restriction enzyme *MspI* (New England Biolabs, MA, USA). The hexanucleotide repeat polymorphism at the start position c.103[212] (Genomic position 6368507 in http://genome.ucsc.edu/) was located in exon 1 of the *SMPD1* gene and detected using the sense primer 5'-GTCAGCCGACTACAGAGAAG-3' and the antisense primer 5'-GGCATCTACAATCCATCACT-3'. The antisense primer was radiolabeled at the 5' end with T4 polynucleotide kinase and [γ-^{32}P] ATP (PerkinElmer, MA, USA) by standard procedures. Polymerase chain reaction products were resolved on a 6% polyacrylamide denaturating gel.

4.3.4 Data analysis

The data was analyzed by examining allele frequencies in subjects with low HDL-cholesterol versus controls. The DeFinetti Program (http://ihg.gsf.de/cgi-bin/hw/hwa1.pl) was employed to test the deviation from Hardy-Weinberg equilibrium and also to compare the frequency of the SNPs between cases and controls. For the G1522A SNP, HDL-cholesterol levels were compared between homozygotes of the common allele and both heterozygotes and homozygotes for the rare allele pooled together. Deviation from Hardy-Weinberg equilibrium for the hexanucleotide repeat polymorphism was tested by PEDSTATS version 0.6.5. The CLUMP program version 2.3 was used to assess the significance of the same marker between cases and controls, by using 1000 simulations in a Monte Carlo approach[218].

Haplotype frequencies (containing both polymorphisms) were estimated for cases and controls, using PHASE Ver.2.02[219].

Power calculations using the Genetic Power Calculator (http://pngu.mgh.harvard.edu/~purcell/gpc/qcc.html) demonstrated that we would have >80% power with our given sample size (controls=230, cases=118) to detect a SNP that accounts for 0.02 or more of the variance in HDL-cholesterol. This assumes that we are directly testing a causitive variant with an allele frequency of 23% with a type I error rate α=0.05.

Statistical analyses were performed with the SAS package version 8 (SAS Institute Inc, NC, USA) and SigmaStat version 2.0 (Jandel Corporation, San Rafael, CA, USA). A χ^2 analysis (GraphPad InStat, CA, USA) was performed

with respect to allele frequencies in each of the HDL-cholesterol groups. Age, body mass index (BMI), and all lipid parameters in both groups were treated as continuous variables. Comparisons were made through generalized linear model procedures (Proc GLM) followed by Duncan's post hoc test. All *p*-values <0.05 were considered significant.

4.4 Results

We analyzed a total of 348 subjects from a pool of control subjects and patients with premature CAD. The selection criterion was an HDL-cholesterol <5th percentile for cases (n=118), and an HDL-cholesterol >25th percentile for controls (n=230). Mean ages were 50±10 and 50±9 years for control and case groups, respectively. Additional demographic and biochemical characteristics are shown in **Table 1**. Subjects with a low HDL-cholesterol had a higher BMI, were more likely to have type II diabetes, hypertension, CAD and a family history of CAD. These correlates of low HDL-cholesterol have been previously well established[18]. The low HDL-cholesterol group had an HDL-cholesterol of 0.67±0.13 mmol/L and the control group had a mean HDL-cholesterol of 1.35±0.33 (p<0.001). Plasma triglyceride levels were higher in the low HDL-cholesterol group than in the controls (3.95±3.35 mmol/L vs. 1.63±0.89 mmol/L, p<0.001).

We examined two polymorphisms at the *SMPD1* gene locus: G1522A (G508R) and a hexanucleotide repeat sequence CTGG (TC)(GT). From our 348 subjects analyzed, 230 controls and 117 cases were successfully genotyped. The

prevalence of the G508R variant in cases and controls is shown in **Table 2**. The presence of the G allele was seen in 75.2% of controls and 78.6% of subjects with low HDL-cholesterol (OR=0.82; *p*=0.317). The G allele has been reported at a frequency of 0.85 in an European population [National Center for Biotechnology Information: http://www.ncbi.nlm.nih.gov; SNP database #rs1050239, NCBI assay ID #ss24058527 from PERLEGEN for European panel]. We also separately analyzed the association of this variant between patients with and without CAD and no significant difference was confirmed (*p*=0.06, data not shown). Genotype frequencies for the GG, GA and AA classes did not differ between subjects with low HDL-cholesterol and controls **(Table 2)**. We examined the association of age, gender, BMI, diabetes mellitus, hypertension, plasma triglycerides, CAD, familial history of CAD, total plasma cholesterol, plasma LDL-cholesterol, between subjects with the GG genotype and subjects with either the AG or AA genotypes in cases and controls **(Table 3)**. We found significant associations between the genotypic classes with familial history of CAD (*p*=0.0003), total cholesterol (*p*=0.03) and LDL-cholesterol (*p*=0.02) in controls. In addition, we found significant associations between the genotypic classes with total cholesterol (*p*=0.009) in the low HDL-cholesterol subjects. We did not find any significant difference in the prevalence of the GG or AG + AA genotypes with the presence of diabetes. The analysis was also carried out separately between each genotypic class (GG, AG and AA), and the results were similar as those presented in **Table 3**.

The second polymorphism consisted of a unique hexanucleotide sequence CTGG(TC)(GT) located within the signal peptide region of the acid sphingomyelinase (corresponding to the hydrophobic sequence LVLALALALALA. The genotype distribution of the hexamer polymorphism was examined and the most frequent allele was the 6 and 7 repeats (respectively 46% and 47% of the control group) **(Table 4)**. We identified 9 genotypes in our study population with the most prevalent being 6/7, 7/7 and 6/6. There was no significant difference in the genotype or allele frequencies between low HDL-cholesterol subjects and controls. We used the CLUMP program to confirm these results with a *p*-value of 0.6 after a 1000-simulation analysis. We examined the most frequent genotypes with respect to age, gender, BMI, diabetes mellitus, CAD, family history of CAD, triglycerides, total cholesterol and LDL-cholesterol levels and we found significant differences between the subgroups of subjects with 6/6, 6/7, 7/7 for hypertension (p=0.04) and triglycerides (p=0.005) in low HDL-cholesterol subjects only **(Table 5)**.

We used the PHASE program to reconstruct haplotypes in cases and controls. Substantial LD was observed as the A allele of G1522A was seen almost exclusively with the hexanucleatide repeat of "6" **(Table 6)**. Overall haplotype frequencies between cases and controls were not significantly different (p=0.5) **(Table 6)**.

4.5 Discussion

The present report suggests that common genetic variability at the *SMPD1* gene locus does not contribute significantly to HDL-cholesterol levels in a French Canadian population. Rare mutations at the *SMPD1* gene can cause Niemann-Pick disease type A or B, which can differ in degrees of neurological impairment. Mutations for both types A and B are distributed throughout the *SMPD1* gene and the structure-function relationship between mutations and disease states is not fully understood. We and others have previously reported that patients with Niemann-Pick disease type A/B have low plasma levels of HDL-cholesterol[189, 210, 220]. More recently, we have shown that cellular cholesterol processing is abnormal in fibroblasts with *SMPD1* mutations[211]. Despite an abnormal lysosomal transport of cholesterol and sphingomyelin, cellular cholesterol efflux onto apolipoprotein AI does not appear to be the rate-limiting step in generating nascent HDL particles. Instead, our data suggests that abnormal composition of nascent HDL particles leads to abnormal LCAT activity and decreased cholesterol esterification when the protein product of the *SMPD1* gene, acid sphingomyelinase, is defective[211]. It has been previously reported that reconstituted HDL particles using proteoliposomes with an increasing ratio of sphingomyelin to phosphatidylcholine inhibits LCAT activity and cholesteryl ester formation[221-223]. This leads to an inability of HDL particles to mature into spherical HDL_3 particles. In turn, current evidence points to an increased catabolism of these nascent, cholesteryl ester-poor HDL particles by the kidney[224].

In a previous report, we have shown that rare mutations of the *SMPD1* gene leads to reduced activity of acid sphingomyelinase and is associated with a low HDL-cholesterol. In addition, the mutations segregate within families with a gene dosage effect. This gene-dosage effect was shown in HDL-cholesterol levels in homozygotes and compound heterozygotes[189]. Here, we have found that the prevalence of the G1522A substitution (G508R) was not significantly different in subjects with a low HDL-cholesterol, compared with controls. We used the polyphen program (http://tux.embl-heidelberg.de/ramensky/) to determine the predicted impact of individual variants on *SMPD1* function and this variant was predicted to be benign.

Moreover, the presence of the 6 and 7 hexanucleotide repeats as well as the 10 different haplotypes in cases and controls were not significantly different. In a previous report, the 5 separate alleles, corresponding to 9, 7, 6, 5 and 4 hexanucleotide repeats were unrelated to Niemann-Pick disease[217]. Corresponding allele frequencies of 0.5%, 12.4%, 50.4%, 34.9% and 1.8% were found in that study, generating 9 different genotypes[6].

Some significant associations were found between *SMPD1* genotypic classes and characteristics of the cases and control subjects. For example, in the control group, carriers of the 1522A had a significant increase in family history of CAD, plasma LDL-cholesterol and total cholesterol levels. However, the 1522A allele was associated with lower total plasma cholesterol concentrations in the cases. Given the lack of consistency of these results between the control and low HDL-cholesterol groups, and the relative statistical weakness of these

associations (not significant or only marginally significant when Bonferroni corrections are applied for multiple testing), the clinical relevance of these findings is uncertain and are probably the result of multiple statistical tests.

This study is limited by the relatively small number of subjects (n=348). However, we did have greater than 80% power to detect a genetic variant that accounts for as little as 2% of the variance of HDL-cholesterol. In addition, we used arbitrary cut-points of an HDL-cholesterol $<5^{th}$ percentile and $>25^{th}$ percentile for cases and controls, respectively. These data should still be confirmed in a large-scale study.

4.6 Conclusion

Our data suggest that while rare mutations at the *SMPD1* locus can cause Niemann-Pick disease types A and B and the concomitant low HDL-cholesterol, the two common coding non-synonymous variants that we examined at this locus do not appear to influence HDL-cholesterol levels to any great extent. Forty-five mutations in *SMPD1* gene causing different forms of Niemann-Pick disease type A and B have been described[225]. Since the incidence of Niemann-Pick disease type B is difficult to estimate due to the lack of enzyme testing in clinic, variability in symptoms and the lack of knowledge of Niemann-Pick disease type B by treating physicians, many patients remain undiagnosed[225]. It remains to be determined if variations in the *SMPD1* gene, affecting the activity of acid sphingomyelinase, might contribute to the modulation of HDL-cholesterol levels in the general population. This study did not examine rare mutations and thus

carrier status for Niemann-Pick disease type B was not ruled out in either group. However, Niemann-Pick disease type B should not be common enough to influence our findings.

4.7 Acknowledgments

This study was supported by the Canadian Institute of Health Research [CIHR grant MOP-62834 (JG) and CIHR grant MOP-74703 (MM)] and the Heart and Stroke Foundation of Québec (MM). MM and J. Engert are research scholars from the Fonds de la Recherche en Santé du Québec. JG holds the McGill University-Novartis Chair in Cardiology. The authors thank their research nurses, Mrs. Colette Rondeau and Natalie Bellavance, for their work with the individuals who participated in this study, and Ms. Carole Dangoisse, summer student, for her technical assistance. The authors would like to thank Ron Do, graduate student, for helpful discussions.

4.8 Tables

Table 1. Baseline characteristics of low HDL-C and control subjects.

	Control subjects (n=230)	Low HDL-C subjects (n=118)	t test *p**
Age (y)	50 ± 9	50 ± 10	0.93
Gender (M/F)	150/80	85/33	0.20
BMI (kg/m^2)	26.3 ± 4.5	28.1 ± 4.9	**0.001**
DM (%)	10.4	21.2	**0.01**
HTN (%)	20.4	45.8	**<0.001**
CAD (%)	49.8	72.3	**<0.001**
FH of CAD (%)	61.3	72.6	**0.04**
TG (mmol/L)	1.63 ± 0.89	3.95 ± 3.35	**<0.001**
T Chol (mmol/L)	5.73 ± 1.41	5.98 ± 2.20	0.26
HDL-C (mmol/L)	1.35 ± 0.33	0.67 ± 0.13	**<0.001**
LDL-C (mmol/L)	3.70 ± 1.33	3.50 ± 1.61	0.27

BMI, body mass index; DM, diabetes mellitus; HTN, hypertension; CAD, coronary artery disease; FH, familial history; TG, plasma triglycerides; T Chol, plasma total cholesterol; HDL-C, plasma high-density lipoprotein cholesterol; LDL-C, plasma low density lipoprotein cholesterol.

* significant *p*-value bolded.

Table 2. Genotype distribution and allele frequency of the G1522A variant in the SMPD1 gene in control and low HDL-cholesterol subjects.

Genotype or allele	Control subjects (n)	Low HDL-cholesterol subjects (n)	χ^2,two-sided *p*-value and O.R.
GG	56.1% (129)	64.1% (75)	0.152[†]
AG	38.3% (88)	29.1% (34)	0.72 (0.45-1.13) [‡]
AA	5.6% (13)	6.8% (8)	
G	75.2%	78.6%	0.317
A	24.8%	21.4%	0.82 (0.57-1.20) [‡]

O.R., odd ratio.

[†] AG and AA genotypes where pooled for statistical analysis.

[‡] 95% confidence interval for O.R.

Table 3. Comparison of biochemical data of the low HDL-C and control subjects between different groups of G1522A genotypes.

	Control subjects		*t test*	*Low HDL-C subjects*		*t test*
	GG	**AG + AA**	***p****	**GG**	**AG + AA**	***p****
n	129	101		75	42	
Age (y)	50 ± 10	50 ± 8	0.5	49 ± 9	51 ± 11	0.4
Gender (M/F)	81/48	69/32	0.3	52/23	32/10	0.4
BMI (kg/m^2)	26.4 ± 4.6	26.3 ± 4.4	0.9	28.3 ± 5.2	27.7 ± 4.5	0.5
DM (%)	10.5	10.9	0.9	20.3	19.0	0.9
HTN (%)	19.3	22.8	0.5	51.3	33.3	0.06
CAD (%)	45.9	56.0	0.1	75.4	65.8	0.3
FH of CAD (%)	51.7	75.8	**0.0003**	75.0	67.5	0.4
TG (mmol/L)	1.60 ± 0.89	1.70 ± 0.89	0.4	4.34 ± 3.89	3.19 ± 1.90	0.07
TChol (mmol/L)	5.55 ± 1.19	5.96 ± 1.62	**0.03**	6.33 ± 2.39	5.25 ± 1.56	**0.009**
HDL-C (mmol/L)	1.34 ± 0.31	1.33 ± 0.33	0.8	0.67 ± 0.13	0.67 ± 0.13	0.9
LDL-C (mmol/L)	3.51 ± 1.10	3.93 ± 1.54	**0.02**	3.71 ± 1.74	3.17 ± 1.34	0.09

BMI, body mass index; DM, diabetes mellitus; HTN, hypertension; CAD, coronary artery disease; FH, familial history; TG, plasma triglycerides; T Chol, total plasma cholesterol; HDL-C, plasma high-density lipoprotein cholesterol; LDL-C, plasma low density lipoprotein cholesterol.

* significant *p*-value bolded.

Table 4. Genotype distribution and allele frequency of the hexanucleotide repeat polymorphism in the SMPD1 gene in low HDL-C and control subjects.

Genotype or allele	Control subjects (%)	Low HDL-C subjects (%)	*P*
5/5	0.4	0	0.527
5/6	6.3	2.6	
5/7	5.4	7.0	
6/6	20.3	25.5	
6/7	45.6	37.7	
7/7	20.3	26.3	
7/8	0.4	0	
7/9	0.9	0.9	
7/10	0.4	0	
5	6.3	4.8	0.619
6	46.2	45.6	
7	46.6	49.1	
8	0.2	0	
9	0.5	0.5	
10	0.2	0	

Table 5. Comparison of biochemical data of the low HDL-C and control subjects between the most prevalent genotypes of the hexanucleotide repeat polymorphism in the SMPD1 gene.

	Control subjects (n=230)			t test	Low HDL-C subjects (n=118)			t test
	6/6	**6/7**	**7/7**	***p***	**6/6**	**6/7**	**7/7**	***p****
n	45	101	45		29	43	30	
Age (y)	50 ± 8	50 ± 9	49 ± 9	0.79	48 ± 10	51 ± 10	49 ± 10	0.51
Gender (M/F)	30/15	68/33	29/16	0.94	22/7	29/14	20/10	0.69
BMI (kg/m^2)	26.1 ± 5.1	26.4 ± 4.0	26.0 ± 4.1	0.86	27.9 ± 4.2	27.8 ± 5.1	29.2 ± 5.6	0.45
DM (%)	15.0	55.0	30.0	0.57	19.0	42.9	38.1	0.47
HTN (%)	23.1	61.5	15.4	0.35	22.2	35.6	42.2†	**0.04**
CAD (%)	26.1	51.1	22.8	0.77	23.9	45.1	31.0	0.16
FH of CAD (%)	28.3	53.3	18.3	0.05	23.6	43.1	33.3	0.15
TG (mmol/L)	1.60 ± 0.68	1.56 ± 0.85	1.68 ± 0.88	0.71	2.73 ± 1.38	4.94 ± 3.65†	3.57 ± 2.58	**0.005**
T Chol (mmol/L)	6.09 ± 1.45	5.61 ± 1.52	5.82 ± 1.32	0.19	5.56 ± 2.21	6.12 ± 1.77	6.03 ± 2.06	0.48
HDL-C (mmol/L)	1.36 ± 0.34	1.33 ± 0.31	1.31 ± 0.28	0.80	0.66 ± 0.11	0.67 ± 0.13	0.69 ± 0.10	0.55
LDL-C (mmol/L)	3.96 ± 1.37	3.60 ± 1.34	3.79 ± 1.39	0.32	3.60 ± 2.08	3.38 ± 1.34	3.68 ± 1.62	0.74

BMI, body mass index; DM, diabetes mellitus; HTN, hypertension; CAD, coronary artery disease; FH, familial history; TG, plasma triglycerides; T Chol, plasma cholesterol; HDL-C, plasma high-density lipoprotein cholesterol; LDL-C, plasma low density lipoprotein cholesterol.

†significantly different from genotype 6/6.

* significant *p* value bolded.

Table 6. Estimation of the haplotype frequency distribution of the G1522A variant and hexanucleotide repeat polymorphism in the SMPD1 gene in low HDL-C and control subjects.

Haplotype	Total (%)	Controls subjects (%)	Low HDL-C subjects (%)
5 G	5.5	5.8	4.7
5 A	0.4	0.5	0.1
6 G	23.3	22.7	24.5
6 A	22.4	23.1	21.1
7 G	46.9	45.8	48.9
7 A	0.8	1.2	0.2
8 G	0.2	0.2	0.0
9 G	0.4	0.4	0.4
10 G	0.2	0.2	0.0

p-value for testing H_0: cases ~ controls=0.5.

Preface to chapter 5

To identify novel associated loci with low HDL-C, we performed a genome scan on French Canadian families. The genome scan was executed using 485 autosomal microsatellite markers covering the human genome with an average marker density of about 6 cM. Using parametric linkage analysis, a locus was identified on chromosome 4q31.21, which was subsequently fine-mapped. This finding has been published in the journal of Arteriosclerosis, Thrombosis, and Vascular Biology: ” *Evidence for a gene influencing high-density lipoprotein cholesterol on chromosome 4q31.2 Arterioscler Thromb Vasc Biol. 2006 Feb;26(2):392-7.*” And will be presented in this chapter.

CHAPTER 5

EVIDENCE FOR A GENE INFLUENCING HDL-C ON CHROMOSOME 4q31.21

Zari Dastani[1], Leigh Quiogue[2], Christopher Plaisier[2], James C. Engert[1], Michel Marcil[3], Jacques Genest[1] and Päivi Pajukanta[2]

From the [1]Departments of Medicine and Human Genetics, McGill University, Montreal, Canada, the [2]Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, USA and the [3]Division of Cardiology, McGill University, Montreal, Canada.

Key words: high-density lipoprotein cholesterol, family study, complex trait, coronary heart disease, gene identification.

Artherioscler Thromb Vasc Biol **2006, 26(2):392-7**

5.1 Abstract

Objective: A low level of plasma high-density lipoprotein cholesterol (HDL-C) is a major risk factor for coronary atherosclerosis. To identify novel genes regulating plasma HDL-C levels, we investigated 13 multigenerational French Canadian families with on average 12 affected individuals per family for genome wide signals which we subsequently fine mapped. **Methods and Results:** We genotyped a total of 362 individuals, including 151 affected subjects for 485 autosomal microsatellite markers. In parametric two-point linkage analyses, the highest two-point lod score of 4.6 was observed with marker D4S424 on chromosome 4q31.21 (at ~142 Mb). The multipoint analysis of this region resulted in a lod score of 3.8 and a lod minus 1 region of 12.2 cM, containing 40 known genes. The results were obtained by allowing for genetic heterogeneity among these extended pedigrees, and approximately 50% of families were linked to this region with the highest single-pedigree lod score being 3.6. We further restricted the linked region from 12.2 cM to 2.9 cM (2.37 Mb) by genotyping 15 additional markers in the three families with the highest LOD-scores. We sequenced four genes with a likely role in lipid metabolism as well as two genes residing directly under the linkage peak, but found no evidence for a causative variant. None of the genes residing in the significantly restricted 2.37-Mb region has previously been associated with HDL-C metabolism. **Conclusion:** This study provides significant evidence for a gene influencing HDL-C on chromosome 4q31.21.

5.2 Introduction

Plasma high-density lipoprotein cholesterol (HDL-C) is a quantitative trait where complex interactions of several genes and environmental factors determine the plasma levels in the general population. A decreased level of plasma HDL-C is a major risk factor for coronary atherosclerosis. Furthermore, aggregation of other cardiovascular risk factors including features of the metabolic syndrome is typical in patients with low HDL-C[202, 208]. The genetic component, estimated to determine approximately 50% of plasma HDL-C variability, has been heavily investigated to identify genes regulating plasma HDL-C levels[226]. Previously, complex quantitative traits have been suggested to be caused by common sequence variants each with a small to moderate phenotypic effect[227, 228]. However, most recent studies suggest that both rare and common variants confer the susceptibility to low plasma levels of HDL-C in the general population[64, 105], although the extent to which each group contributes to this susceptibility is currently not known.

HDL has a key role in reverse cholesterol transport (RCT), mobilizing cholesterol from the peripheral tissues to liver. This mechanism contributes to the cardioprotective effect of HDL. In RCT, the ATP-binding cassette transporter A1 (ABCA1) protein controls efflux of intracellular cholesterol to lipid-poor apolipoprotein A-I, the major apolipoprotein of HDL. Interestingly, ABCA1 mutations were discovered to cause Tangier disease, a rare recessive HDL deficiency[101-103], and sequence variants in ABCA1 also appear to contribute to variations in plasma HDL-C levels in the general population[64, 105]. Recent

evidence further shows that two other ABC transporters, ABCG1 and ABCG4, mediate the efflux of cellular cholesterol to smaller and larger HDL subclasses, HDL2 and HDL3 that constitute the bulk of the plasma HDL[229]. Genes implicated in rare Mendelian forms of HDL deficiency include apolipoprotein AI (APOA1) and lecithin cholesterol acyltransferase (LCAT)[134, 230]; and genes implicated in candidate gene studies include apolipoproteins and the enzymes remodeling HDL[231, 232]. The previous genome wide scans have also identified six chromosomal loci for HDL-C: a locus on chromosome 8q in Mexican Americans and Finns[92, 176]; loci on 15q and 9p in Mexican Americans[92, 233]; a locus on 11q23 in 105 Utah CHD families[234]; and loci on 16q and 20q in Finns[175, 176]. However, DNA sequence variants contributing to variation in plasma levels of HDL-C in the general population are largely unknown, especially regarding the prevalence of variants with major effects.

We performed a genome scan in 13 extended multigenerational French Canadian families, each exhibiting affected individuals in consecutive generations. Importantly, families with a defect in cellular cholesterol efflux or previously known ABCA1 mutations were excluded. This was done to minimize confounding factors and to identify genes with novel functions. We also anticipated that extended families with on average 12 HDL-C affected individuals per family could provide a powerful approach for the identification of genes contributing to the complex HDL-C trait, especially when allowing for heterogeneity and thus, assuming that the same gene is causative within a multigenerational family but not necessarily between all families.

5.3 Methods

5.3.1 Subjects

A total of 13 multigenerational French Canadian families consisting of 362 genotyped family members were collected in the Cardiovascular Genetics Laboratory, McGill University Health Centre, Royal Victoria Hospital, Montreal, Canada. All subjects provided separate informed consent forms for plasma and DNA sampling, isolation, and storage. The research protocol was approved by the Research Ethics Board of the McGill University Health Centre.

The selection criterion for probands was an HDL-C level < the 5^{th} age/sex-specific population percentile based on the Lipid Research Clinics Population Studies Data Book, as previously described[235]. Exclusion criteria for the probands were as follows: severe hypertriglyceridemia (>10 mmol/l), cellular cholesterol efflux or phospholipid efflux defect in skin fibroblasts, or previously known ABCA1 mutations. All available living relatives were invited to participate in the study. Family members were sampled after a 12-hour fast and discontinuation of lipid modifying medications for at least 4 weeks. Demographic and clinical information, medications, and lipoprotein profiles were determined on all subjects. In the 13 extended families, there were 151 subjects affected using the HDL-C $\leq 10^{th}$ age/sex specific population percentile (see **Table 1** for the phenotypic characteristics of the family members).

5.3.2 Biochemical measurements

Lipids and lipoproteins were measured using standardized techniques as described previously[9, 22]. Low-density lipoprotein cholesterol (LDL-C) was calculated according to the Friedewald formula [LDL-C (mmol/L) = Total cholesterol – (triglycerides /2.19 + HDL-C)], unless the triglycerides were >4.5 mmol/L. In this case, ultracentrifugation of plasma was used and lipoprotein lipid concentration measured directly. Cellular cholesterol efflux and phospholipid efflux assays have been carried out on the probands as described previously[101, 203, 236].

5.3.3 Genotyping and sequencing

The genome scan was executed using 485 autosomal microsatellite markers covering the human genome with an average marker density of about 6 cM. DNA from 362 individuals in the 13 French Canadian HDL-C families was genotyped by deCODE Genetics, Reykjavik, Iceland using Applied Biosystems Automated DNA Sequencing System. The genome-wide scan marker sets by deCODE Genetics are based partly on the ABI JD Marker Linkage set and partly on the in-house designed and validated markers originally selected from the Marshfield genetic map. For fine mapping of chromosome 4q31.21, a total of 15 additional fine mapping markers with heterozygosity values >0.50 were analyzed in 155 individuals of the three families with the highest LOD-scores. Their genotyping was also performed by the deCODE Genetics using Applied Biosystems Automated DNA Sequencing System.

Sequencing of the RAB33B, member RAS oncogene family (RAB33B), TBDN100 transcriptional coactivator tubedown-100 (TBDN100), uncoupling protein 1 (mitochondrial, proton carrier) (UCP1), ATP-binding cassette, sub-family E (OABP), member 1 (ABCE1), solute carrier family 7, (cationic amino acid transporter, y+ system) member 11 (SLC7A11) and protocadherin 18 (PCDH18) genes was performed using the dye termination method and Applied Biosystems Automated DNA Sequencing System.

5.3.4 Statistical analyses

We carried out parametric two-point linkage analyses of all 485 microsatellites using the Mendel program[237] under the assumption of a recessive mode of inheritance. The Mendel program was selected because it allowed us to keep these extended families intact in the two-point analyses. For these analyses, subjects were coded as affected or unaffected based on the 10th age/sex specific percentiles for HDL-C. We performed the analyses of this complex trait by allowing for heterogeneity because ignoring locus heterogeneity significantly decreases the power to detect linkage[238]. We used a gene frequency of 8% based on an estimated ~1% disease prevalence for low HDL-C, and allowed for a 2% phenocopy rate to address the problem of incomplete penetrance.

We also conducted parametric multipoint analyses using the Location Score option of SimWalk2 (version 2.90)[239] on chromosomal regions with significant two-point lod scores (above 3.0) to further explore the linkage evidence obtained in the initial two-point screening of all markers. Assuming

again a recessive mode of inheritance, we incorporated the same parameters into the model as in the two-point linkage analyses described above. SimWalk2's Location Score option calculates lod scores both under homogeneity (alpha = 1.0) and by allowing for heterogeneity. The heterogeneity lod scores are maximized over a grid of possible alpha values. The SimWalk2 program was also used to build the gene-specific haplotypes of the identified SNPs.

5.4 Results

We performed a genome-wide scan of 13 extended multigenerational French Canadian families with 362 genotyped individuals (an average of 28 genotyped individuals per family), to identify genes regulating plasma HDL-C levels. These 13 families included 151 affected subjects, resulting in an average of 12 affected individuals per family. All of the families had 3 or more generations of which DNA is available for family members. The detailed subject and family characteristics are shown in **Tables 1** and **2**.

A total of 485 autosomal microsatellite markers were genotyped and tested for linkage using a recessive mode of inheritance. In the initial screening we performed a two-point linkage analysis for each marker on each chromosome by allowing for heterogeneity. We were able to keep all these extended families intact in the linkage analysis by using the Mendel program and thus to obtain all information about vertical transmissions in the families. The results of the two-point analyses are shown in **Figure 1**. The highest two-point heterogeneity lod score of 4.6 was observed with marker D4S424 on chromosome 4q31.21 (at ~142

Mb), and approximately 50% of families were linked to this chromosomal region (**Figure 1**). It is noteworthy that separately one of the families with 86 genotyped individuals of which 39 were affected provided a lod score of 3.6 (with a posterior probability of linkage = 0.9998) with marker D4S424 using the Mendel program. The next most significant family resulted in a two-point lod score of 1.8 (with a posterior probability of linkage = 0.9834) with marker D4S424. Peak markers for other chromosomes resulting in lod scores >1.0 are also indicated in **Figure 1**. Markers on chromosomes 2, 12, 14, 15 and 16 resulted in lod scores > 1.0 (**Figure 1**).

We performed the multipoint analyses using the Location Score option of SimWalk2 (version 2.90) only for the regions that showed significant evidence for linkage in the initial screening with the two-point linkage analysis to further explore the interesting signals of potential linkage evidence by multipoint. The multipoint analysis of the 4q31.21 region resulted in a lod score of 3.8 and a 12.2-cM lod -1 region (**Figure 2**). Separately, the same family that produced a two-point lod score of 3.6 produced a multipoint lod score of 3.8. Again approximately 50% of the families were linked to the region in the multipoint analysis. It is worth noting that the average marker density for this linked region in the genome scan was 1 marker / 9.4 cM, affecting the information content of the multipoint analysis. The description of the gene symbols and exact positions of the genes residing in the lod -1 region are shown in **Supplementary Table 1** that will be available at our web site:
(http://www.genetics.ucla.edu/labs/pajukanta/QuebecHDL/).

We fine mapped the 12.2- cM region on 4q31.21 by genotyping a total of 15 additional microsatellite markers. In the initial scan we had genotyped three markers (D4S1527, D4S424 and D4S2962) for this region. This strategy resulted in a marker density of 1 marker / 0.68 cM. We genotyped the additional 15 markers in 155 individuals of the three families with the highest LOD-scores and performed the multipoint analysis again using the Location Score option of SimWalk2 (version 2.90) with the same parameters as in the initial multipoint analysis. As a result, the lod -1 region was further restricted from 12.2 cM to 2.9 cM (~2.37 Mb) (**Figure 2**). Of the 40 previous candidate genes, 10 reside within this restricted region of 2.37 Mb. Two of these genes, PCDH18 and SLC7A11, reside directly under the peak (**Figure 2**). The appearance of two peaks in this lod -1 region is more likely to be related to the marker frequencies, i.e. the information content of the microsatellites, rather than to biological processes, because nine of the 22 genotyped microsatellite markers have a relatively low heterozygosity value between 0.5-0.7. Furthermore, there is a gap of 3.6 cM between markers D4S1586 and D4S3008, that reside in close vicinity of the second peak (**Figure 2**), where no informative microsatellites (heterozygosity value>0.5) were found using the Marshfield, DeCode or UCSC Human Genome Browser databases.

We sequenced probands of two linked families, including the proband of the family that produced the significant linkage signal alone, for the coding regions and exon-intron boundaries of all regional candidate genes with a probable role in lipid metabolism. Accordingly, the RAB33B, TBDN100, UCP1

and ABCE1 genes were sequenced. We also sequenced the PCDH18 and SLC7A11 genes that are located directly under the fine map peak (**Figure 2**). None of the 17 identified variants in these six genes (**Supplementary Table 2** that will be available at http://www.genetics.ucla.edu/labs/pajukanta/hdl/) represent a missense or nonsense variant with obvious functional consequences, although the possibility of their regulatory role cannot be excluded without actual functional analyses. We investigated these 17 variants first in five affected individuals / family. The segregation of 10 of the 17 variants with the low HDL-C trait could not be ruled out by genotyping the five affected subjects per family (for rs numbers of the 10 variants, see **Supplementary Table 2**). Therefore, we then genotyped these 10 SNPs in all family members of the particular family/families. None of the SNPs segregated with the low HDL-C trait in these extended families, nor did we find significant evidence for linkage with any of the SNPs using the Mendel program as all 10 SNPs resulted in lod scores <1.1. Finally no gene-specific haplotypes segregating with the low HDL-C trait were observed using the SimWalk2 program.

5.5 Discussion

We identified a significant signal of linkage (a lod score of 4.6) for low plasma HDL-C on chromosome 4q31.21 in French Canadian families with low HDL-C. These extended families with an average of 12 low HDL-C affected family members and an average of 28 genotyped individuals per family are very powerful to test for linkage due to the large number of potentially informative

meioses in several consecutive generations. In the linkage analyses, we allowed for heterogeneity, because plasma HDL-C is a quantitative trait likely regulated by variants of multiple genes interacting with environmental factors. We hypothesized that the underlying segregating variants conferring susceptibility to low levels of plasma HDL-C in these extended multigenerational families may well differ between the families but less likely within the families. The latter assumption is supported by the fact that these families are French Canadian, a regional population group known to exhibit a higher degree of genetic homogeneity when compared to more mixed populations[240]. Approximately 50% of families were linked to chromosome 4q31.21. After observing significant evidence for linkage in the total study sample, we investigated the families separately. In one single family with 86 genotyped individuals and 39 HDL-C affected individuals we observed a significant lod score of 3.6 for this chromosomal region on 4q31.21.

Six previous genome-wide scans have identified six chromosomal loci for HDL-C on 8q, 9p, 11q23, 15q, 16q and 20q[92, 175, 176, 233, 234]. In the current study, we observed lod scores > 1.0 for two of these 6 loci: the loci on 15q and 16q. In addition, we also detected the region on 2q previously observed for HDL-C in the Utah and Finnish families[241, 242]. However, the only significant signal in the present study was observed for 4q31.21.

To select the region for further fine mapping, we used the lod -1 criterion, a heuristic approach that can generally be considered as somewhat arbitrary when investigating complex traits, because this 4q31.21 region provided the only

significant signal in this genome scan and furthermore, when analyzing the families separately, one extended family alone resulted in a significant lod score of 3.6. Taken together these results suggest that a gene in the region may contribute significantly to the low HDL-C levels in a manner resembling a single gene disorder at least in this one family, resulting in a significant lod score separately. If the employment of the lod -1 does not result in gene identification within this region, we will extend the investigated region to lod -2 in future steps of the fine mapping.

Fine mapping of the 4q31.21 region with 15 additional microsatellites and a marker density of 1 marker / 0.68 cM allowed us to further restrict the lod -1 region from 12.2 cM, including 40 genes to 2.9 cM (~2.37 Mb), including ten genes. However, as described above we recognize that the application of the lod -1 strategy is somewhat arbitrary when investigating complex traits. Therefore, we sequenced all genes among the 40 genes that have a likely role in lipid metabolism, i.e. TBDN100, RAB33B, UCP1 and ABCE1, but found no evidence for a variant co-segregating with low-HDL-C in the two linked families examined. We also sequenced the two genes residing precisely under the fine map peak, PCDH18 and SLC7A11. The PCDH18 gene is a potential calcium-dependent cell-adhesion protein expressed in several tissues, and the SLC7A11 gene is a member of a heteromeric Na+-independent anionic amino acid transport system, highly specific for cystine and glutamate. Again, no evidence for a co-segregating variant was observed. In future studies, the 34 remaining regional genes need to be investigated. Finally, all SNPs identified in the linked families

as well as the tag SNPs, provided by the HapMap Project and capturing most of the genetic variation in this region, could be tested for association with HDL-C in a Canadian population-based case-control study sample to identify the underlying gene. By utilizing this strategy for the identification of the DNA sequence variants, we can also identify rare variants in addition to the common variants identified by the HapMap Project. This may be of importance because rare variants may have a role in individual families, and because recent studies suggest that both rare and common variants confer the susceptibility to low plasma levels of HDL-C in the general population[6, 7]. Since none of the probands in this study had a cellular cholesterol efflux defect, the functional role of the underlying gene for HDL-C in this region is unlikely to be directly linked to cellular efflux of cholesterol.

5.6 Acknowledgments

We thank the French Canadian families that are participating in the study. This work was supported by Grant MOP 15042 from the Canadian Institutes of Health Research (CIHR). JCE and MM are Chercheur-Boursiers of the Fonds de la Recherche en Sante du Quebec (FRSQ). PP is supported by the NIH grants HL-28481 and HL-70150 as well as by the American Heart Association grant 0430180N. The authors would like to thank J Faith, M Lemire, and JC Loredo-Osti for their advice and assistance.

5.7 Figures

Figure 1. Two-point linkage results of the genome wide scan of the 13 French-Canadian families with low HDL-C when allowing for heterogeneity. Alpha value and position are indicated for the peak marker on chromosomes with lod scores > 1.0

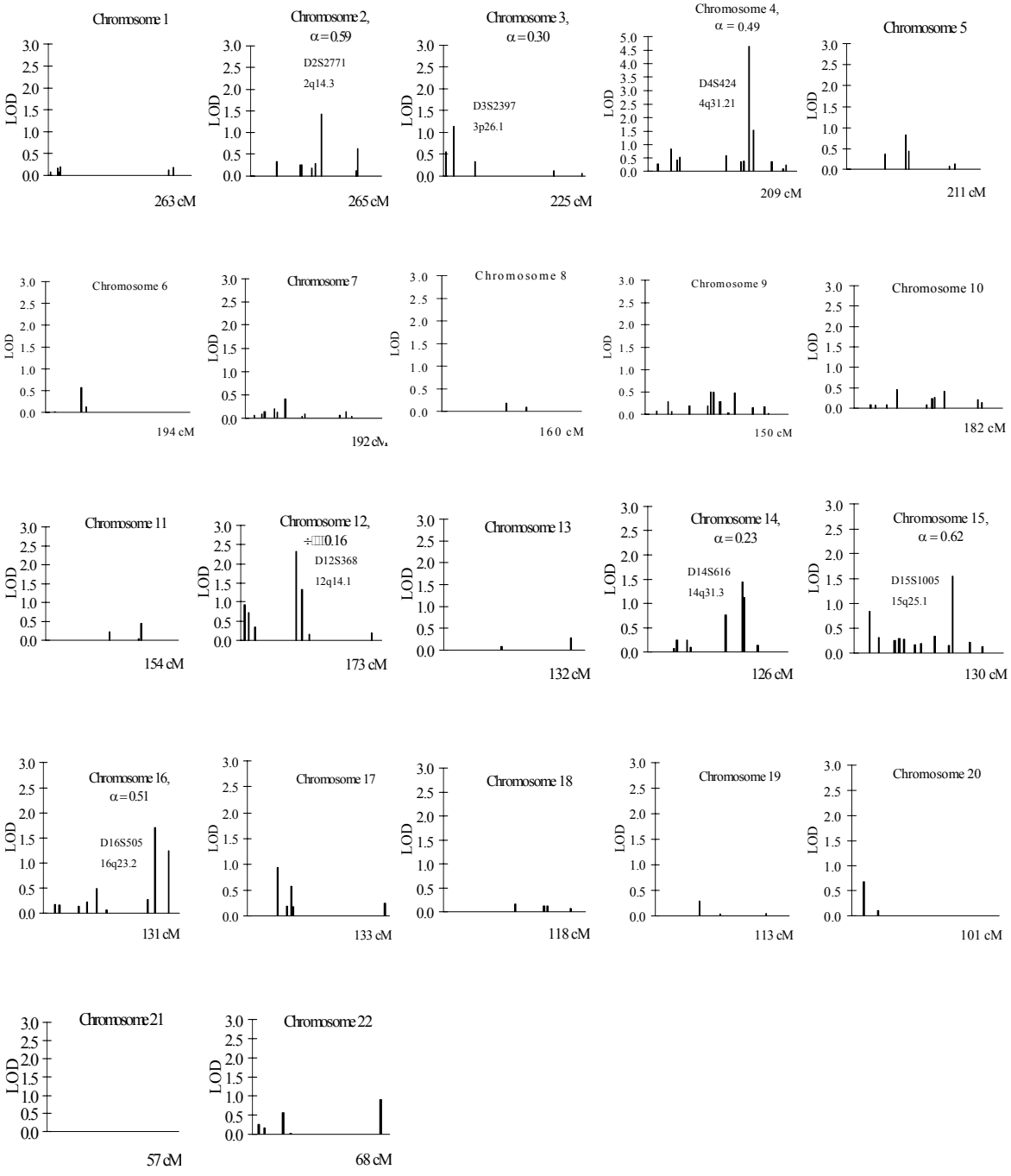
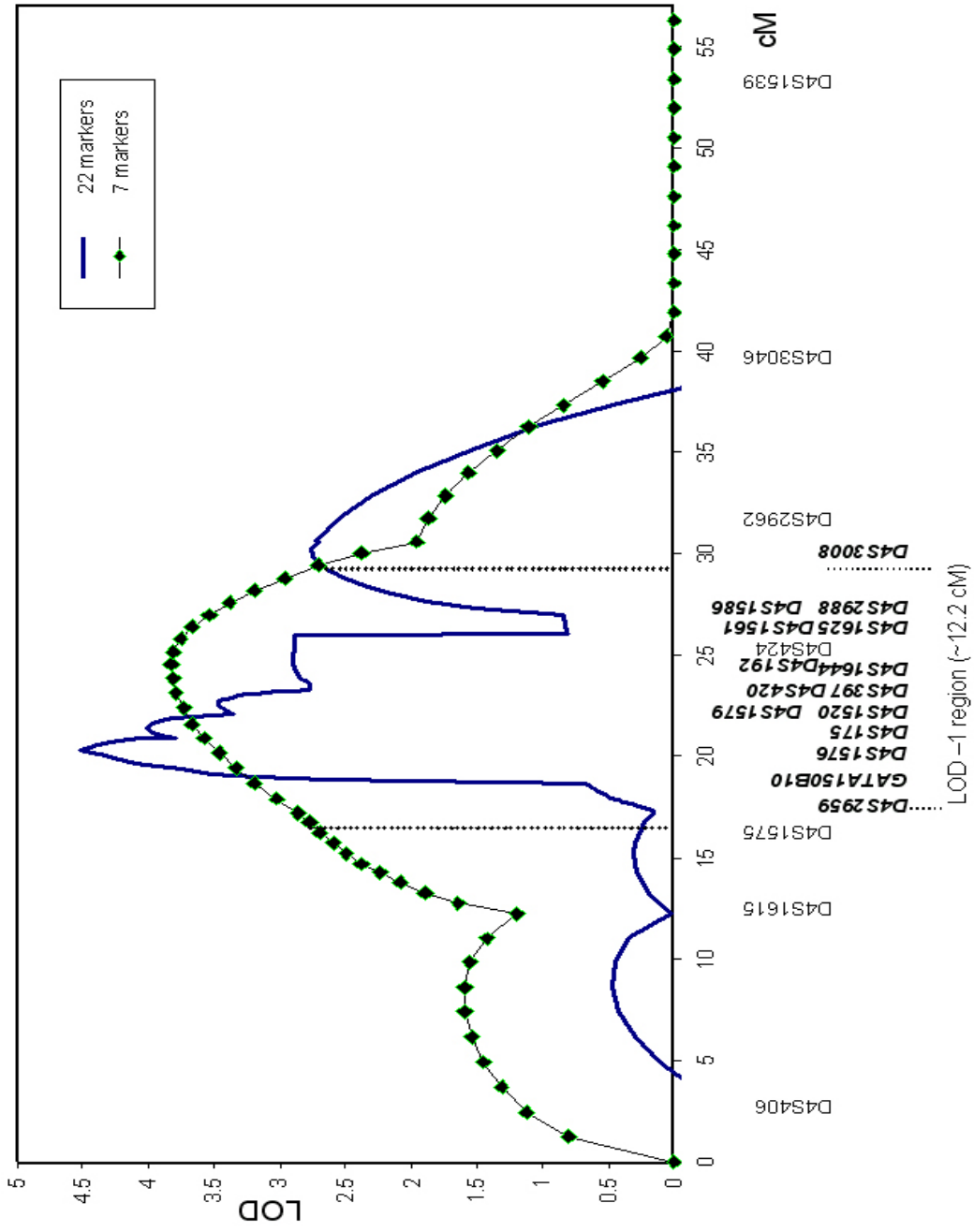


Figure 2. Multipoint analyses of the genome-wide scan and fine mapping for HDL-C on chromosome 4q31.21.

The curve of seven markers represents the multipoint analysis using all families, and the curve of 22 markers the multipoint analysis using a subset of three families. For a list of genes residing in the LOD -1 region, see **Supplementary Table 1**.



5.8 Tables

Table 1. Characteristics of the 13 French Canadian low HDL-C families.

Number of individuals with DNA available; M/F*	362; 163 M, 199 F
No. of probands; M/F	13; 11 M, 2 F
No. of families in each group / Mean no. of individuals of which DNA is available in each family type:	
2 generations	0
3 generations	3 / 7
4 generations	9 / 28
5 generations	1 / 86
Families with ≥3 generations (%)	100 %
No. of affected subjects, M/F; Mean no. of affected subjects in families is given in parenthesis:	
HDL-C ≤10th percentile	151; 72 M, 79 F (12)
No. of affected sibpairs (independent):	
HDL ≤10th percentile	66

* M indicates males; and F, females.

Table 2. Phenotypic characteristics of the 13 French Canadian families with low HDL-C shown separately in the affected individuals, unaffected individuals, spouses, and probands.

French Canadian HDL-C families		Affected* M/F		Unaffected M/F		Spouses M/F		Probands M/F	
Individuals (M/F)		72	79	91	120	25	33	11	2
Age	(years)	38.7±19.8	39.2±19.9	37.6±15.7	40.4±19.5	51.7±11.1	50.8±11.8	49.5±12.2	44.0±5.7
BMI‡	(kg/m^2)	26.8±5.3	23.8±4.7	24.8±4.4	23.2±4.7	27.7±3.6	25.9±5.3	26.7±4.9	25.3±5.4
TG§	(mmol/l)	2.8±2.3	2.1±2.1	1.5±0.9	1.4±0.7	1.7±1.0	1.7±1.1	4.3±3.5	3.7±0.1
HDL-C	(mmol/l)	0.7±0.2	0.9±0.2	1.1±0.2	1.3±0.3	1.1±0.3	1.4±0.4	0.6±0.1	0.6±0.1

* The affection status is determined as HDL-C $\leq$ 10th age-sex specific population percentile. † M indicates males; F, females; ‡ BMI, body mass index; and § TG triglycerides.

Supplementary Table 1. Gene symbol, description, and position of the genes found in the LOD -1 region (40 genes) and surrounding regions (5 genes) on chromosome 4q31.21
(see Figure 2). The positions are based on the UCSC Human Genome Browser (build 35) (http://genome.ucsc.edu).

	Gene symbol	Description	Position (bp)
1.	PDZK6	PDZ domain containing 6	128,911,724 - 128,995,535
2.	APG-1	Heat shock protein (hsp110 family)	129,061,057 - 129,112,127
3.	PLK4	Polo-like kinase 4	129,159,705 - 129,177,532
4.	PGRMC2	Progesterone receptor membrane component 2	129,548,903 - 129,566,571
5.	PHF17	PHD finger protein 17	130,088,574 - 130,152,032
6.	PCDH10	Protocadherin 10	134,428,074 - 134,432,009
7.	PCDH18	Protocadherin 18	138,798,464 - 138,811,234
8.	SLC7A11	Solute carrier family 7, (cationic amino acid transporter, y+ system) member 11	139,442,854 - 139,520,963
9.	CCRN4L	CCR4 carbon catabolite repression 4-like	140,294,547 - 140,324,511
10.	ELF2	E74-like factor 2	140,336,475 - 140,363,071
11.	OSAP	Ovary-specific acidic protein	140,544,921 - 140,559,097
12.	NDUFC1	NADH dehydrogenase (ubiquinone) 1, subcomplex unknown 1, 6kDa	140,568,717 - 140,574,564
13.	TBDN100	Transcriptional coactivator tubedown-100	140,580,281 - 140,649,041
14.	RAB33B	RAB33B, member RAS oncogene family	140,732,565 - 140,754,672
15.	SET7	SET domain-containing protein 7	140,784,797 - 140,835,163
16.	MGST2	Microsomal glutathione S-transferase 2	140,944,620 - 140,983,002
17.	MAML3	Mastermind-like 3	140,997,246 - 141,432,838
18.	SCOC	Short coiled-coil protein	141,622,205 - 141,660,909
19.	CLGN	Calmegin	141,667,213 - 141,706,387
20.	UCP1	Uncoupling protein 1 (mitochondrial, proton carrier)	141,838,654 - 141,847,488
21.	RNF150	Ring finger protein 150	142,144,329 - 142,412,221
22.	ZNF330	Zinc finger protein 330	142,499,701 - 142,513,453
23.	IL15	Interleukin 15	142,915,359 - 143,012,216
24.	INPP4B	Inositol polyphosphate-4-phosphatase, type II, 105kDa	143,307,498 - 143,710,137
25.	USP38	Ubiquitin specific protease 38	144,463,702 - 144,500,244
26.	GAB1	GRB2-associated binding protein 1	144,615,654 - 144,748,883
27.	SMARCA5	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 5	144,792,469 - 144,832,171
28.	GYPE	Glycophorin E precursor	145,149,625 - 145,184,317
29.	GYPA	Glycophorin A	145,390,124 - 145,419,393
30.	GPERIK	Glycophorin Erik (STA)	145,390,151 - 145,419,393
31.	HGPMz	Glycophorin Mz II-IV fragment	145,393,446 - 145,419,393
32.	HHIP	Hedgehog interacting protein	145,924,777 - 146,017,486
33.	ANAPC10	Anaphase promoting complex subunit 10	146,273,920 - 146,376,953
34.	ABCE1	ATP-binding cassette, sub-family E (OABP), member 1	146,377,098 - 146,407,916
35.	HIN1	Putative HIV-1 induced protein HIN-1	146,437,893 - 146,456,915
36.	MADH1	SMAD, mothers against DPP homolog 1	146,760,630 - 146,836,836
37.	MMAA	Methylmalonic aciduria type A protein, mitochondrial precursor	146,917,831 - 146,934,342

38.	LSM6	LSM6 homolog, U6 small nuclear RNA associated	147,454,485 - 147,468,667
39.	POU4F2	POU domain, class 4, transcription factor 2	147,917,692 - 147,921,228
40.	NYD-SP14	NYD-SP14 protein	147,985,785 - 148,224,585
41.	SPAR	Surfactant protein A binding protein	148,304,677 - 148,308,829
42.	EDNRA	Endothelin receptor type A	148,759,734 - 148,823,710
43.	TMEM34	Transmembrane protein 34	148,911,657 - 148,914,275
44.	GRAF-2	Rho GTPase activating protein 10	149,011,058 - 149,351,531
45.	NR3C2	Nuclear receptor subfamily 3, group C, member 2	149,357,524 - 149,721,128

Supplementary Table 2. DNA sequence variants identified by sequencing the four regional candidate genes with a probable role in lipid metabolism.

Gene Symbol	Location	SNP Position on Genome Browser (Build 35) and Proband's genotype	rs number	Amino-Acid Change	Allele Frequency
ABCE1					
Proband 1					
Proband 2	Intron 11	146400050 GG/GC	rs10519725		0.965/0.035 (G>C)
UCP1					
Proband 1					
Proband 2					
TBDN100					
Proband 1	Intron 6	140623226 AA/AA	rs13146944		NA
	Intron 13	140639546 GG/AA			NA
	Intron 19	140666589 CC/AA	rs998310		0.533/0.467 (C>A)
Proband 2	Intron 6	140623226 AA/AG	rs13146944		NA
	Intron 13	140639546 GG/AA			NA
	Intron 19	140666589 CC/CA	rs998310		0.533/0.467 (C>A)
RAB33B					
Proband 1	Exon 1	140732750 GG/CC	rs13126617	5' UTR	NA
	Exon 1	140732901 CC/GG	rs13128486	5' UTR	NA
Proband 2	Exon 1	140732750 GG/GC	rs13126617	5' UTR 5'	NA
	Exon 1	140732901 CC/CG	rs13128486	UTR	NA
PCDH18					
Proband 1	Exon 3	138807288 CC/TT	rs10018837	Syn	NA
Proband 1	Exon 4	138799135 CC/TT	rs9286351	3' UTR	NA
Proband 2	Exon 1	138811058 CC/CT		5' UTR	NA
Proband 2	Intron 2	138807417 GG/CC	rs10006580		NA
Proband 2	Intron 4	138798380 GG/GT	rs9330354		NA
SLC7A11					
Proband 1	Intron 11	139450833 CC/CT	rs4504306		0.567/0.433(T>C)
Proband 1	Exon 8	139462019 GG/GA		Syn	NA
Proband 1	Exon 8	139450324 CC/CT	rs13120371	3' UTR	NA
Proband 2	Intron 3	139510922 TT/AA	rs6823642		NA
Proband 2	Exon 5	139498099 GG/CC	rs6838248	Syn	NA
Proband 2	Intron 5	139493404 AA/AG	rs11944083		NA
Proband 2	Intron 11	139450833 CC/CT	rs4504306		0.567/0.433(T>C)
Proband 2	Exon 8	139450324 CC/TT	rs13120371	3' UTR	NA

ABCE1 indicates ATP-binding cassette, sub-family E (OABP), member 1 gene; UCP1, uncoupling protein 1 (mitochondrial, proton carrier) gene; TBDN100, TBDN100 transcriptional coactivator tubedown-100 gene; and RAB33B indicates RAB33B, member RAS oncogene family gene; PCDH18, protocadherin 18 gene; and SLC7A1, solute carrier family 7, (cationic amino acid transporter, y+ system) member 11 gene. Syn indicates a synonymous SNP; and NA, not available. A

total of 10 of these 17 variants (rs10519725, the new G/A variant in intron 13 of TBDN100, the new C/T variant in 5' UTR of PCDH18, rs10006580, rs10018837, rs9286351, rs9330354, rs11944083, the new synonymous G/A variant in exon 8 in SLC7A11, rs4504306) were genotyped in all family members of the particular family/families but none of them segregated with the low HDL-C trait in these extended families.

Preface to chaper 6

In addition to the parametric linkage analysis, we performed a genome-wide QTL analysis to identify region(s) of linkage to the HDL-C trait. Based on this approach a locus on chromosome 16q was identified that had a LOD score of 2.69 at 95 cM. I fine-mapped this region on chromosome 16q23-24 and sequenced all coding regions of the 39 genes located in this region. The results of this work has been resubmitted in revision form to the European Journal of Human Genetics: "*Fine Mapping and Association Studies of a High-Density Lipoprotein Cholesterol Linkage Region on Chromosome 16 in French-Canadian Subjects"* and will be presented in this chapter.

CHAPTER 6

FINE MAPPING AND ASSOCIATION STUDIES OF A HIGH-DENSITY LIPOPROTEIN CHOLESTEROL LINKAGE REGION ON CHROMOSOME 16 IN FRENCH-CANADIAN SUBJECTS

Zari Dastani[1,2], Päivi Pajukanta[4,] Michel Marcil[1], Nicholas Rudzicz[3], Isabelle Ruel[1], Swneke D. Bailey[2], Jenny C. Lee [4], Mathieu Lemire[5,11], Janet Faith[5], Jill Platko[6,9], John Rioux[6,10], Thomas J. Hudson[2,5,7,11] , Daniel Gaudet[8], James C. Engert[2,7] and Jacques Genest[1,2,7]

[1]Cardiovascular Genetics Laboratory, McGill University Health Centre, Royal Victoria Hospital, Montreal, Canada. [2]Department of Human Genetics, McGill University, Montreal, Canada. [3]Department of Computer Science, McGill University, Montreal, Canada. [4]Department of Human Genetics, David Geffen School of Medicine at UCLA, University of California, Los Angeles, USA. [5]McGill University and Genome Quebec Innovation Centre, Montreal, Canada. [6]Whitehead Institute for Biomedical Research, MIT, Cambridge, MA, USA. [7]Department of Medicine, McGill University, Montreal, Canada. [8]Dyslipidemia, Diabetes and Atherosclerosis Group and Community Genomics Research Center, Université de Montréal and Complexe hospitalier de la Sagamie, Chicoutimi, Canada. [9]Present address: Massachusetts General Hospital. [10]Present address: Broad Institute of MIT and Harvard, Cambridge, MA; Université de Montréal,

Montreal, Canada; Montreal Heart Institute, Montreal, Canada. [11]Present address: Ontario Institute for Cancer Research, Toronto, Ontario, Canada.

Key words: high-density lipoprotein cholesterol, family study, complex traits, coronary heart disease, gene identification.

***European Jouranl of Human Genetics* 2009, Resubmitted with revision (23, June, 2009)**

6.1 Abstract

Low levels of high-density lipoprotein cholesterol (HDL-C) are an independent risk factor for cardiovascular disease. To identify novel genetic variants that contribute to HDL-C, we performed genome-wide scans and quantitative association studies in two study samples: a Quebec-wide study consisting of 11 multi-generational families and a study of 61 families from the Saguenay–Lac St-Jean (SLSJ) region of Quebec. The heritability of HDL-C in these study samples was 0.73 and 0.49 respectively. Variance components linkage methods identified a LOD score of 2.61, at 98 cM near the marker D16S515 in the Quebec-wide families and a LOD score of 2.96 at 86 cM, near the marker D16S2624 in the SLSJ families. In the Quebec-wide sample, four families demonstrated segregation over a 25.5 cM (18 Mb) region, which was further reduced to 6.6 Mb with additional markers. The coding regions of all genes within this region were sequenced. A missense variant in CHST6, segregated in four families and with additional families we observed a *p* value of 0.015 for this variant. However, an association study of this SNP in unrelated Quebec-wide samples was not significant. We also identified a SNP (rs11646677) in the same region, which was significantly associated with low HDL-C (p=0.016) in the SLSJ study sample. In addition, RT-PCR results from cultured cells demonstrated a significant difference in expression of CHST6 and KIAA1576, another gene in the region. Our data constitute additional evidence for a locus on chromosome 16q23-24 that affects HDL-C levels in two independent French Canadian studies.

6.2 Introduction

Coronary artery disease (CAD) is the most common cause of death in Western societies as well as in emerging market economies.[191, 243] Epidemiological studies have consistently shown an association between low levels of high-density lipoprotein cholesterol (HDL-C) and the prevalence of CAD.[20, 244, 245] HDL-C has a key role in reverse cholesterol transport, mobilizing cholesterol from peripheral tissues to the liver, a mechanism contributing to the cardio-protective effect of HDL-C.

Several studies have shown modest to high estimates of heritability (0.24 to 0.83) for plasma HDL-C levels,[209, 246] underlining the importance of genetic factors. However, until recently, only a few specific genes had been identified (see Dastani *et al* [209] for review). Rare mutations in the ATP binding cassette A1 (*ABCA1*), lecithin: cholesterol acyltransferase (*LCAT*) and apolipoprotein A-I (*APOA1*) genes and other genes involved in HDL metabolism have been reported in individuals with low levels of HDL-C[64, 108]. However, these account for approximately ~26% of familial forms of HDL deficiency in Quebec.[106, 247] Identifying novel genes that regulate HDL-C levels may provide further insights into lipoprotein metabolism and potential pathways for pharmacological modulation of HDL-C.

In the present study, we examined the genetic determinants of low HDL-C in well-characterized families and individuals of French Canadian descent. Specifically, we performed genome-wide linkage followed by association

analyses of SNPs located on chromosome 16q23-24 with the HDL-C trait, and examined the expression of a number of candidate genes in this region.

6.3 Methods

6.3.1 Subjects

Probands for the QUE study sample with HDL-C levels < the 5th percentile, (age/sex-specific) based on the Lipid Research Clinics population studies data book, as previously described,[235] were selected from the Preventive Cardiology/Lipid Clinic of the McGill University Health Center, Royal Victoria Hospital (Montreal, Canada). Exclusion criteria for the probands were: severe hypertriglyceridemia (>10 mmol/L), cellular cholesterol or phospholipid efflux defect in skin fibroblasts,[236] or the presence of known *ABCA1* or *ApoA1* mutations based on exonic sequencing. A total of 11 families out of 29 recruited multigenerational French-Canadian families, each exhibiting at least 3 individuals with low HDL in consecutive generations (n=412 subjects) were selected for genome-wide linkage analysis. The size of the families varied between 4 and 109 individuals, with a median of 41. Family members were sampled after a 12-hour fast and discontinuation of lipid modifying medications for at least 2 weeks. 136 cases (HDL<5%) and 218 controls (HDL>25%) derived from the QUE families and additional CAD patients were selected for an association study. These additional CAD patients consisted of unrelated French Canadian patients with premature CAD [202]. All subjects provided informed consent for plasma and DNA sampling, isolation, and storage. The Research Ethics Board of the McGill

University Health Centre approved the research protocol. The SLSJ sample was composed of 410 phenotyped individuals from 61 nuclear families from the Saguenay-Lac Saint-Jean region of Quebec that were ascertained as part of ongoing family-based studies of CAD or type 2 diabetes. Both studies required that at least two siblings be affected. Biochemical measurements were performed after a 3-week medication washout period for all hypocholesterolemic and antihypertensive drugs. The SLSJ case/control sample was predominantly derived from the families and consisted of 94 individuals with an HDL-C $<5^{th}$ percentile and 94 individuals with an HDL-C $>40^{th}$ percentile. **Table 1** summarizes the subjects for this study.

6.3.2 Biochemical measurements

HDL and triglycerides were measured using standardized techniques as described previously.[236]

6.3.3 Genome scan markers and data quality control

The genotyping for the QUE families was performed by DeCode Genetics (Reykjavik, Iceland) with 485 autosomal microsatellite markers from the Decode map (an average marker density of 6 cM) on 412 subjects from the QUE families (n=11, mean family size=41). The genome scan of on 410 phenotyped individuals from 61 SLSJ families was executed as described in Rioux *et al*.[248] by using either 312 or 396 microsatellite markers, with an average marker density of 11.9 and 9 cM respectively. These markers were modified versions of the Co-operative

Human linkage Center Screening Set version 6.0, that were supplemented with additional Genethon markers to increase the information content[249]. Overall, 34 markers were common between the two marker sets used, however there were no markers in common on chromosome 16.

For both the SLSJ and the QUE studies, the software GRR[250] was used to identify mispaternities and sample mix-ups. In total, three mispaternities and one sample mix-up were identified and appropriate adjustments were made before further analysis.

6.3.4 Sequencing and genotyping

Sequencing of all 39 known genes in narrowed region was performed using the Thermo-Sequenase BigDye Direct Cycle-Sequencing kit (Applied Biosystems). We also selected 200 tagging SNPs (tSNPs) that spanned the region from 55574820 to 77244563 bp (NCBI, build 36.1) corresponding to the linkage peak on chromosome 16q13-23.1. T The selected SNPs were based on $r^2<0.8$ between tSNPs and minor allele frequency (MAF) >0.05 from the European samples of the HapMap project in genes that were differentially expressed after cholesterol or 22-hydroxycholesterol stimulation of human fibroblasts, using Affymetrix (Santa Clara, California, USA) micro-array data (**supplementary Table 1**). The SNPs were genotyped in the SLSJ study sample using the SNPstream Genotyping System (Beckman Coulter, Inc). We also genotyped 51 tSNPs in the QUE sample in a parallel study as previously described.[251]

6.3.5 Haplotype construction

Ten Families with a positive LOD score (LOD>0) on chromosome 16 were chosen for haplotype analysis. Haplotypes were constructed with the program Merlin[252] using both microsatellite markers and SNPs based on HDL-C less than 5% as the affected status. The haplotype analysis for SNPs in unrelated subjects from the SLSJ and the QUE study sample was performed using Haploview v4.1 [251, 253]

6.3.6 RNA extraction and RT-PCR

Skin fibroblasts were available for three out of four probands from the QUE families with the highest linkage scores. These and three control cell lines were cultured based on a standard protocol.[236] Total RNA was prepared from cultured fibroblasts using the RNeasy mini RNA extraction kit (Qiagen) according to the manufacturer's instructions. Total RNA (100 ng) in a 14 µL-reaction was reverse transcribed using the cDNA Qiagen Kit according to the manufacturer's protocols. Real-Time quantitative PCR assays were performed using the Quantitect SYBR® Green PCR kit (Qiagen) on an ABI PRISM® 7300 Sequence Detection System (Applied Biosystems). The amplifications were carried out in a 96-well plate with 50 µL reaction volumes and 40 amplification cycles (94°C, 15 s; 55°C, 30 s; 72°C, 34 s). Experiments were performed in triplicate and the mRNA expression was taken as the mean of three different experiments. The expression of each gene was normalized to succinate dehydrogenase expression levels. Fold changes relative to controls were

determined using ΔΔCt method. A mean fold change ($2^{-\Delta\Delta Ct}$) and the standard error of the mean value were determined and significance was determined with a two tailed *t*-test.

6.3.7 Statistical analyses

SIMWALK2 version 2.91 was used to estimate the multipoint inheritance vectors in the QUE families.[239, 254] Genome scans were performed using the variance components approach as implemented in SOLAR. For multipoint linkage analysis, the HDL-C phenotype was adjusted for triglycerides, sex and age X sex using SOLAR (version 4.2.0). Association tests in families were performed using the Mendel version 9.0.0[255] with the option for association given linkage. Case/control association analyses were performed with Haploview version 4.1. Heritability of HDL-C was analyzed with the method implemented in the SOLAR version 4.2.0 software package with the " ascertainment correction" option. While we realize that using this correction may still lead to biased results, the heritability we report for our samples is consistent with previous report of HDL.

6.4 Results

In the QUE families, heritability of HDL-C was estimated to be 73% ($p=9.89X10^{-16}$) in the genotyped individuals. Sex, age X sex, and triglycerides were significant contributors ($p<0.05$) to the variance of HDL-C in this study

sample. The heritability estimate for the SLSJ families was 49% ($p = 6.55X\ 10^{-22}$). Sex and triglycerides were significant covariates in this study sample. The results of the linkage analysis for HDL-C in the SLSJ and QUE studies are shown in **supplementary Figure 1**.

In the genome scan for the QUE study sample, we identified four chromosomal regions that may harbor QTL for HDL-C (LOD scores >1). We observed a LOD score of 2.61 at 98 cM (based on the Decode Map[256]) near the marker D16S515 in chromosome 16q23-24. In the SLSJ study sample, the highest multipoint LOD score (2.96) was observed at 86 cM on chromosome 16, located 1.25cM away from D16S2624 in chromosome 16q22-23. Chromosomal regions with a multipoint LOD score ≥1 in either of the two genome scans are summarized in **Table 2**. **Figure 1** shows the multipoint LOD scores on chromosome 16 for the SLSJ and QUE study samples as well as two previously reported findings in the region.

Haplotype construction identified a 25.5 cM (18 Mb) region on 16q23-24 that segregated with the low HDL-C phenotype in the four QUE families with the highest HDL-C LOD scores. The segregating region was further restricted by additional microsatellite markers from 25.5 cM to 18.1 cM (~7.8 Mb) in the family with the highest LOD score (**Figure 2a**). Using the UCSC genome browser (http://genome.ucsc.edu/), we identified 39 genes residing within this 7.8 Mb region.

In the SLSJ study sample, we tested for association between HDL-C affection and 200 tSNPs selected from genes in the region that were differentially

expressed in human fibroblasts treated with cholesterol or 22-hydrocholesterol and thus appeared to be sterol-regulated (microarray data not shown). One SNP, rs11646677, which was located within the shared haplotype from the QUE families, showed evidence of association with HDL-C affection (p=0.016) (**supplementary Figure 2**).

We sequenced all coding regions of the 39 genes located in this region (**Figure 2a**). The *LCAT* gene was excluded from the candidate region by a recombination event (subject # 407, **supplementary Figure 3a**) and by sequencing in all QUE probands. We identified several non-synonymous variants within the 39 genes. These were genotyped in family members and checked for segregation. We found three non-synonymous variants within the CHST6 gene and its homologue CHST5, located in the same region. One of the variants in the CHST6 gene, R162G, segregated with HDL-C in the four QUE families that had the highest linkage score (**supplementary Figure 3 a, b, c and d**). In association analysis of these families and additional families recruited later using Mendel with the option association given linkage we observed a p value of 0.015 for this variant. However, an association study in 136 cases and 218 controls did not confirm this finding (data not shown). This SNP and four other telomeric SNPs (spanning from 74070744 to 74822026 bps build 36.1) comprise a haplotype (**Figure 2b**) that is found in the four families but had a haplotype frequency of only 0.05 in our unrelated samples from Quebec.

The non-synonymous variants found in all the sequenced genes are listed in **Table 3**. Segregation analyses with these SNPs narrowed down the region to 6.6 Mb in the family with the highest LOD score.

Because none of the other non-synonymous variants identified by sequence analysis segregated in our families, we investigated the potential influence of regulatory variants on HDL-C levels, using RT-PCR analyses of nine genes within the 6.6-Mb region, between the genes CHST6 and WWOX. Three of those genes (ADAMTS18, CLEC3A, CNTNAP4) were not significantly expressed in fibroblasts. The results of those genes that are expressed in fibroblasts are shown in Figure 3. The expression of the CHST6 and KIAA1576 genes were found to be increased in cases as compared to controls ($p<0.001$). The CHST5 and WWOX genes were not differentially expressed, while the NUDT7 gene had lower expression in cases compared with controls but this was not statistically significant.

6.5 Discussion

The inverse relation between HDL-C plasma levels and CAD has been well described in multiple epidemiological studies. Though environmental influences play an important role in determining serum HDL-C levels, genetics plays a major role in determining HDL-C levels. The genetic heterogeneity underlying HDL-C has been underscored by the discovery and identification of several genes involved in HDL metabolism[209]. We performed two genome-wide scans in large families with probands ascertained for type 2 diabetes, CHD or low

HDL-C. We obtained suggestive evidence for linkage [257] on chromosome 16 from both of these French Canadian study samples. The region on chromosome 16q22-24 has been associated with HDL-C in several studies. Specifically, two previous studies, one involving Mexican Americans [181], and the other using Finnish and Dutch family samples [175], both provided evidence of linkage to this region on chromosome 16 and both produced linkage peaks that were located less than 12 cM from the peaks observed in our analysis of Quebec families. Indeed, at least eight previous genome-wide scan studies have identified chromosomal loci for HDL-C on chromosome 16 with LOD scores greater than 1.[97, 176, 177, 183, 258-261] In addition, Mehrabian *et al*.[260] found a QTL for HDL-C on mouse chromosome eight at genetic markers D8Mit12 (LOD=3.6), and D8Mit14 (LOD=3.5) that is syntenic to human chromosomal region 16q22.1-16q24. Moreover, in a study of the mouse strains C57BL/6J and 129S1/SvImJ, Ishimori *et al.*[261] identified a suggestive QTL on chromosome 8 (cM 44, LOD 2.6) consistent with the locus identified by Mehrabian *et al*.[260] Thus, these mouse QTLs for HDL-C provide additional confirmation of the involvement of this region of human chromosome 16 in HDL-C metabolism.

The *LCAT* and *CETP* genes, two well-known genes involved in HDL metabolism, are located near chromosome 16q23-24. LCAT mediates the formation of cholesteryl esters within HDL particles, which allows the maintenance of a gradient of free cholesterol from the plasma membrane to HDL particles and ensures a net movement of cholesterol from the cell to HDL. We sequenced the *LCAT* gene in the probands of the QUE families and no coding

variants were identified. CETP promotes the transfer of cholesteryl esters from HDL particles to triglyceride-rich (VLDL and IDL) particles in exchange for triglycerides.[192, 262, 263] Many association studies have revealed a significant association between CETP gene polymorphisms and levels of plasma HDL-C (reviewed in ref. [264, 265]). Fine mapping of the 16q23-24 region with 12 additional microsatellites markers excluded both the *LCAT* and *CETP* genes from the linked region.

Our previous analysis of common variants in the region demonstrated an association between an intronic SNP in the WWOX gene (rs2548861) and HDL-C in the QUE sample *(p*=0.02 using a dominant model). [251] However this SNP, explains only 1.5% of the variance in HDL-C concentrations at the population level [251] and is not likely to be a major determinant of HDL-C in families that demonstrate a strong pattern of segregation.

Using both microsatellite markers and SNPs, we were able to identify a minimal haplotype in one linked family of approximately 6.6 Mb on chromosome 16q23-24. Of note, a genome-wide association study from the Diabetes Genetics Initiative (http://www.broad.mit.edu/diabetes/) also showed evidence of association (p value < 0.0001) with 6 SNPs (rs8057477, rs4888535, rs6564359, rs7404386, rs3924497, rs8049365) between the positions 75307621-75376819, all residing within the linked region. , We sequenced all coding and exon-intron boundaries of the 39 genes found within this region. One variant in *CHST6* showed evidence of co-segregation with low HDL-C and was investigated further. CHST6 belongs to the sulfotransferase family. It catalyzes the transfer of

sulfate to position 6 of non-reducing N-acetylglucosamine (GlcNAc) residues of keratin.[266] It is located on the Golgi apparatus membrane and is mainly expressed in cornea and brain but also in the spinal cord and trachea. Defects in the *CHST6* gene are the cause of macular corneal dystrophy [MIM: 217800]. In the present study, *CHST6* was considered a good candidate gene for the low HDL-C trait because ocular manifestations have been previously associated with other gene defects causing a low HDL-C phenotype (e.g. *LCAT*, *APOA1*). However, there is currently no published data characterizing the lipid profile of patients with macular corneal dystrophy or directly relating the function of this gene to lipid metabolism.

We resequenced all other genes in the 6.6 Mb region and found no segregating variants but could not rule out variants found in noncoding regions of these loci, which could have an effect on the transcription of these genes. Therefore we examined the expression of genes in the three probands of the families linked to this region. An increase in the level of *CHST6* expression was found in these probands and we hypothesize that it might contribute to HDL metabolism. However, our RT-PCR results should be interpreted with caution as fibroblasts may not be the most appropriate cell model. Furthermore, we cannot exclude the possibility that the observed linkage and association at this locus may be caused by another nearby gene.

In a proteomic analysis of HDL by Rezaee *et al*., C-type lectin domain family 3 member A (*CLEC3A*) has been identified as a HDL-associated protein.[267] The gene coding for CLEC3A protein also resides in the fine-mapped

region on chromosome 16q23-24. Sequencing of this gene in probands revealed one non-synonymous (Q197L) variant in one family. However this variant did not segregate with low HDL-C in this family and as *CLEC3A* is not expressed in fibroblasts, we were not able to investigate the differential expression level of its mRNA.

In summary, we have fine-mapped a region on chromosome 16q23-24 that is likely to harbour a gene for low plasma levels of HDL-C. Recently, genome-wide association studies have been performed for lipid traits, including HDL-C.[185, 268] Although these studies have revealed polymorphisms associated with the variation in HDL-C level in a number of European populations, such variants do not account for a significant fraction of their variance and are not likely to be the cause of extreme levels of HDL-C that segregate in some families. More likely, multiple rare genetic variants that have a large effect lead to the extreme levels of HDL-C observed in families. These may sometimes be found in the same genes as the common variants that effect HDL-C, but perhaps not always. Therefore the identification of the causal genes for more severe forms of low HDL-C may contribute to further understanding the regulation of HDL-C levels and subsequently provide new therapeutic targets for CAD prevention. With recent advances in sequencing technology, it should be feasible to resequence this particular region in its entirety to identify the regulatory variants that contribute to HDL-C levels.

6.6 Abbreviations

ABCA1, ATP binding cassette A1; *APOA1*, apolipoprotein A-I; CAD, coronary artery disease; FBAT, family-based association test; HDL-C, high-density lipoprotein cholesterol; *LCAT*, lecithin:cholesterol acyltransferase; LDL-C, low-density lipoprotein cholesterol; MAF, minor allele frequency; SLSJ, Saguenay-Lac St-Jean region of Quebec, Canada; SNPs, single nucleotide polymorphisms; tSNPs, tagging SNPs.

6.7 Acknowledgments

We thank all patients and their families for participating in the studies. This work was supported by Grant CIHR MOP 62834 from the Canadian Institutes of Health Research (CIHR). Jacques Genest holds the McGill University-Novartis Chair in Medicine. Zari Dastani is a recipient of a Doctoral research award from the Heart and Stroke Foundation of Canada (HSFC). Michel Marcil is a research scholar from the *Fonds de la recherche en santé du Québec* (FRSQ). Päivi Pajukanta is supported by NIH grants HL 28481, HL08276, and HL095056.

6.8 Figures

Figure 1. Multipoint LOD scores for HDL-C on chromosome 16 in the SLSJ and QUE study samples as well as previously reported findings in the region. The filled line and dashed line correspond to results from the QUE samples near marker D16S515 at 16q23-24 and SLSJ samples near marker D16S2624 at 16q22-23. ● reported peak in Mexican American families near markers D16S2624- D16S518 at 16q22.3-23.1 [181]. ■ reported peak in Finnish and Dutch families near marker D16S3096 at 16q22.1-24.3[175].

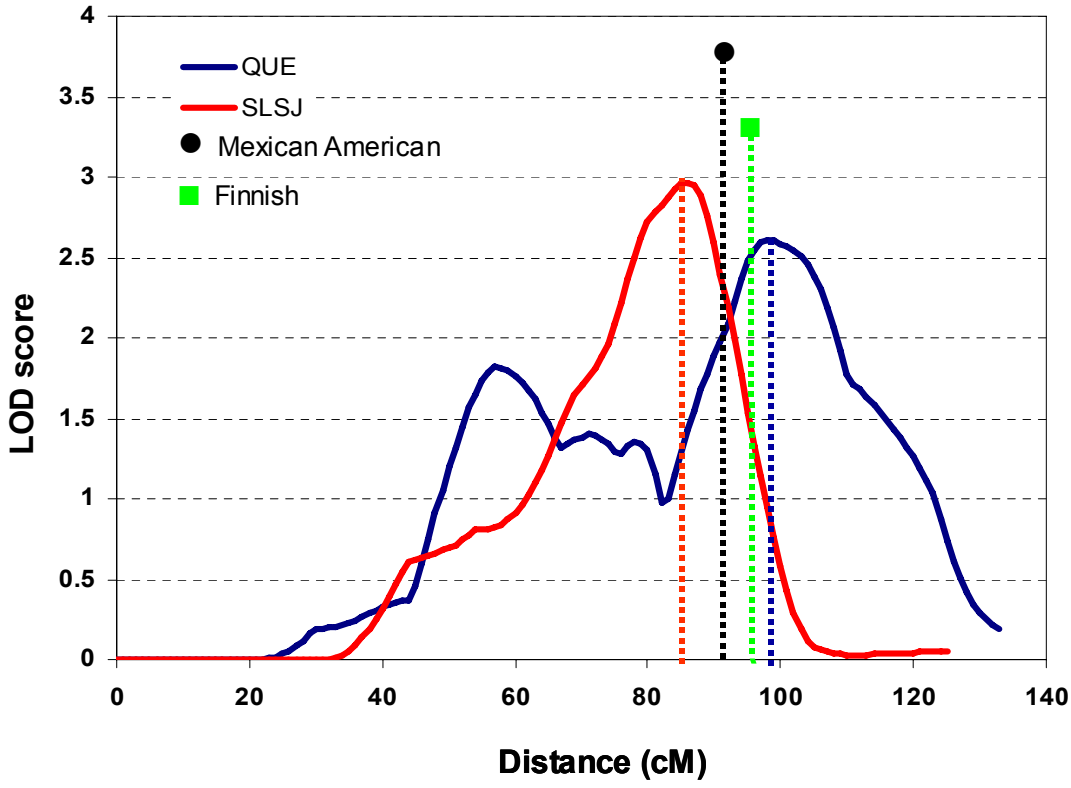
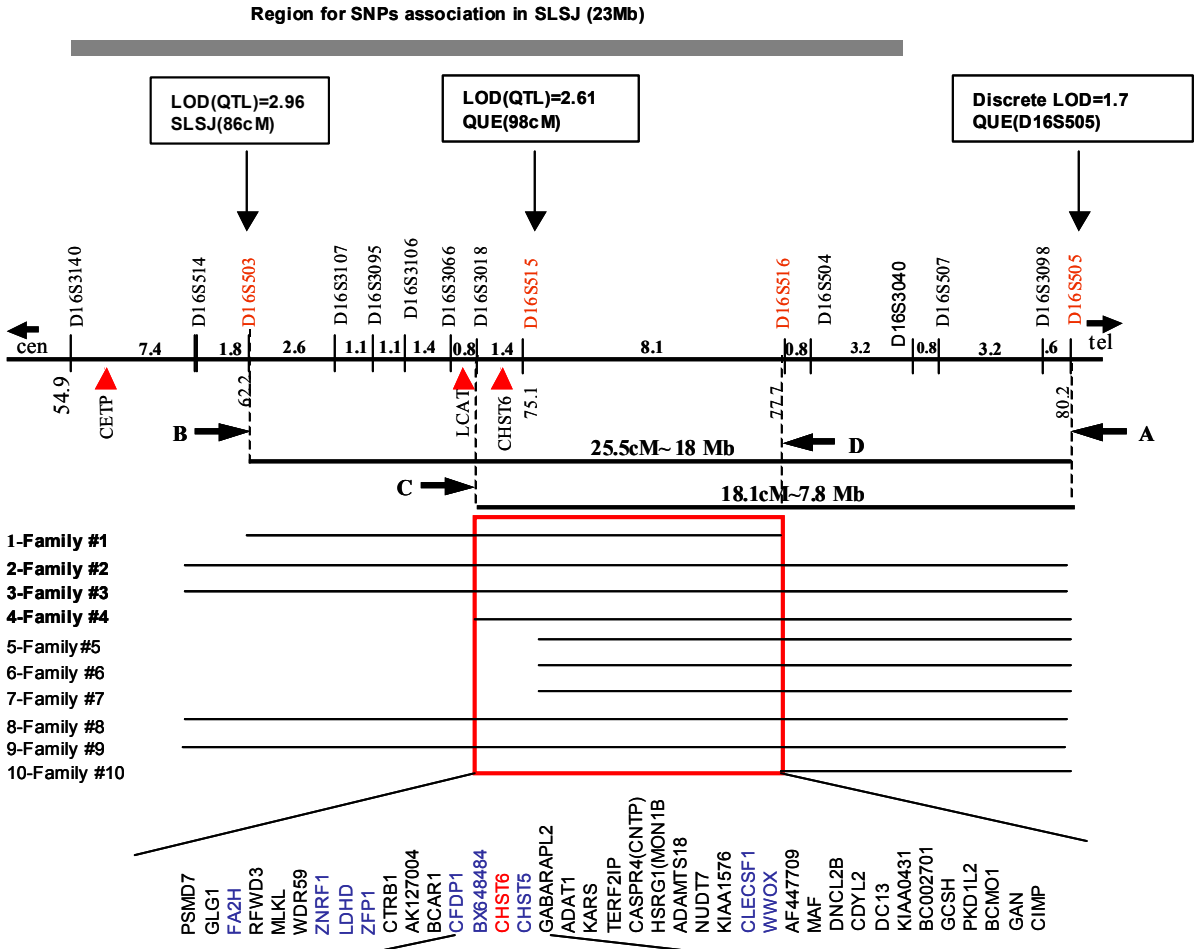


Figure 2. Linked region for HDL-C on chromosome 16q21-24.

Panel A. The position of microsatellite markers with the interval distance (cM) between them is indicated and below the line is actual the physical locations (Mb). Brown markers are genome-wide scan markers and other markers are from the subsequent fine mapping. Position of the LOD score peak is shown in the box. Family identification numbers are indicated on the left side and the segregating haploytypes for each family are shown. The position of *CETP*, *LCAT*, and *CHST6* are indicated by the arrowhead. The distance between arrows A and B corresponds to the initially identified linked region, which was narrowed down to the region between A and C. The known genes in this region are listed shown in panel B. Families indicated in bold carry the R162G mutation in *CHST6*.

Panel b. Haplotype with a frequency of only 0.05 in our unrelated samples from Quebec.

A.


Region for SNPs association in SLSJ (23Mb)
LOD(QTL)=2.96
SLSJ(86cM)
LOD(QTL)=2.61
QUE(98cM)
Discrete LOD=1.7
QUE(D16S505)
D16S3140
D16S514
D16S503
D16S3107
D16S3095
D16S3106
D16S3066
D16S3018
D16S515
D16S516
D16S504
D16S3040
D16S507
D16S3098
D16S505
cen
tel
7.4
1.8
2.6
1.1
1.1
1.4
0.8
1.4
8.1
0.8
3.2
0.8
3.2
.6
54.9
CETP
62.2
LCAT
CHST6
75.1
77.7
80.2
B
A
25.5cM~ 18 Mb
D
C
18.1cM~7.8 Mb
1-Family #1
2-Family #2
3-Family #3
4-Family #4
5-Family#5
6-Family #6
7-Family #7
8-Family #8
9-Family #9
10-Family #10
PSMD7
GLG1
FA2H
RFWD3
MLKL
WDR59
ZNRF1
LDHD
ZFP1
CTRB1
AK127004
BCAR1
CFDP1
BX648484
CHST6
CHST5
GABARAPL2
ADAT1
KARS
TERF2IP
CASPR4(CNTP)
HSRG1(MON1B
ADAMTS18
NUDT7
KIAA1576
CLECSF1
WWOX
AF447709
MAF
DNCL2B
CDYL2
DC13
KIAA0431
BC002701
GCSH
PKD1L2
BCMO1
GAN
CIMP

B

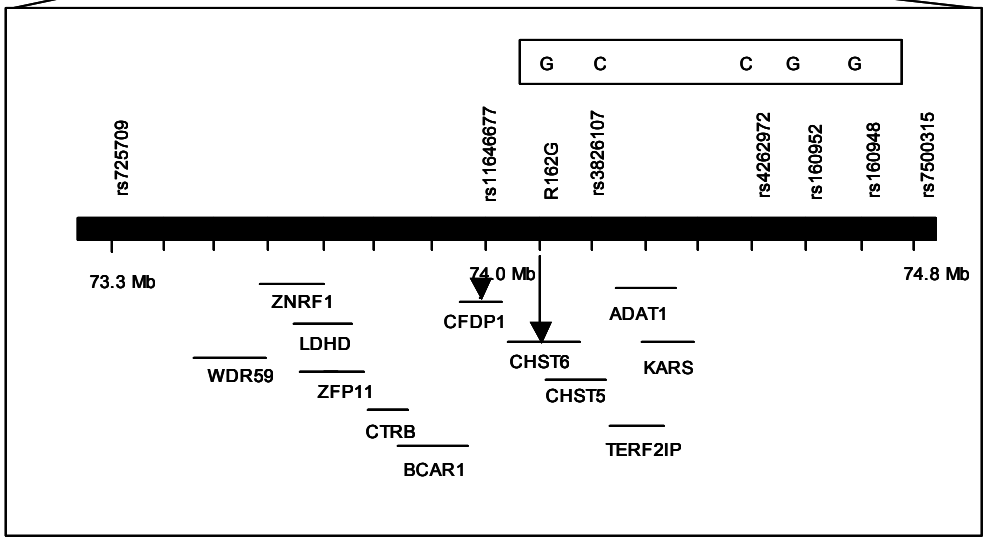
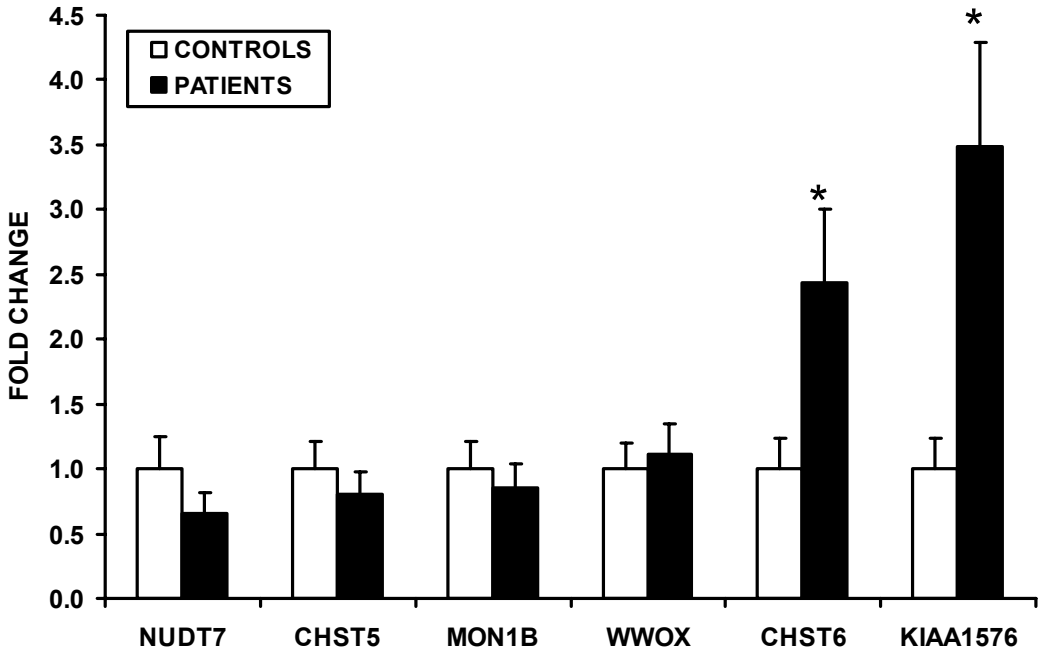

G C C G G
rs725709
rs11646677
R162G
rs3826107
rs4262972
rs160952
rs160948
rs7500315
73.3 Mb
74.0 Mb
74.8 Mb
ZNRF1
LDHD
WDR59
ZFP11
CTRB
BCAR1
CFDP1
CHST6
CHST5
ADAT1
KARS
TERF2IP

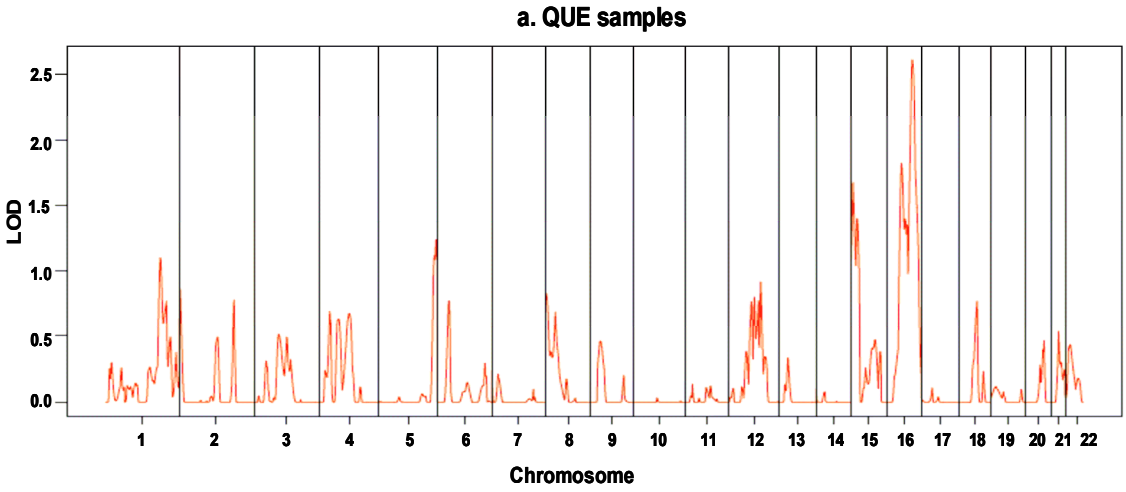
Figure 3. mRNA expression levels of select candidate genes found in the linked region.

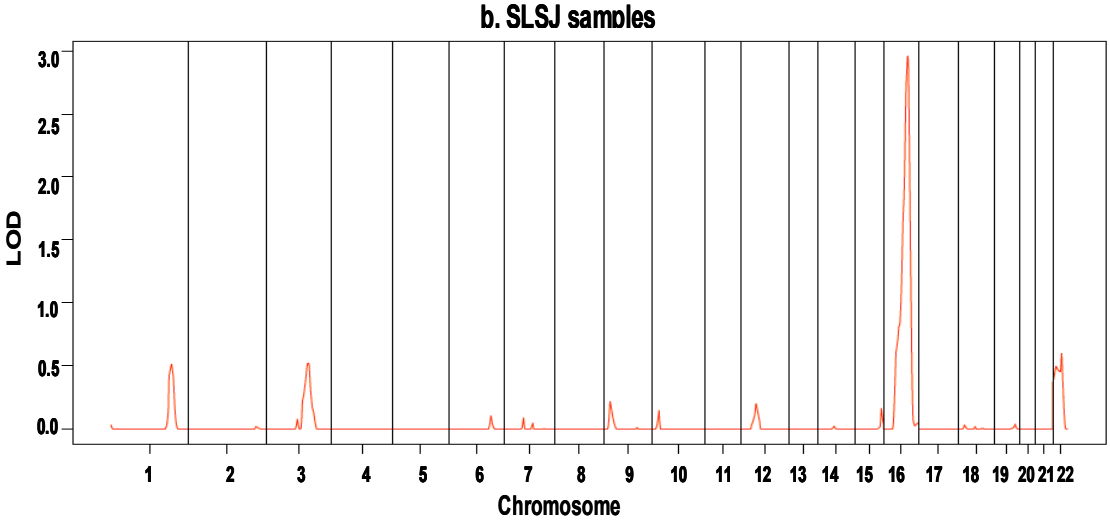
mRNA expression of the genes (normalized to succinate dehydrogenase) from the skin fibroblasts of three probands are compared with the mean mRNA expression from three controls. The mRNA expression level was taken as the mean fold change ($2^{-\Delta\Delta Ct}$). The standard error of the mean values of triplicate experiments are shown.



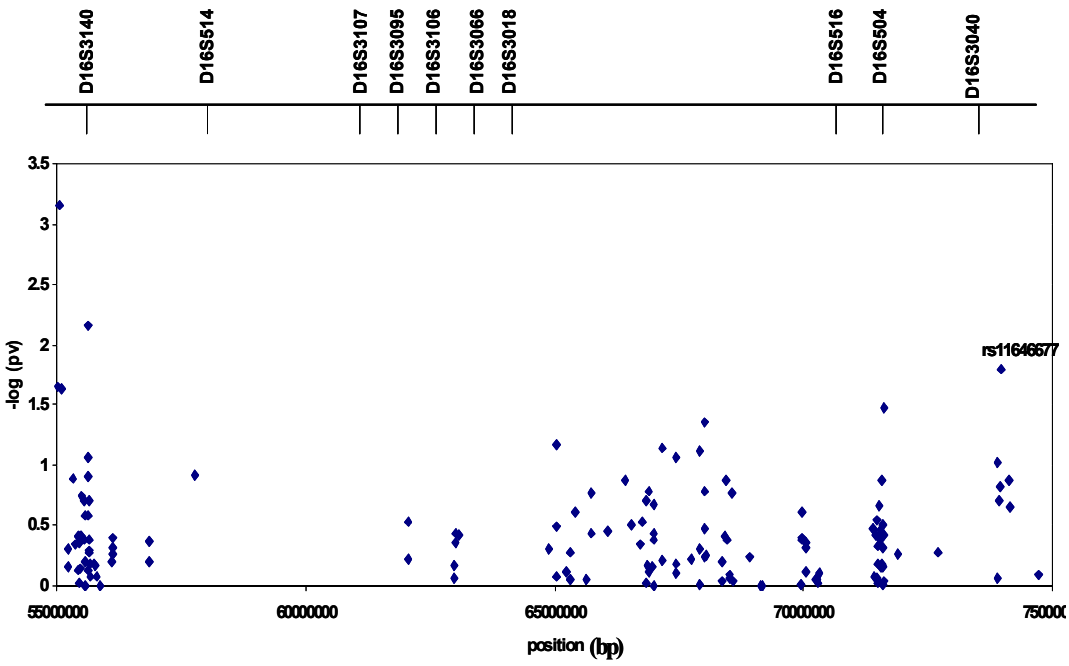
Supplementary Figure 1. Results of linkage analysis from SOLAR in the a) QUE and b)SLSJ study samples.

Genome-wide multipoint linkage results for HDL-C with LOD scores on the *y* axis and chromosomal position for all chromosomes on *x* axis.

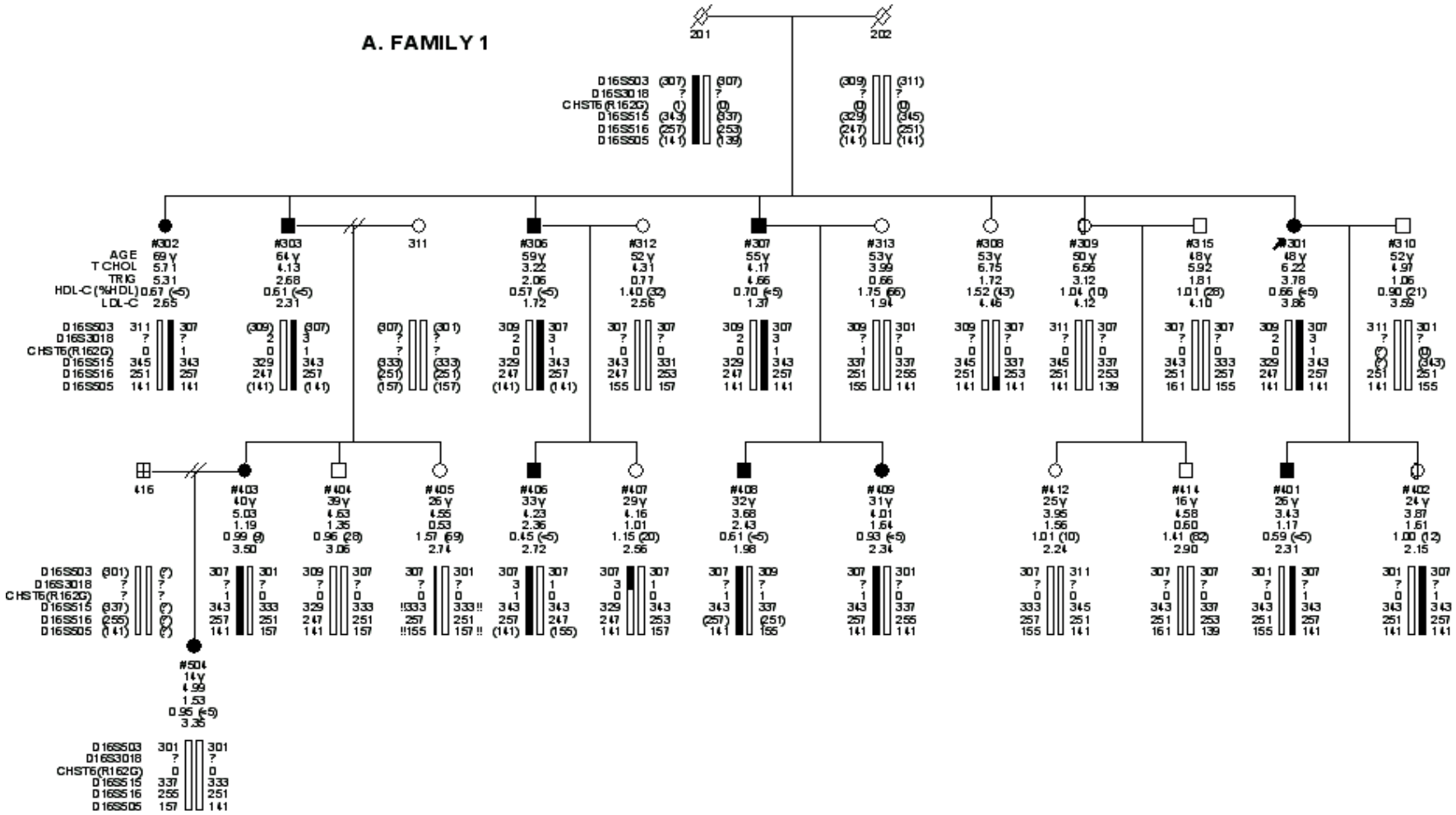




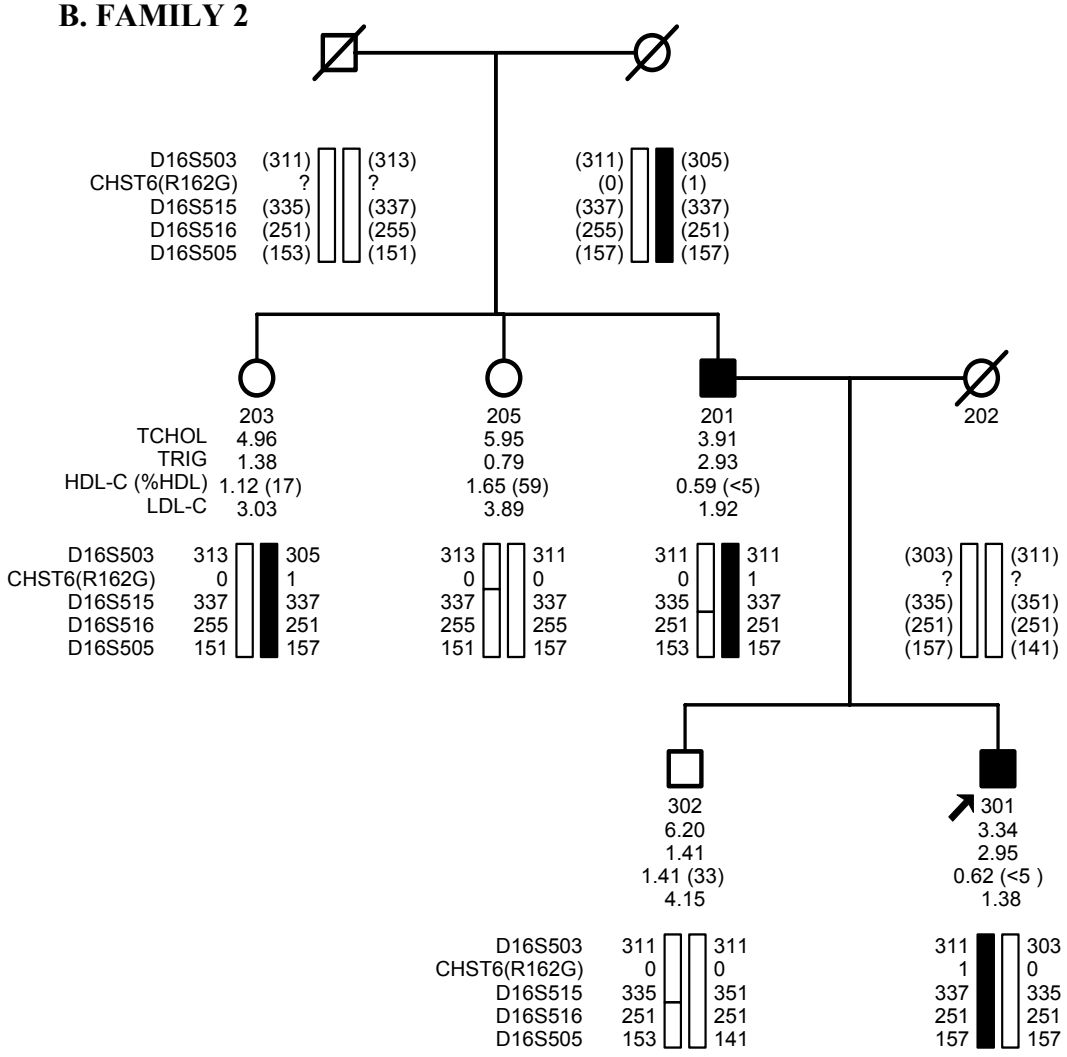
Supplementary Figure 2. SNPs vs – log(p-value) in the SLSJ cohort.
x axis shows the physical distance in bp. *y* axis is –log P value. Corresponding microsatellite markers in this region are indicated at the top of the figure. The indicated SNP with a corresponding P value is the nearest SNP to the *CHST6* gene.



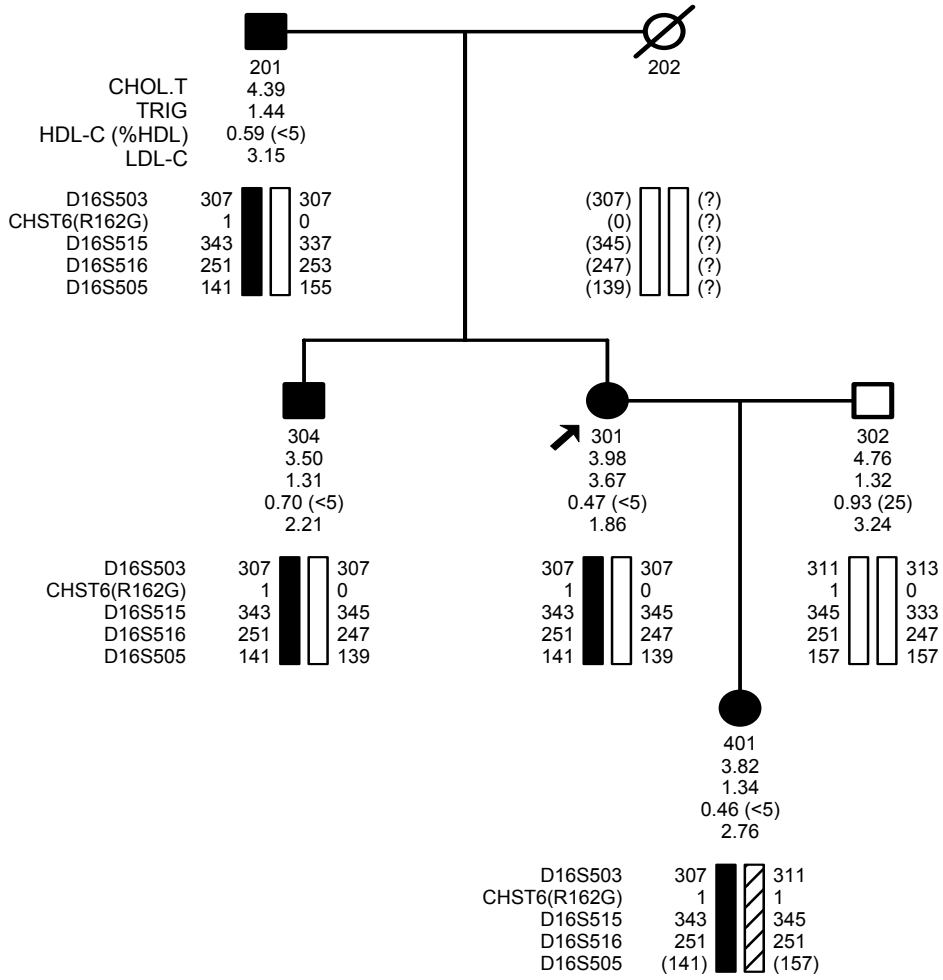
Supplementary Figure 3. Co-segregation of the five SNP haplotype that includes the variant of the *CHST6* gene and the low HDL-C trait.
(A) pedigree of family 1; (B) pedigree of family 2; (C) pedigree of family 3; (D) pedigree of family 4. Probands are identified by black arrows. Diagonal and crossed lines through the symbols indicate deceased individuals and unavailable data, respectively. Filled bars indicate alleles bearing the mutation. Haplotypes around the *CHST6* gene locus include four microsatellite markers flanking the gene and the R162G mutation itself (0, absence; 1, presence). Not shown are four other SNPs that are on the same haplotype as R162G (see Figure 2, Panel B) in all four families. Horizontal lines and partly filled bars represent recombination. Individual, age, total cholesterol (TCHOL), triglyceride (TRIG), HDL-C followed by percentile ranking for age and gender, and LDL-C are listed respectively below each symbol. Black symbols indicate individuals with a HDL-C value below the 5^{th} percentile. Inferred marker genotypes are indicated in brackets.


A. FAMILY 1
201
202
D16S503 (307) (307)
D16S3018 ? ?
CHST6(R162G) (C) (T)
D16S515 (343) (337)
D16S516 (257) (253)
D16S505 (141) (139)
(309) (311)
? ?
(329) (345)
(247) (251)
(141) (141)
#302
69 y
5.71
5.31
0.67 (<5)
2.65
#303
64 y
4.13
2.68
0.61 (<5)
2.31
311
#306
59 y
3.22
2.06
0.57 (<5)
1.72
#312
52 y
4.31
0.77
1.40 (32)
2.56
#307
55 y
4.17
4.66
0.70 (<5)
1.37
#313
53 y
3.99
0.66
1.75 (66)
1.94
#308
53 y
6.75
1.72
1.52 (43)
4.46
#309
50 y
6.56
3.12
1.04 (10)
4.12
#315
48 y
5.92
1.81
1.01 (28)
4.10
#301
48 y
6.22
3.78
0.66 (<5)
3.86
#310
52 y
4.97
1.06
0.90 (21)
3.59
AGE
TCHOL
TRIG
HDL-C (%HDL)
LDL-C
D16S503 311 307
D16S3018 ? ?
CHST6(R162G) 0 1
D16S515 345 343
D16S516 251 257
D16S505 141 141
(309) (307)
2 3
0 1
329 343
247 257
(141) (141)
(307) (301)
? ?
(339) (333)
(251) (251)
(157) (157)
309 307
2 3
0 1
329 343
247 257
(141) (141)
307 307
? ?
0 0
343 331
247 253
155 157
309 307
2 3
0 1
329 343
247 257
141 141
309 301
? ?
1 0
337 337
251 255
155 141
309 307
? ?
0 0
345 337
251 253
141 141
311 307
? ?
0 0
345 337
251 253
141 139
307 307
? ?
0 0
343 333
251 257
161 155
309 307
2 3
0 1
329 343
247 257
141 141
311 301
? ?
(343) (343)
251 251
141 155
416
#403
40 y
5.03
1.19
0.99 (9)
3.50
#404
39 y
4.53
1.35
0.96 (28)
3.06
#405
26 y
4.55
0.53
1.57 (69)
2.74
#406
33 y
4.23
2.36
0.45 (<5)
2.72
#407
29 y
4.16
1.01
1.15 (20)
2.56
#408
32 y
3.68
2.43
0.61 (<5)
1.98
#409
31 y
4.01
1.64
0.93 (<5)
2.34
#412
25 y
3.95
1.56
1.01 (10)
2.24
#414
16 y
4.58
0.60
1.41 (82)
2.90
#401
26 y
3.43
1.17
0.59 (<5)
2.31
#402
24 y
3.87
1.61
1.00 (12)
2.15
D16S503 (301) (?)
D16S3018 ? ?
CHST6(R162G) ? ?
D16S515 (337) (?)
D16S516 (255) (?)
D16S505 (141) (?)
307 301
? ?
1 0
343 333
257 251
141 157
309 307
? ?
0 0
329 333
247 251
141 157
307 301
? ?
0 0
!!333 333 !!
257 251
!!155 157 !!
307 307
? ?
3 1
1 0
343 343
257 247
(141) (155)
307 307
? 1
3 1
0 0
329 343
247 253
141 157
307 309
? ?
1 1
343 337
(257) (251)
141 155
307 301
? ?
1 0
343 337
257 255
141 141
307 311
? ?
0 0
333 345
257 251
155 141
307 307
? ?
0 0
343 337
251 253
161 139
301 307
? ?
0 1
343 343
251 257
155 141
301 307
? ?
0 1
343 343
251 257
141 141
#504
14 y
4.99
1.53
0.95 (<5)
3.35
D16S503 301 301
D16S3018 ? ?
CHST6(R162G) 0 0
D16S515 337 333
D16S516 255 251
D16S505 157 141

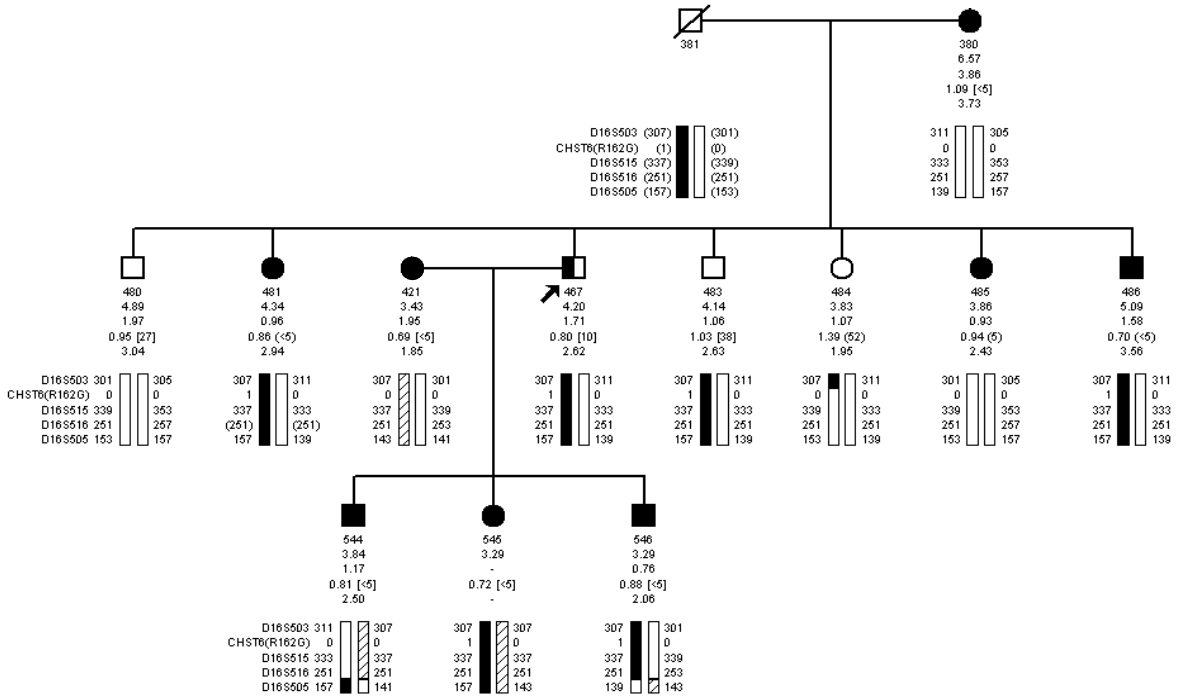
B. FAMILY 2


D16S503 (311) (313)
CHST6(R162G) ? ?
D16S515 (335) (337)
D16S516 (251) (255)
D16S505 (153) (151)
(311) (305)
(0) (1)
(337) (337)
(255) (251)
(157) (157)
203
TCHOL 4.96
TRIG 1.38
HDL-C (%HDL) 1.12 (17)
LDL-C 3.03
205
5.95
0.79
1.65 (59)
3.89
201
3.91
2.93
0.59 (<5)
1.92
202
D16S503 313 305
CHST6(R162G) 0 1
D16S515 337 337
D16S516 255 251
D16S505 151 157
313 311
0 0
337 337
255 255
151 157
311 311
0 1
335 337
251 251
153 157
(303) (311)
? ?
(335) (351)
(251) (251)
(157) (141)
302
6.20
1.41
1.41 (33)
4.15
301
3.34
2.95
0.62 (<5)
1.38
D16S503 311 311
CHST6(R162G) 0 0
D16S515 335 351
D16S516 251 251
D16S505 153 141
311 303
1 0
337 335
251 251
157 157

C. FAMILY 3


201
CHOL.T 4.39
TRIG 1.44
HDL-C (%HDL) 0.59 (<5)
LDL-C 3.15
D16S503 307 307
CHST6(R162G) 1 0
D16S515 343 337
D16S516 251 253
D16S505 141 155
202
(307) (?)
(0) (?)
(345) (?)
(247) (?)
(139) (?)
304
3.50
1.31
0.70 (<5)
2.21
D16S503 307 307
CHST6(R162G) 1 0
D16S515 343 345
D16S516 251 247
D16S505 141 139
301
3.98
3.67
0.47 (<5)
1.86
307 307
1 0
343 345
251 247
141 139
302
4.76
1.32
0.93 (25)
3.24
311 313
1 0
345 333
251 247
157 157
401
3.82
1.34
0.46 (<5)
2.76
D16S503 307 311
CHST6(R162G) 1 1
D16S515 343 345
D16S516 251 251
D16S505 (141) (157)

D. FAMILY 4


381
380
6.57
3.86
1.09 [<5]
3.73
D16S503 (307) (301)
CHST6(R162G) (1) (0)
D16S515 (337) (339)
D16S516 (251) (251)
D16S505 (157) (153)
480
4.89
1.97
0.95 [27]
3.04
481
4.34
0.96
0.86 (<5)
2.94
421
3.43
1.95
0.69 [<5]
1.85
467
4.20
1.71
0.80 [10]
2.62
483
4.14
1.06
1.03 [38]
2.63
484
3.83
1.07
1.39 (52)
1.95
485
3.86
0.93
0.94 (5)
2.43
486
5.09
1.58
0.70 (<5)
3.56
544
3.84
1.17
-
0.81 [<5]
2.50
545
3.29
-
0.72 [<5]
-
546
3.29
0.76
-
0.88 [<5]
2.06

6.9 Tables

Table 1. Subjects in this study

Subjects	QUE	SLSJ
Family		
# families	11	61
# individuals	412	410
Case/Control		
# cases	136	94
# controls	218	94

QUE: samples from French Canadian families from across Quebec
SLSJ: samples from the SLSJ region of Quebec

Table 2. Whole genome scan results with LOD score > 1.

Cohort	Chromosome	Location (cM)	LOD score
QUE	16	98	2.61
	15	8	1.68
	5	217	1.24
	1	209	1.08
SLSJ	16	86	2.96

Table 2. Non-synonymous variants identified by sequencing.

Gene	aa change	rs number	Physical Distance*
FA2H	Pro97Ala	rs35874850	73331496
RFWD3	Thr90Asn	rs8058922	73252580
RFWD3	Ile564Val	rs7193541	73222244
MLKL	Pro132Ser	rs35589326	73286763
MLKL	Met364Thr	rs34389205	73267111
CFDP1	Asn204Asp		73986529
BX648484	Met 44Ile		74043126
CHST6	Arg162Gly		74070744
CHST5	thr297met	rs3826107	74120831
CHST5	Leu 6Val	rs3743601	74147593
TERF21P	Lys324Glu	rs4888444	74247780
CASPR4	Leu279Val	rs34251012	75040248
CASPR4	Gln789Arg	rs12933808	75090084
CASPR4	Ala922Pro		75126951
CASPR4	Asp1158Glu	rs7202925	75144703
ADAMTS18	Tyr191His	rs11643211	75959046
ADAMTS18	Leu625Ile	rs11640912	75917420
ADAMTS18	Leu769Ile	rs9930984	75911474
ADAMTS18	Ala946Ser	rs12935394	75886491
ADAMTS18	Ser1080Arg	rs35478105	75882826
ADAMTS18	Ser1159Thr	rs3743749	75880736
NUDT7	Are183His	rs308925	76327335
CLEC3A	Gln197 Lys	rs2072663	76622234
WWOX	Ala179Thr	rs12918952	76978276
KIAA0431	Ser 600Thr		79635403
PKD1L2	Val20Ala	rs9924530	79811418
PKD1L2	Trp73Arg	rs9924371	79811260
PKD1L2	Gln120Leu	rs7191351	79807455
PKD1L2	Gly129Asp	rs7185774	79807428
PKD1L2	Lue173Ser	rs8060294	79806246
PKD1L2	Val183Ile	rs12933806	79806217
PKD1L2	Pro301Ala	rs11150370	79798601
PKD1L2	Lys416Gln	rs7194871	79790065
PKD1L2	Pro512Lue	rs7205673	79789776
PKD1L2	Leu711Pro	rs4889261	76770879
PKD1L2	Gly785Cys	rs9935113	79768997
PKD1L2	Arg849His	rs1869349	79766748
PKD1L2	Ala863Val	rs12596941	79766016
PKD1L2	Leu1036Pro	rs12597040	79757056
BCM01	Arg267Ser	rs12934922	79859195
BCM01	Ala379Val	rs7501331	79871997
PLCG2	His244Arg	rs11548656	80474413
PLCG2	Arg268Trp	rs17537869	80480314

*Physical distance is based on positions from the UCSC Genome Browser (http://genome.ucsc.edu) (Build 36.1 March 2006).

Supplementary Table1: List of differentially expressed genes using Affymetrix micro-array data in area under the linkage peak on chromosome 16.

Position based on Build 36.1

Gene	Position(bp)
CPSF5/ NUDT21	55020549-55042762
BBS2	55089850-55111696
MT1K/MT1M	55258154-55259478
NUP93	55373242-55436178
SLC12A3	55456620-55507263
CETP	55553336-55575257
NOD27	55625581-55674937
NIP30	55763664-55769970
TM4SF11	55847510-55876072
HSPC065	56841364-56875235
CMTM4	65206154-65288111
DNCLI2	65312300-65343026
APPBP1	65394413-65422282
RRAD	65513083-65516940
CBFB	65620551-65692459
CENPT	66419561-66425504
NFATC3	66676876-66818338
LYPLA3	66836748-66852462
SLC7A6	66855924-66893223
SMPD3	66949733-67039910
ZFP90	67131162-67158540
TCMO7	67435010-67676586
CIRH1A	67724000-67760438
VPS4A	67902788-67916446
CYB-5M	68016064-68055208
WWP2	68353775-68533144
PDXDC2	68567703-68634407
DDX19L	68938325-68964782
SF3B3	69115202-69163702
CALB2	69950127-69981843
ZNF23	70039004-70053618
PHLPPL	70240294-70315344
ATBF1	71374287-71639775
CFDP1	73885109-74024888
CHST5	74119929-74126569
GABARAPL2	74157750-74169280
MON1B	75782832-75792754
ADAMTS18	75873526-76026512

CHAPTER 7

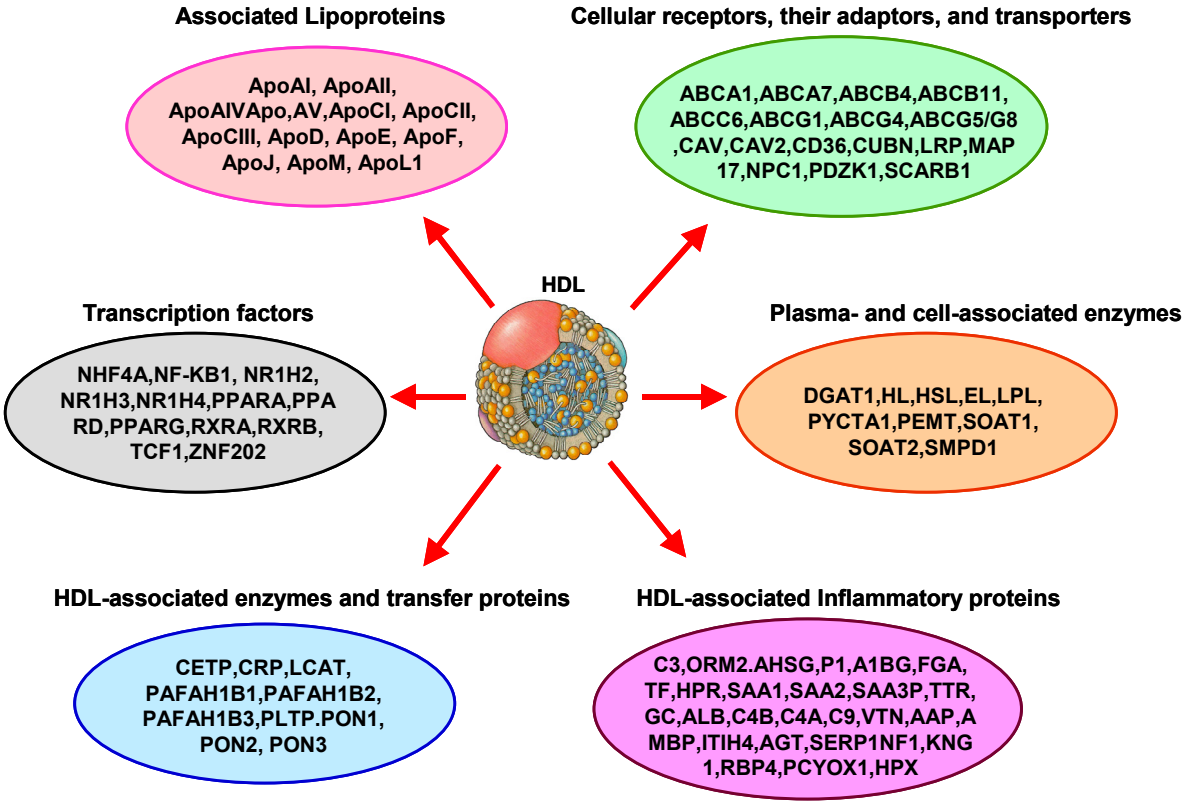
DISCUSSION AND CONCLUSION

HDL-C levels are inversely correlated with risk of CAD. The prevalence of low HDL-C in patients with CAD has been examined in several studies. To accentuate the importance of plasma lipoproteins in the pathogenesis of CAD and acute myocardial infarction, the largest case-control study of myocardial infarction (INTERHEART) has shown that the apo B/apo AI ratio (respectively an index of atherogenic lipoproteins and protective lipoproteins) accounts for approximately 49% of the population attributable risk of acute myocardial infarction[269]. It is estimated that approximately 40% of patients with premature CAD have HDL-C below the10th percentile of age- and sex-specific norms [190, 202, 208] and this represents the most common lipoprotein disorder in patients with CAD. The athero-protective effects of HDL are multifactorial and include: removing cholesterol from lipid-laden macrophages in the atherosclerotic plaque, anti-inflammatory effects, preventing oxidation of low-density lipoproteins (LDL), anti-thrombotic properties, vasomotor tone modulation and improving endothelial cell survival (by preventing apoptosis), migration and proliferation[270]. Nevertheless, the major athero-protective effect of HDL has been attributed to its key role in RCT, a process in which cholesterol from peripheral tissues such as lipid-laden macrophages (foam cells) is selectively returned to the liver for excretion in the bile. Mutations in any of the proteins regulating this complex metabolic pathway may potentially decrease HDL-C levels and accelerate CAD.

Although, a previously study reported that heterozygocity at the ABCA1 locus increases the risk of CAD 3.5 fold [271], in a study of the Copenhagen Heart study, Frikke-Schmidt examined the prevalence of rare variants at the ABCA1 gene locus in CAD cases compared with controls and concluded that ABCA1 mutations causing a low HDL-C were not increased in CAD cases [107]. This discrepancy can be explained in part by the ascertainment bias in former study that recruited patients from a premature CAD clinic. Thus these data challenge the existing concept that genetic forms of HDL deficiency increase cardiovascular risk.

Over 100 proteins have been genetically or biochemically demonstrated to be involved in HDL metabolism. These are summarized in **Figure 1**. Based on a proteomic study of HDL proteins, at least 56 different proteins have been identified, by a combination of one- and two-dimension electrophoresis, MALDI-TOF and isotope-coded affinity tag (ICAT)[267]. All of the apolipoproteins (with the exception of apo B) are found in HDL particles, as are several enzymes involved in lipoprotein processing, inflammation, free-radical enzymes and many others whose function is yet to be characterized. To these, must be added a nearly equal number of proteins, enzymes, receptors, and transcription factors known to be involved in HDL metabolism, but not found in HDL particles. Wang and Paigent reported 54 genes involved in HDL metabolism and function in the mouse; of these, apo AI, CETP, LCAT, EL, HL, LPL, ABCA1 and ABCG5/8 have been shown to regulate HDL in humans[100].

Figure1. Candidate Genes for HDL Metabolism and Functions
(Modified from Wang X and Paigen B Circ Res 2005;96:27; Rezaee F. *et al.* Proteinomics 2006;6:722 and Vaisar T *et al*. J Clin Invest 2007;117:746-756.)



Many studies have examined the role of genetic variability at many of these genes, either by DNA sequencing or by association studies and have shown that both rare and common variants in these genes affect the level of HDL-C (see review by Boes et al.[108])

Additionally, many studies confirmed that plasma levels of HDL-C have a strong inherited basis. It estimated that more than 50% of the variability in lipid and lipoprotein levels between individual is a result of the effect of genes. Linkage analysis has been very successfully used to discover the genetic basis of monogenic traits that segregate in a Mendelian manner. Such investigations for monogenic lipid disorders have also successfully identified many genes important for lipoprotein metabolism[272]. An example is the isolation of the ABCA1 gene as the causal gene for Tangier disease in patients who have extremely low HDL-C[101, 103, 104]. In the past few years, numerous family studies using parametric or quantitative linkage have revealed many loci harbouring genes involved in HDL metabolism. The apparent lack of consistency observed in linkage studies can be explained, in part, by the oligogenetic and multivariate nature of HDL regulation, but also by pleiotropy and variable ascertainment used for the selection of probands. Yet, most of the genes and mutations, underlying these findings remain to be defined.

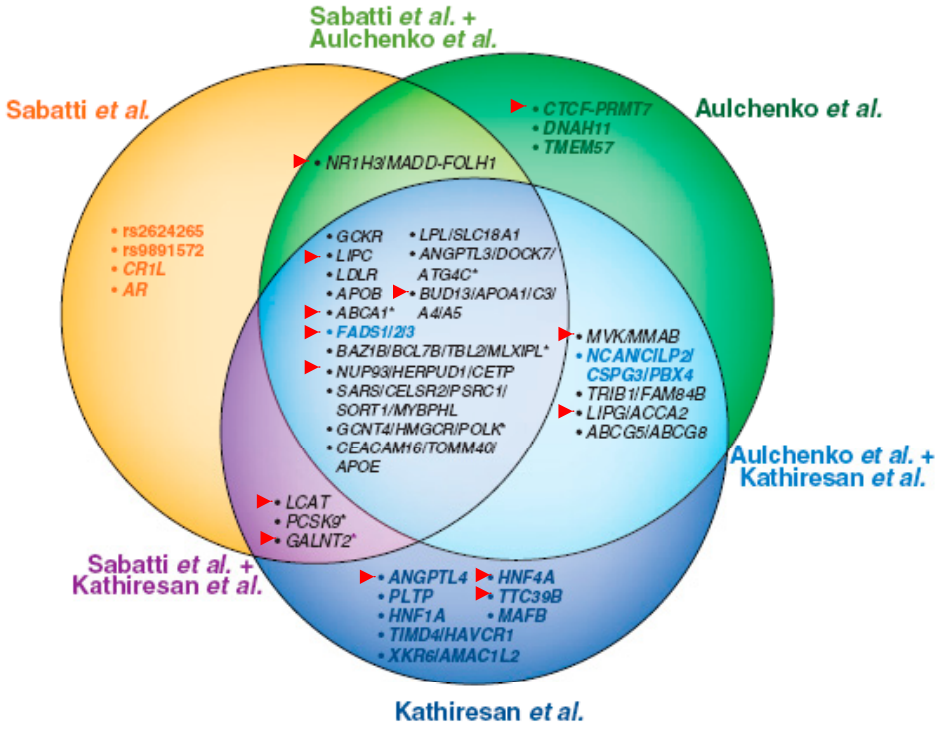
The previous genome wide scans have identified several chromosomal loci including a locus on chromosome 8q in Mexican Americans and Finns[92, 242]; 5q and 9p in Mexican Americans[92, 233]; a locus on 11q23 in 105 Utah CHD families[234]; and loci on 16q and 20q in Finns[175, 176].

In humans, at least 21 studies have been published, identifying over 45 QTLs for HDL-C. As some of the QTLs map to similar chromosomal localizations, it is possible that they represent the same genes. In their review, Wang and Paigen (2005) estimated that 30 QTL's with LOD score more than 1.4 for HDL-C have been identified in humans. Most of these QTLs (> 90%) are concordant with mice QTLs in syntenic regions. The high degree of concordance between mouse and human QTLs suggests a strong likelihood that the underlying genes will be the same. These authors therefore argue that an examination of HDL-C QTLs in mice will be a more efficient way to identify potential chromosomal regions of interest in the human genome[100].

Linkage studies have good statistical power for discovering less frequent alleles. However, association studies will have more power to discover common alleles. It has been proposed that a significant portion of the risk for complex trait, such as low HDL-C, is the result of common genetic variants (the common-disease common-variant (CDCV) hypothesis)[273]. Recently, GWAS approach has been successfully applied to identify several new, reproducible associations for HDL-C. The GWAS were successful in identifying many common variants in known genes or loci but also allowed the identification of new loci, presumably important in the regulation of lipoprotein metabolism. Many of those loci with common variants have been identified in genes that have been shown to also harbour rare mutations that lead to monogenic low HDL-C (or, in the case of the CETP gene, high HDL-C) including: ABCA1[101-103],ApoA5[274], ApoB[275], CETP[276], LDLR[275], HL[277], LPL[278]. Thus, the novel loci might also harbour rare variants

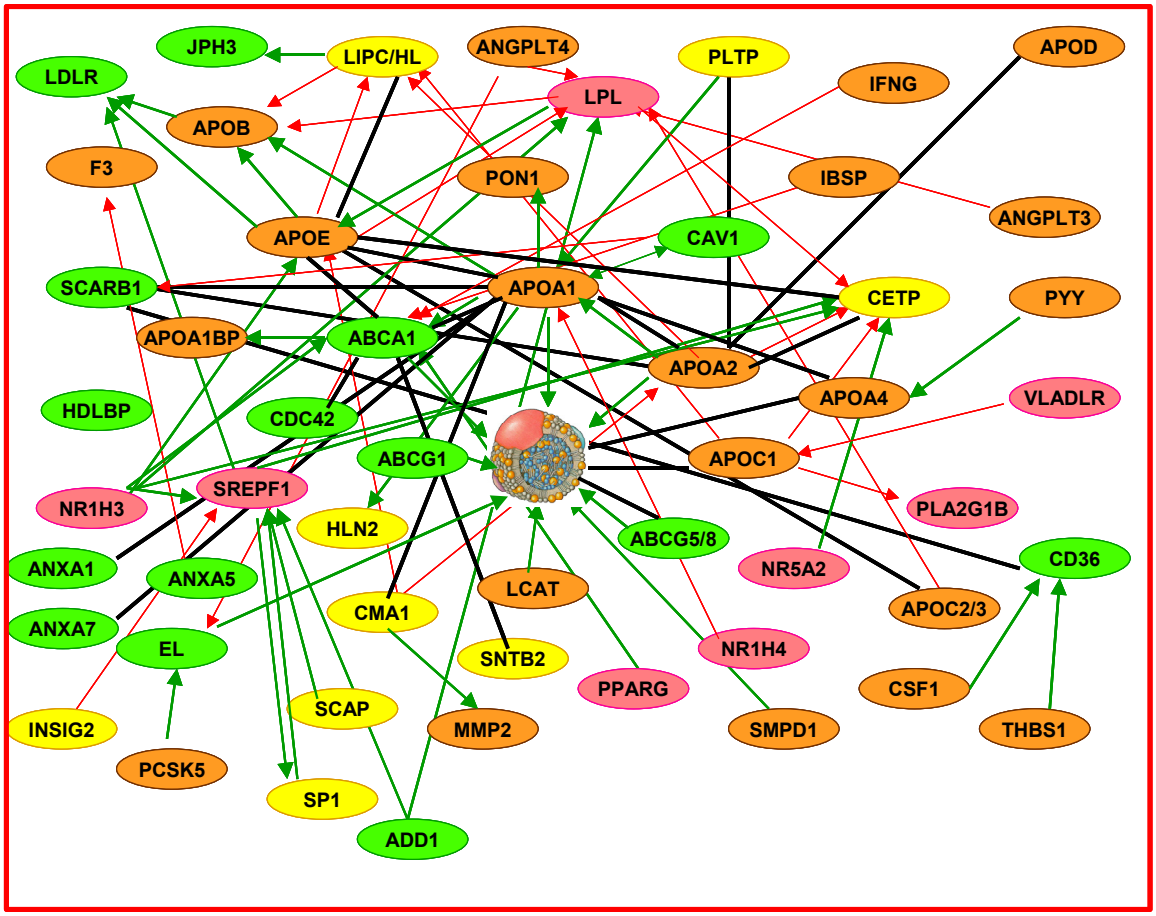
and be strong candidates for Mendelian dyslipidemias. By the GWAS approach, investigators have identified and/or confirmed 19 loci with common variants that influence plasma levels of LDL-C, TG, and HDL-C. **Figure 2** shows these loci. Of those, the nine novel loci that affect the level of plasma HDL-C are near the genes: GALANT2, MVK/ MMAB, NR1H3, MADD/ FOLH1, FAD1-3, ANGPLT4, HNF4A, TTC39B, CTCF-PRMT7. The associated SNPs at the new loci are all noncoding and explain only a minor fraction of variance in HDL-C level (up to2.42 mg/dl). So the causal variants or even the causal genes are not yet clear, but these new candidate genes and their nearby genes should certainly be used as clues to search for variants with major effect that can explain severe form of HDL-C deficiency in the families. To identify the causal variants deep sequencing of the genes at the new loci are required. Ultimately, functional studies involving positional candidate genes are essential to uncover the mechanism of these novel genetic factors that regulate HDL metabolism.

Figure2. Loci reported to be associated with lipid variables
(LDL-C, TG, and HDL-C). Modified from comments on GWAS by Teri A Manolio in *Nature Genetics* 41, 5 - 6 (2009).[279] Arrow heads show associated loci with HDL-C.



Despite the remarkable progress in gene identification, we are still faced with the major challenge of investigating the function of these genes. Additionally, network analysis could help elucidate the signalling pathways, gene regulatory networks, and protein interactions. The network for HDL genes as we know it today is shown in **Figure 3**.

Figure 3. Biological network for HDL.



Green arrows: positive effect; red arrows: negative effect; black lines show binding effect; cell localization of genes marked as yellow: cytoplasm, Orange: extra-cellular, pink: nucleus, green: cell membrane.
(http://www.jmdbase.jp/JmdBaseExt/Top.aspx)

As mentioned above, the known genes in HDL metabolism have been identified by the candidate gene approach, family linkage studies and more recently GWAS. However, because only a small percentage of the variance and of the heritability has been accounted for, it is clear that there are other genes yet to be identified those modulate HDL-C levels in man.

Because of the existence of pleiotropy and heterogeneity in the nature of HDL-C, which hampers the task of identifying new genes for this trait, I believe that choosing an isolated population could increase our chances to discover new genes. This is because an isolated population will have reduced heterogeneity and hence increased power to discover susceptibility genes [280, 281]. Therefore, our study has been conducted on the population of Quebec, which is descended from approximately 5,000 settlers from France, who arrived between 1608 and 1759. The advantages of limited number of settlers, a high birth rate, and relatively little genetic admixing because of language, cultural and religious barriers make the French Canadian population ideally suited for genetic studies.

To identify novel genetic determinants that contribute to HDL-C levels in the Quebec population, we focused our research on a sample of well-characterized families who were recruited for low HDL-C. I applied conventional candidate gene and family linkage approaches to pursue novel genes.

Firstly, searching the known candidate genes involved in the metabolism of HDL-C is a reasonable approach to identify variants that influence HDL-C levels. Several candidate gene association studies revealed variants that affect the level of HDL-C. However, due to allelic heterogeneity the genes need to be

sequenced to identify causative rare variants. Based on this idea, we sequenced well-known genes involved in HDL metabolism. Causal variants in ABCA1 were identified in thirteen of our probands[104, 106]. I turned to two other strong candidates, the ApoA1 and LCAT genes. Direct sequencing of genomic DNA revealed a novel mutation of the apoA-I gene, apoAI_{E136X}, identified in 3/54 patients (6%) with low HDL-C. In two available families, the E136X mutation showed a strong co-segregation with the phenotype of low HDL-C level. Interestingly, there are some individuals with low HDL-C in these families, who not bearing apoA-I_{E136X} mutation. This puts emphasis on the complexity of HDL that could be the results of phenocopy or heterogeneity in HDL. When added to the ABCA1 families, we have now accounted for 26% of the low HDL-C families in Quebec (assuming our families are representative). Afterwards, the identification of these genetic susceptibility to low HDL-C and the development of CAD may provide a more comprehensive understanding of the genetic basis for CAD. In turn, this has the potential to lead to the development of novel therapies for subjects with HDL deficiency.

The remaining families with no identified mutations in these candidate genes were subjects for further investigations to identify new gene(s) involved in the metabolism of HDL-C. Thus, we performed a genome scan in 13 extended multigenerational French Canadian families. In a parametric linkage analysis we identified a significant signal of linkage for low plasma HDL-C on chromosome 4q31.21. We also observed LOD scores > 1.0 for two loci on 15q and 16q. Because the only significant signal in this study was observed for 4q31.21, this

region was selected for further fine mapping. Following fine mapping, we sequenced all genes in LOD-1 region. However no evidence for a co-segregating variant was observed in our families. Harrap et al. in a study of healthy families from the general population in Australia also observed suggestive evidence for linkage of HDL-C to chromosome 4q32.3[282]. While their linkage peak and fine mapping analyses is approximately 30 cM further towards the q-telomere, it is in the vicinity of evidence for HDL-C in our samples. Further independent studies will be required to verify this locus.

Given that there is genetic heterogeneity underlying HDL-C and environmental influences also contribute to determining serum HDL-C levels, QTL analysis may be a more suitable approach to identify chromosomal regions that are associated with variation in HDL-C levels. Several QTLs have been identified for HDL-C. The inconsistency of the results can be explained by the difference in study design, measurement of HDL-C, ethnicity of population, and covariates used in the analysis. The only reproducible region in QTL analysis is located on chromosome 16q23-24. In fact, at least eight previous genome-wide scan studies have identified chromosomal loci for HDL-C on chromosome 16 with LOD scores greater than one[97, 176, 177, 180, 182].

In our QTL analysis, we obtained suggestive evidence for linkage on chromosome 16 from two French Canadian study samples. In our sample, four families demonstrated segregation over a 25.5 cM (18 Mb) region, which was further reduced to 6.6 Mb with additional markers. The coding regions of all genes within this region were sequenced. A missense variant in the gene CHST6,

segregated in all four families demonstrating segregation. However, an association study of this SNP in unrelated Quebec-wide samples was not significant but these results might be underpowered because of the small number of samples used in this association study. An association study in the SLSJ study sample identified a SNP (rs11646677) located in the same region, which was significantly associated with low HDL-C (p=0.016).

In addition, RT-PCR results from cultured cells demonstrated a significant difference in the expression level of the CHST6 and KIAA1576 genes in this region. Thus my data constitutes additional evidence for a locus on chromosome 16q23-24 that affects HDL-C levels in French Canadian subjects.

My research has lead to specific research approaches. Using next generation DNA sequencing the genomic sequence of this narrow region will be sequenced to identify the variants that are present in cases with low HDL-C and not present in a non-low HDL-C individual in the family. It is hoped that the association of those variants or particular haplotypes with low HDL-C in the families could explain the signal of linkage on chromosome 16q23-24.

In conclusion, I believe that by identifying casual SNPs for low HDL-C it may be possible to define individuals at increased risk for cardiovascular disease. Moreover, the identification of novel genes involved in HDL metabolism may help identify novel metabolic pathways and possibly provide novel therapeutic targets to modulate HDL-C in humans and test whether genotype-driven targeting of therapeutic interventions can effectively reduce their risk for heart disease.

REFRENCE LIST

(1) Castelli WP. Epidemiology of coronary heart disease: the Framingham study. *Am J Med* 1984 February 27;76(2A):4-12.

(2) Rosamond W, Flegal K, Furie K, Go A, Greenlund K, Haase N, Hailpern SM, Ho M, Howard V, Kissela B, Kittner S, Lloyd-Jones D, McDermott M, Meigs J, Moy C, Nichol G, O'Donnell C, Roger V, Sorlie P, Steinberger J, Thom T, Wilson M, Hong Y. Heart disease and stroke statistics--2008 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation* 2008 January 29;117(4):e25-146.

(3) Lusis AJ, Mar R, Pajukanta P. Genetics of atherosclerosis. *Annu Rev Genomics Hum Genet* 2004;5:189-218.

(4) Grundy SM. *Cholesterol and Atherosclerosis Diagnosis and Treatment*. Philadelphia: Lippincott; 1990.

(5) Davis RA. Lipoprotein structure and secretion. In: Vance DE, Vance JE, eds. *Biochemistry of lipids, lipoproteins and membranes*.New York: Elsevier; 1991. p. 403-26.

(6) Guerra R, Wang J, Grundy SM, Cohen JC. A hepatic lipase (LIPC) allele associated with high plasma concentrations of high density lipoprotein cholesterol. *Proc Natl Acad Sci U S A* 1997 April 29;94(9):4532-7.

(7) Myant NB. *Cholesterol metabolism, LDL, and the LDL receptor.* Totonto: Academic Press; 1990.

(8) Schneider WJ. Removal of lipoproteins from plasma. In: Vance DE, Vance JE, eds. *Biochemistry of lipids, lipoproteins, and membranes*.New York: Elsevier; 1996. p. 517-41.

(9) Goldstein JL, Brown MS, Anderson RG, Russell DW, Schneider WJ. Receptor-mediated endocytosis: concepts emerging from the LDL receptor system. *Annu Rev Cell Biol* 1985;1:1-39.

(10) Brown MS, Goldstein JL. A receptor-mediated pathway for cholesterol homeostasis. *Science* 1986 April 4;232(4746):34-47.

(11) Schneider WJ. The low density lipoprotein receptor. *Biochim Biophys Acta* 1989 May 9;988(2):303-17.

(12) Goldstein JL, Brown MS. The low-density lipoprotein pathway and its relation to atherosclerosis. *Annu Rev Biochem* 1977;46:897-930.

(13) Glomset JA. The plasma lecithins:cholesterol acyltransferase reaction. *J Lipid Res* 1968 March;9(2):155-67.

(14) Benditt EP, Eriksen N. Amyloid protein SAA is associated with high density lipoprotein from human serum. *Proc Natl Acad Sci U S A* 1977 September;74(9):4025-8.

(15) Lee NS, Brewer HB, Jr., Osborne JC, Jr. beta 2-Glycoprotein I. Molecular properties of an unusual apolipoprotein, apolipoprotein H. *J Biol Chem* 1983 April 25;258(8):4765-70.

(16) Vaisar T, Pennathur S, Green PS, Gharib SA, Hoofnagle AN, Cheung MC, Byun J, Vuletic S, Kassim S, Singh P, Chea H, Knopp RH, Brunzell J, Geary R, Chait A, Zhao XQ, Elkon K, Marcovina S, Ridker P, Oram JF, Heinecke JW. Shotgun proteomics implicates protease inhibition and complement activation in the antiinflammatory properties of HDL. *J Clin Invest* 2007 March;117(3):746-56.

(17) Lindgren FT, Elliott HA, Gofman JW. The ultracentrifugal characterization and isolation of human blood lipids and lipoproteins, with applications to the study of atherosclerosis. *J Phys Colloid Chem* 1951 January;55(1):80-93.

(18) Srivastava RA, Srivastava N. High density lipoprotein, apolipoprotein A-I, and coronary artery disease. *Mol Cell Biochem* 2000 June;209(1-2):131-44.

(19) Asztalos BF, Sloop CH, Wong L, Roheim PS. Two-dimensional electrophoresis of plasma lipoproteins: recognition of new apo A-I-containing subpopulations. *Biochim Biophys Acta* 1993 September 8;1169(3):291-300.

(20) Gordon T, Castelli WP, Hjortland MC, Kannel WB, Dawber TR. High density lipoprotein as a protective factor against coronary heart disease. The Framingham Study. *Am J Med* 1977 May;62(5):707-14.

(21) Kannel WB. High-density lipoproteins: epidemiologic profile and risks of coronary artery disease. *Am J Cardiol* 1983 August 22;52(4):9B-12B.

(22) Gordon DJ, Rifkind BM. High-density lipoprotein--the clinical implications of recent studies. *N Engl J Med* 1989 November 9;321(19):1311-6.

(23) Castelli WP, Anderson K, Wilson PW, Levy D. Lipids and risk of coronary heart disease. The Framingham Study. *Ann Epidemiol* 1992 January;2(1-2):23-8.

(24) Multiple risk factor intervention trial. Risk factor changes and mortality results. Multiple Risk Factor Intervention Trial Research Group. *JAMA* 1982 September 24;248(12):1465-77.

(25) Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). *JAMA* 2001 May 16;285(19):2486-97.

(26) De Backer G, Ambrosioni E, Borch-Johnsen K, Brotons C, Cifkova R, Dallongeville J, Ebrahim S, Faergeman O, Graham I, Mancia G, Cats

VM, Orth-Gomer K, Perk J, Pyorala K, Rodicio JL, Sans S, Sansoy V, Sechtem U, Silber S, Thomsen T, Wood D. European guidelines on cardiovascular disease prevention in clinical practice: third joint task force of European and other societies on cardiovascular disease prevention in clinical practice (constituted by representatives of eight societies and by invited experts). *Eur J Cardiovasc Prev Rehabil* 2003 August;10(4):S1-S10.

(27) Genest J, Frohlich J, Fodor G, McPherson R. Recommendations for the management of dyslipidemia and the prevention of cardiovascular disease: summary of the 2003 update. *CMAJ* 2003 October 28;169(9):921-4.

(28) Joy T, Hegele RA. Is raising HDL a futile strategy for atheroprotection? *Nat Rev Drug Discov* 2008 February;7(2):143-55.

(29) Pouliot MC, Despres JP, Nadeau A, Moorjani S, Prud'Homme D, Lupien PJ, Tremblay A, Bouchard C. Visceral obesity in men. Associations with glucose tolerance, plasma insulin, and lipoprotein levels. *Diabetes* 1992 July;41(7):826-34.

(30) Patsch JR, Karlin JB, Scott LW, Smith LC, Gotto AM, Jr. Inverse relationship between blood levels of high density lipoprotein subfraction 2 and magnitude of postprandial lipemia. *Proc Natl Acad Sci U S A* 1983 March;80(5):1449-53.

(31) Ballantyne FC, Clark RS, Simpson HS, Ballantyne D. High density and low density lipoprotein subfractions in survivors of myocardial infarction and in control subjects. *Metabolism* 1982 May;31(5):433-7.

(32) Asztalos BF, Roheim PS, Milani RL, Lefevre M, McNamara JR, Horvath KV, Schaefer EJ. Distribution of ApoA-I-containing HDL subpopulations in patients with coronary heart disease. *Arterioscler Thromb Vasc Biol* 2000 December;20(12):2670-6.

(33) Timmins JM, Lee JY, Boudyguina E, Kluckman KD, Brunham LR, Mulya A, Gebre AK, Coutinho JM, Colvin PL, Smith TL, Hayden MR, Maeda N, Parks JS. Targeted inactivation of hepatic Abca1 causes profound hypoalphalipoproteinemia and kidney hypercatabolism of apoA-I. *J Clin Invest* 2005 May;115(5):1333-42.

(34) Brunham LR, Kruit JK, Iqbal J, Fievet C, Timmins JM, Pape TD, Coburn BA, Bissada N, Staels B, Groen AK, Hussain MM, Parks JS, Kuipers F, Hayden MR. Intestinal ABCA1 directly contributes to HDL biogenesis in vivo. *J Clin Invest* 2006 April;116(4):1052-62.

(35) Nakamura K, Kennedy MA, Baldan A, Bojanic DD, Lyons K, Edwards PA. Expression and regulation of multiple murine ATP-binding cassette transporter G1 mRNAs/isoforms that stimulate cellular cholesterol efflux to high density lipoprotein. *J Biol Chem* 2004 October 29;279(44):45980-9.

(36) Wang N, Lan D, Chen W, Matsuura F, Tall AR. ATP-binding cassette transporters G1 and G4 mediate cellular cholesterol efflux to high-density lipoproteins. *Proc Natl Acad Sci U S A* 2004 June 29;101(26):9774-9.

(37) Yancey PG, Bortnick AE, Kellner-Weibel G, Llera-Moya M, Phillips MC, Rothblat GH. Importance of different pathways of cellular cholesterol efflux. *Arterioscler Thromb Vasc Biol* 2003 May 1;23(5):712-9.

(38) Marcil M, O'Connell B, Krimbou L, Genest J, Jr. High-density lipoproteins: multifunctional vanguards of the cardiovascular system. *Expert Rev Cardiovasc Ther* 2004 May;2(3):417-30.

(39) Navab M, Berliner JA, Subbanagounder G, Hama S, Lusis AJ, Castellani LW, Reddy S, Shih D, Shi W, Watson AD, Van Lenten BJ, Vora D, Fogelman AM. HDL and the inflammatory response induced

by LDL-derived oxidized phospholipids. *Arterioscler Thromb Vasc Biol* 2001 April;21(4):481-8.

(40) Shih DM, Xia YR, Wang XP, Miller E, Castellani LW, Subbanagounder G, Cheroutre H, Faull KF, Berliner JA, Witztum JL, Lusis AJ. Combined serum paraoxonase knockout/apolipoprotein E knockout mice exhibit increased lipoprotein oxidation and atherosclerosis. *J Biol Chem* 2000 June 9;275(23):17527-35.

(41) Tward A, Xia YR, Wang XP, Shi YS, Park C, Castellani LW, Lusis AJ, Shih DM. Decreased atherosclerotic lesion formation in human serum paraoxonase transgenic mice. *Circulation* 2002 July 23;106(4):484-90.

(42) Reddy ST, Wadleigh DJ, Grijalva V, Ng C, Hama S, Gangopadhyay A, Shih DM, Lusis AJ, Navab M, Fogelman AM. Human paraoxonase-3 is an HDL-associated enzyme with biological activity similar to paraoxonase-1 protein but is not regulated by oxidized lipids. *Arterioscler Thromb Vasc Biol* 2001 April;21(4):542-7.

(43) Arai H. Platelet-activating factor acetylhydrolase. *Prostaglandins Other Lipid Mediat* 2002 August;68-69:83-94.

(44) Osterud B, Bjorklid E. Role of monocytes in atherogenesis. *Physiol Rev* 2003 October;83(4):1069-112.

(45) Davenport P, Tipping PG. The role of interleukin-4 and interleukin-12 in the progression of atherosclerosis in apolipoprotein E-deficient mice. *Am J Pathol* 2003 September;163(3):1117-25.

(46) Carlos TM, Schwartz BR, Kovach NL, Yee E, Rosa M, Osborn L, Chi-Rosso G, Newman B, Lobb R, . Vascular cell adhesion molecule-1 mediates lymphocyte adherence to cytokine-activated cultured human endothelial cells. *Blood* 1990 September 1;76(5):965-70.

(47) Blankenberg S, Barbaux S, Tiret L. Adhesion molecules and atherosclerosis. *Atherosclerosis* 2003 October;170(2):191-203.

(48) Lawrence MB, Springer TA. Leukocytes roll on a selectin at physiologic flow rates: distinction from and prerequisite for adhesion through integrins. *Cell* 1991 May 31;65(5):859-73.

(49) Davies MJ, Gordon JL, Gearing AJ, Pigott R, Woolf N, Katz D, Kyriakopoulos A. The expression of the adhesion molecules ICAM-1, VCAM-1, PECAM, and E-selectin in human atherosclerosis. *J Pathol* 1993 November;171(3):223-9.

(50) Springer TA. Adhesion receptors of the immune system. *Nature* 1990 August 2;346(6283):425-34.

(51) Lesnik P, Chapman MJ. A new dimension in the vasculoprotective function of HDL: progenitor-mediated endothelium repair. *Arterioscler Thromb Vasc Biol* 2006 May;26(5):965-7.

(52) Barter PJ, Rye KA. High density lipoproteins and coronary heart disease. *Atherosclerosis* 1996 March;121(1):1-12.

(53) Barter PJ, Baker PW, Rye KA. Effect of high-density lipoproteins on the expression of adhesion molecules in endothelial cells. *Curr Opin Lipidol* 2002 June;13(3):285-8.

(54) Navab M, Imes SS, Hama SY, Hough GP, Ross LA, Bork RW, Valente AJ, Berliner JA, Drinkwater DC, Laks H, . Monocyte transmigration induced by modification of low density lipoprotein in cocultures of human aortic wall cells is due to induction of monocyte chemotactic protein 1 synthesis and is abolished by high density lipoprotein. *J Clin Invest* 1991 December;88(6):2039-46.

(55) Sugano M, Tsuchida K, Makino N. High-density lipoproteins protect endothelial cells from tumor necrosis factor-alpha-induced apoptosis. *Biochem Biophys Res Commun* 2000 June 16;272(3):872-6.

(56) Mineo C, Deguchi H, Griffin JH, Shaul PW. Endothelial and antithrombotic actions of HDL. *Circ Res* 2006 June 9;98(11):1352-64.

(57) Honda HM, Wakamatsu BK, Goldhaber JI, Berliner JA, Navab M, Weiss JN. High-density lipoprotein increases intracellular calcium levels by releasing calcium from internal stores in human endothelial cells. *Atherosclerosis* 1999 April;143(2):299-306.

(58) Kuvin JT, Ramet ME, Patel AR, Pandian NG, Mendelsohn ME, Karas RH. A novel mechanism for the beneficial vascular effects of high-density lipoprotein cholesterol: enhanced vasorelaxation and increased endothelial nitric oxide synthase expression. *Am Heart J* 2002 July;144(1):165-72.

(59) Yuhanna IS, Zhu Y, Cox BE, Hahner LD, Osborne-Lawrence S, Lu P, Marcel YL, Anderson RG, Mendelsohn ME, Hobbs HH, Shaul PW. High-density lipoprotein binding to scavenger receptor-BI activates endothelial nitric oxide synthase. *Nat Med* 2001 July;7(7):853-7.

(60) Levkau B, Hermann S, Theilmeier G, van der GM, Chun J, Schober O, Schafers M. High-density lipoprotein stimulates myocardial perfusion in vivo. *Circulation* 2004 November 23;110(21):3355-9.

(61) Saku K, Ahmad M, Glas-Greenwalt P, Kashyap ML. Activation of fibrinolysis by apolipoproteins of high density lipoproteins in man. *Thromb Res* 1985 July 1;39(1):1-8.

(62) Naqvi TZ, Shah PK, Ivey PA, Molloy MD, Thomas AM, Panicker S, Ahmed A, Cercek B, Kaul S. Evidence that high-density lipoprotein

cholesterol is an independent predictor of acute platelet-dependent thrombus formation. *Am J Cardiol* 1999 November 1;84(9):1011-7.

(63) Griffin JH, Kojima K, Banka CL, Curtiss LK, Fernandez JA. High-density lipoprotein enhancement of anticoagulant activities of plasma protein S and activated protein C. *J Clin Invest* 1999 January;103(2):219-27.

(64) Cohen JC, Kiss RS, Pertsemlidis A, Marcel YL, McPherson R, Hobbs HH. Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science* 2004 August 6;305(5685):869-72.

(65) Morton NE. Sequential tests for the detection of linkage. *Am J Hum Genet* 1955 September;7(3):277-318.

(66) Lander ES, Schork NJ. Genetic dissection of complex traits. *Science* 1994 September 30;265(5181):2037-48.

(67) Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, Boutin P, Vincent D, Belisle A, Hadjadj S, Balkau B, Heude B, Charpentier G, Hudson TJ, Montpetit A, Pshezhetsky AV, Prentki M, Posner BI, Balding DJ, Meyre D, Polychronakos C, Froguel P. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 2007 February 22;445(7130):881-5.

(68) McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, Cox DR, Hinds DA, Pennacchio LA, Tybjaerg-Hansen A, Folsom AR, Boerwinkle E, Hobbs HH, Cohen JC. A common allele on chromosome 9 associated with coronary heart disease. *Science* 2007 June 8;316(5830):1488-91.

(69) Barrett JC, Hansoul S, Nicolae DL, Cho JH, Duerr RH, Rioux JD, Brant SR, Silverberg MS, Taylor KD, Barmada MM, Bitton A, Dassopoulos T, Datta LW, Green T, Griffiths AM, Kistner EO, Murtha MT, Regueiro

MD, Rotter JI, Schumm LP, Steinhart AH, Targan SR, Xavier RJ, Libioulle C, Sandor C, Lathrop M, Belaiche J, Dewit O, Gut I, Heath S, Laukens D, Mni M, Rutgeerts P, Van GA, Zelenika D, Franchimont D, Hugot JP, de VM, Vermeire S, Louis E, Cardon LR, Anderson CA, Drummond H, Nimmo E, Ahmad T, Prescott NJ, Onnie CM, Fisher SA, Marchini J, Ghori J, Bumpstead S, Gwilliam R, Tremelling M, Deloukas P, Mansfield J, Jewell D, Satsangi J, Mathew CG, Parkes M, Georges M, Daly MJ. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet* 2008 August;40(8):955-62.

(70) Feinleib M, Garrison RJ, Fabsitz R, Christian JC, Hrubec Z, Borhani NO, Kannel WB, Rosenman R, Schwartz JT, Wagner JO. The NHLBI twin study of cardiovascular disease risk factors: methodology and summary of results. *Am J Epidemiol* 1977 October;106(4):284-5.

(71) Duccini Dal Colletto GM, Krieger H, Magalhaes JR. Estimates of the genetical and environmental determinants of serum lipid and lipoprotein concentrations in Brazilian twins. *Hum Hered* 1981;31(4):232-7.

(72) Dahlen G, Ericson C, de Faire U, Iselius L, Lundman T. Genetic and environmental determinants of cholesterol and HDL-cholesterol concentrations in blood. *Int J Epidemiol* 1983 March;12(1):32-5.

(73) Whitfield JB, Martin NG. Plasma lipids in twins. Environmental and genetic influences. *Atherosclerosis* 1983 September;48(3):265-77.

(74) Austin MA, King MC, Bawol RD, Hulley SB, Friedman GD. Risk factors for coronary heart disease in adult female twins. Genetic heritability and shared environmental influences. *Am J Epidemiol* 1987 February;125(2):308-18.

(75) O'Connell DL, Heller RF, Roberts DC, Allen JR, Knapp JC, Steele PL, Silove D. Twin study of genetic and environmental effects on lipid levels. *Genet Epidemiol* 1988;5(5):323-41.

(76) Hunt SC, Hasstedt SJ, Kuida H, Stults BM, Hopkins PN, Williams RR. Genetic heritability and common environmental components of resting and stressed blood pressures, lipids, and body mass index in Utah pedigrees and twins. *Am J Epidemiol* 1989 March;129(3):625-38.

(77) Heller DA, de Faire U, Pedersen NL, Dahlen G, McClearn GE. Genetic and environmental influences on serum lipid levels in twins. *N Engl J Med* 1993 April 22;328(16):1150-6.

(78) Knoblauch H, Busjahn A, Munter S, Nagy Z, Faulhaber HD, Schuster H, Luft FC. Heritability analysis of lipids and three gene loci in twins link the macrophage scavenger receptor to HDL cholesterol concentrations. *Arterioscler Thromb Vasc Biol* 1997 October;17(10):2054-60.

(79) Bucher KD, Friedlander Y, Kaplan EB, Namboodiri KK, Kark JD, Eisenberg S, Stein Y, Rifkind BM. Biological and cultural sources of familial resemblance in plasma lipids: a comparison between North America and Israel--the Lipid Research Clinics Program. *Genet Epidemiol* 1988;5(1):17-33.

(80) Edwards KL, Mahaney MC, Motulsky AG, Austin MA. Pleiotropic genetic effects on LDL size, plasma triglyceride, and HDL cholesterol in families. *Arterioscler Thromb Vasc Biol* 1999 October;19(10):2456-64.

(81) Hamsten A, Iselius L, Dahlen G, de Faire U. Genetic and cultural inheritance of serum lipids, low and high density lipoprotein cholesterol and serum apolipoproteins A-I, A-II and B. *Atherosclerosis* 1986 June;60(3):199-208.

(82) Iselius L, Carlson LA, Morton NE, Efendic S, Lindsten J, Luft R. Genetic and environmental determinants for lipoprotein concentrations in blood. *Acta Med Scand* 1985;217(2):161-70.

(83) Namboodiri KK, Kaplan EB, Heuch I, Elston RC, Green PP, Rao DC, Laskarzewski P, Glueck CJ, Rifkind BM. The Collaborative Lipid Research Clinics Family Study: biological and cultural determinants of familial resemblance for plasma lipids and lipoproteins. *Genet Epidemiol* 1985;2(3):227-54.

(84) Peacock JM, Arnett DK, Atwood LD, Myers RH, Coon H, Rich SS, Province MA, Heiss G. Genome scan for quantitative trait loci linked to high-density lipoprotein cholesterol: The NHLBI Family Heart Study. *Arterioscler Thromb Vasc Biol* 2001 November;21(11):1823-8.

(85) Perusse L, Despres JP, Tremblay A, Leblanc C, Talbot J, Allard C, Bouchard C. Genetic and environmental determinants of serum lipids and lipoproteins in French Canadian families. *Arteriosclerosis* 1989 May;9(3):308-18.

(86) Perusse L, Rice T, Despres JP, Bergeron J, Province MA, Gagnon J, Leon AS, Rao DC, Skinner JS, Wilmore JH, Bouchard C. Familial resemblance of plasma lipids, lipoproteins and postheparin lipoprotein and hepatic lipases in the HERITAGE Family Study. *Arterioscler Thromb Vasc Biol* 1997 November;17(11):3263-9.

(87) Rao DC, Laskarzewski PM, Morrison JA, Khoury P, Kelly K, Wette R, Russell J, Glueck CJ. The Cincinnati Lipid Research Clinic family study: cultural and biological determinants of lipids and lipoprotein concentrations. *Am J Hum Genet* 1982 November;34(6):888-903.

(88) Rice T, Vogler GP, Perry TS, Laskarzewski PM, Rao DC. Familial aggregation of lipids and lipoproteins in families ascertained through

random and nonrandom probands in the Iowa Lipid Research Clinics family study. *Hum Hered* 1991;41(2):107-21.

(89) Steinmetz J, Boerwinkle E, Gueguen R, Visvikis S, Henny J, Siest G. Multivariate genetic analysis of high density lipoprotein particles. *Atherosclerosis* 1992 February;92(2-3):219-27.

(90) Vogler GP, Wette R, Laskarzewski PM, Perry TS, Rice T, Province MA, Rao DC. Heterogeneity in the biological and cultural determinants of high-density lipoprotein cholesterol in five North American populations: the Lipid Research Clinics Family Study. *Hum Hered* 1989;39(5):249-57.

(91) Mahaney MC, Blangero J, Rainwater DL, Comuzzie AG, VandeBerg JL, Stern MP, MacCluer JW, Hixson JE. A major locus influencing plasma high-density lipoprotein cholesterol levels in the San Antonio Family Heart Study. Segregation and linkage analyses. *Arterioscler Thromb Vasc Biol* 1995 October;15(10):1730-9.

(92) Almasy L, Hixson JE, Rainwater DL, Cole S, Williams JT, Mahaney MC, VandeBerg JL, Stern MP, MacCluer JW, Blangero J. Human pedigree-based quantitative-trait-locus mapping: localization of two genes influencing HDL-cholesterol metabolism. *Am J Hum Genet* 1999 June;64(6):1686-93.

(93) Byard PJ, Borecki IB, Glueck CJ, Laskarzewski PM, Third JL, Rao DC. A genetic study of hypoalphalipoproteinemia. *Genet Epidemiol* 1984;1(1):43-51.

(94) Friedlander Y, Kark JD, Stein Y. Complex segregation analysis of low levels of plasma high-density lipoprotein cholesterol in a sample of nuclear families in Jerusalem. *Genet Epidemiol* 1986;3(5):285-97.

(95) Kronenberg F, Coon H, Ellison RC, Borecki I, Arnett DK, Province MA, Eckfeldt JH, Hopkins PN, Hunt SC. Segregation analysis of HDL cholesterol in the NHLBI Family Heart Study and in Utah pedigrees. *Eur J Hum Genet* 2002 June;10(6):367-74.

(96) Karlin S, Williams PT, Haskell WL, Wood PD. Genetic analysis of the Stanford LRC family study data. II. Structured exploratory data analysis of lipids and lipoproteins. *Am J Epidemiol* 1981 March;113(3):325-37.

(97) Amos CI, Wilson AF, Rosenbaum PA, Srinivasan SR, Webber LS, Elston RC, Berenson GS. An approach to the multivariate analysis of high-density-lipoprotein cholesterol in a large kindred: the Bogalusa Heart Study. *Genet Epidemiol* 1986;3(4):255-67.

(98) Feitosa MF, Province MA, Heiss G, Arnett DK, Myers RH, Pankow JS, Hopkins PN, Borecki IB. Evidence of QTL on 15q21 for high-density lipoprotein cholesterol: the National Heart, Lung, and Blood Institute Family Heart Study (NHLBI FHS). *Atherosclerosis* 2007 January;190(1):232-7.

(99) Cupples LA, Myers RH. Segregation analysis for high density lipoprotein in the Berkeley data. *Genet Epidemiol* 1993;10(6):629-34.

(100) Wang X, Paigen B. Genetics of variation in HDL cholesterol in humans and mice. *Circ Res* 2005 January 7;96(1):27-42.

(101) Brooks-Wilson A, Marcil M, Clee SM, Zhang LH, Roomp K, van Dam M, Yu L, Brewer C, Collins JA, Molhuizen HO, Loubser O, Ouelette BF, Fichter K, Ashbourne-Excoffon KJ, Sensen CW, Scherer S, Mott S, Denis M, Martindale D, Frohlich J, Morgan K, Koop B, Pimstone S, Kastelein JJ, Genest J, Jr., Hayden MR. Mutations in ABC1 in Tangier disease and familial high-density lipoprotein deficiency. *Nat Genet* 1999 August;22(4):336-45.

(102) Bodzioch M, Orso E, Klucken J, Langmann T, Bottcher A, Diederich W, Drobnik W, Barlage S, Buchler C, Porsch-Ozcurumez M, Kaminski WE, Hahmann HW, Oette K, Rothe G, Aslanidis C, Lackner KJ, Schmitz G. The gene encoding ATP-binding cassette transporter 1 is mutated in Tangier disease. *Nat Genet* 1999 August;22(4):347-51.

(103) Rust S, Rosier M, Funke H, Real J, Amoura Z, Piette JC, Deleuze JF, Brewer HB, Duverger N, Denefle P, Assmann G. Tangier disease is caused by mutations in the gene encoding ATP-binding cassette transporter 1. *Nat Genet* 1999 August;22(4):352-5.

(104) Marcil M, Brooks-Wilson A, Clee SM, Roomp K, Zhang LH, Yu L, Collins JA, van Dam M, Molhuizen HO, Loubster O, Ouellette BF, Sensen CW, Fichter K, Mott S, Denis M, Boucher B, Pimstone S, Genest J, Jr., Kastelein JJ, Hayden MR. Mutations in the ABC1 gene in familial HDL deficiency with defective cholesterol efflux. *Lancet* 1999 October 16;354(9187):1341-6.

(105) Frikke-Schmidt R, Nordestgaard BG, Jensen GB, Tybjaerg-Hansen A. Genetic variation in ABC transporter A1 contributes to HDL cholesterol in the general population. *J Clin Invest* 2004 November;114(9):1343-53.

(106) Alrasadi K, Ruel IL, Marcil M, Genest J. Functional mutations of the ABCA1 gene in subjects of French-Canadian descent with HDL deficiency. *Atherosclerosis* 2005 December 7.

(107) Frikke-Schmidt R, Nordestgaard BG, Stene MC, Sethi AA, Remaley AT, Schnohr P, Grande P, Tybjaerg-Hansen A. Association of loss-of-function mutations in the ABCA1 gene with high-density lipoprotein cholesterol levels and risk of ischemic heart disease. *JAMA* 2008 June 4;299(21):2524-32.

(108) Boes E, Coassin S, Kollerits B, Heid IM, Kronenberg F. Genetic-epidemiological evidence on genes associated with HDL cholesterol levels: A systematic in-depth review. *Exp Gerontol* 2008 November 17.

(109) Zannis VI, Karathanasis SK, Keutmann HT, Goldberger G, Breslow JL. Intracellular and extracellular processing of human apolipoprotein A-I: secreted apolipoprotein A-I isoprotein 2 is a propeptide. *Proc Natl Acad Sci U S A* 1983 May;80(9):2574-8.

(110) Ajees AA, Anantharamaiah GM, Mishra VK, Hussain MM, Murthy HM. Crystal structure of human apolipoprotein A-I: insights into its protective effect against cardiovascular diseases. *Proc Natl Acad Sci U S A* 2006 February 14;103(7):2126-31.

(111) Catte A, Patterson JC, Jones MK, Jerome WG, Bashtovyy D, Su Z, Gu F, Chen J, Aliste MP, Harvey SC, Li L, Weinstein G, Segrest JP. Novel changes in discoidal high density lipoprotein morphology: a molecular dynamics study. *Biophys J* 2006 June 15;90(12):4345-60.

(112) Jonas A. Lecithin cholesterol acyltransferase. In: Gotto AM, ed. *Plasma Lipoproteins.*Amsterdam: Elsevier; 1987. p. 299-333.

(113) Denis M, Haidar B, Marcil M, Bouvier M, Krimbou L, Genest J, Jr. Molecular and cellular physiology of apolipoprotein A-I lipidation by the ATP-binding cassette transporter A1 (ABCA1). *J Biol Chem* 2004 February 27;279(9):7384-94.

(114) McLachlan AD. Repeated helical pattern in apolipoprotein-A-I. *Nature* 1977 June 2;267(5610):465-6.

(115) Luo CC, Li WH, Moore MN, Chan L. Structure and evolution of the apolipoprotein multigene family. *J Mol Biol* 1986 February 5;187(3):325-40.

(116) Sorci-Thomas M, Kearns MW, Lee JP. Apolipoprotein A-I domains involved in lecithin-cholesterol acyltransferase activation. Structure:function relationships. *J Biol Chem* 1993 October 5;268(28):21403-9.

(117) Roberts LM, Ray MJ, Shih TW, Hayden E, Reader MM, Brouillette CG. Structural analysis of apolipoprotein A-I: limited proteolysis of methionine-reduced and -oxidized lipid-free and lipid-bound human apo A-I. *Biochemistry* 1997 June 17;36(24):7615-24.

(118) Rogers DP, Brouillette CG, Engler JA, Tendian SW, Roberts L, Mishra VK, Anantharamaiah GM, Lund-Katz S, Phillips MC, Ray MJ. Truncation of the amino terminus of human apolipoprotein A-I substantially alters only the lipid-free conformation. *Biochemistry* 1997 January 14;36(2):288-300.

(119) Burgess JW, Frank PG, Franklin V, Liang P, McManus DC, Desforges M, Rassart E, Marcel YL. Deletion of the C-terminal domain of apolipoprotein A-I impairs cell surface binding and lipid efflux in macrophage. *Biochemistry* 1999 November 2;38(44):14524-33.

(120) Sorci-Thomas MG, Thomas MJ. The effects of altered apolipoprotein A-I structure on plasma HDL concentration. *Trends Cardiovasc Med* 2002 April;12(3):121-8.

(121) Freeman DJ, Griffin BA, Holmes AP, Lindsay GM, Gaffney D, Packard CJ, Shepherd J. Regulation of plasma HDL cholesterol and subfraction distribution by genetic and environmental factors. Associations between the TaqI B RFLP in the CETP gene and smoking and obesity. *Arterioscler Thromb* 1994 March;14(3):336-44.

(122) Takata K, Saku K, Ohta T, Takata M, Bai H, Jimi S, Liu R, Sato H, Kajiyama G, Arakawa K. A new case of apoA-I deficiency showing codon 8 nonsense mutation of the apoA-I gene without evidence of

coronary heart disease. *Arterioscler Thromb Vasc Biol* 1995 November;15(11):1866-74.

(123) Romling R, von EA, Funke H, Motti C, Fragiacomo GC, Noseda G, Assmann G. A nonsense mutation in the apolipoprotein A-I gene is associated with high-density lipoprotein deficiency and periorbital xanthelasmas. *Arterioscler Thromb* 1994 December;14(12):1915-22.

(124) Matsunaga T, Hiasa Y, Yanagi H, Maeda T, Hattori N, Yamakawa K, Yamanouchi Y, Tanaka I, Obara T, Hamaguchi H. Apolipoprotein A-I deficiency due to a codon 84 nonsense mutation of the apolipoprotein A-I gene. *Proc Natl Acad Sci U S A* 1991 April 1;88(7):2793-7.

(125) Nakata K, Kobayashi K, Yanagi H, Shimakura Y, Tsuchiya S, Arinami T, Hamaguchi H. Autosomal dominant hypoalphalipoproteinemia due to a completely defective apolipoprotein A-I gene. *Biochem Biophys Res Commun* 1993 October 29;196(2):950-5.

(126) Pisciotta L, Miccoli R, Cantafora A, Calabresi L, Tarugi P, Alessandrini P, Bittolo BG, Franceschini G, Cortese C, Calandra S, Bertolini S. Recurrent mutations of the apolipoprotein A-I gene in three kindreds with severe HDL deficiency. *Atherosclerosis* 2003 April;167(2):335-45.

(127) Moriyama K, Sasaki J, Takada Y, Matsunaga A, Fukui J, Albers JJ, Arakawa K. A cysteine-containing truncated apo A-I variant associated with HDL deficiency. *Arterioscler Thromb Vasc Biol* 1996 December;16(12):1416-23.

(128) Ikewaki K, Matsunaga A, Han H, Watanabe H, Endo A, Tohyama J, Kuno M, Mogi J, Sugimoto K, Tada N, Sasaki J, Mochizuki S. A novel two nucleotide deletion in the apolipoprotein A-I gene, apoA-I Shinbashi, associated with high density lipoprotein deficiency, corneal opacities, planar xanthomas, and premature coronary artery disease. *Atherosclerosis* 2004 January;172(1):39-45.

(129) Funke H, von EA, Pritchard PH, Karas M, Albers JJ, Assmann G. A frameshift mutation in the human apolipoprotein A-I gene causes high density lipoprotein deficiency, partial lecithin: cholesterol-acyltransferase deficiency, and corneal opacities. *J Clin Invest* 1991 January;87(1):371-6.

(130) Sirtori CR, Calabresi L, Franceschini G, Baldassarre D, Amato M, Johansson J, Salvetti M, Monteduro C, Zulli R, Muiesan ML, gabiti-Rosei E. Cardiovascular status of carriers of the apolipoprotein A-I(Milano) mutant: the Limone sul Garda study. *Circulation* 2001 April 17;103(15):1949-54.

(131) Jeenah M, Kessling A, Miller N, Humphries S. G to A substitution in the promoter region of the apolipoprotein AI gene is associated with elevated serum apolipoprotein AI and high density lipoprotein cholesterol concentrations. *Mol Biol Med* 1990 June;7(3):233-41.

(132) Brousseau ME. Common variation in genes involved in HDL metabolism influences coronary heart disease risk at the population level. *Rev Endocr Metab Disord* 2004 December;5(4):343-9.

(133) Evans D, Beil FU. The association of the R219K polymorphism in the ATP-binding cassette transporter 1 (ABCA1) gene with coronary heart disease and hyperlipidaemia. *J Mol Med* 2003 April;81(4):264-70.

(134) Kuivenhoven JA, Pritchard H, Hill J, Frohlich J, Assmann G, Kastelein J. The molecular pathology of lecithin:cholesterol acyltransferase (LCAT) deficiency syndromes. *J Lipid Res* 1997 February;38(2):191-205.

(135) Kuivenhoven JA, Groenemeyer BE, Boer JM, Reymer PW, Berghuis R, Bruin T, Jansen H, Seidell JC, Kastelein JJ. Ser447stop mutation in lipoprotein lipase is associated with elevated HDL cholesterol levels in

normolipidemic males. *Arterioscler Thromb Vasc Biol* 1997 March;17(3):595-9.

(136) Kasid A, Rhyne J, Zeller K, Pritchard H, Miller M. A novel TC deletion resulting in Pro(260)-->Stop in the human LCAT gene is associated with a dominant effect on HDL-cholesterol. *Atherosclerosis* 2001 May;156(1):127-32.

(137) Winder AF, Owen JS, Pritchard PH, Lloyd-Jones D, Vallance DT, White P, Wray R. A first British case of fish-eye disease presenting at age 75 years: a double heterozygote for defined and new mutations affecting LCAT structure and expression. *J Clin Pathol* 1999 March;52(3):228-30.

(138) Funke H, von EA, Pritchard PH, Albers JJ, Kastelein JJ, Droste C, Assmann G. A molecular defect causing fish eye disease: an amino acid exchange in lecithin-cholesterol acyltransferase (LCAT) leads to the selective loss of alpha-LCAT activity. *Proc Natl Acad Sci U S A* 1991 June 1;88(11):4855-9.

(139) Kuivenhoven JA, Stalenhoef AF, Hill JS, Demacker PN, Errami A, Kastelein JJ, Pritchard PH. Two novel molecular defects in the LCAT gene are associated with fish eye disease. *Arterioscler Thromb Vasc Biol* 1996 February;16(2):294-303.

(140) Pare G, Serre D, Brisson D, Anand SS, Montpetit A, Tremblay G, Engert JC, Hudson TJ, Gaudet D. Genetic analysis of 103 candidate genes for coronary artery disease and associated phenotypes in a founder population reveals a new association between endothelin-1 and high-density lipoprotein cholesterol. *Am J Hum Genet* 2007 April;80(4):673-82.

(141) Zhang K, Zhang S, Zheng K, He Y, Zhang L, Su Z, Sun Y, Shi J, Kong X, Tong Y. [Study on the association of lecithin cholesterol

acyltransferase gene polymorphisms with the lipid metabolism in coronary atherosclerotic heart disease]. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* 2003 April;20(2):135-7.

(142) Carlquist J, Anderson JL. Inconsistencies in the genetic prediction of HDL cholesterol versus atherosclerosis. *Curr Opin Cardiol* 2007 July;22(4):352-8.

(143) Matsuura F, Wang N, Chen W, Jiang XC, Tall AR. HDL from CETP-deficient subjects shows enhanced ability to promote cholesterol efflux from macrophages in an apoE- and ABCG1-dependent pathway. *J Clin Invest* 2006 May;116(5):1435-42.

(144) Agerholm-Larsen B, Tybjaerg-Hansen A, Schnohr P, Steffensen R, Nordestgaard BG. Common cholesteryl ester transfer protein mutations, decreased HDL cholesterol, and possible decreased risk of ischemic heart disease: The Copenhagen City Heart Study. *Circulation* 2000 October 31;102(18):2197-203.

(145) Dachet C, Poirier O, Cambien F, Chapman J, Rouis M. New functional promoter polymorphism, CETP/-629, in cholesteryl ester transfer protein (CETP) gene related to CETP mass and high density lipoprotein cholesterol levels: role of Sp1/Sp3 in transcriptional regulation. *Arterioscler Thromb Vasc Biol* 2000 February;20(2):507-15.

(146) Le GW, Guerin M, Nicaud V, Dachet C, Luc G, Arveiler D, Ruidavets JB, Evans A, Kee F, Morrison C, Chapman MJ, Thillet J. A novel cholesteryl ester transfer protein promoter polymorphism (-971G/A) associated with plasma high-density lipoprotein cholesterol levels. Interaction with the TaqIB and -629C/A polymorphisms. *Atherosclerosis* 2002 April;161(2):269-79.

(147) Lu H, Inazu A, Moriyama Y, Higashikata T, Kawashiri MA, Yu W, Huang Z, Okamura T, Mabuchi H. Haplotype analyses of cholesteryl

ester transfer protein gene promoter: a clue to an unsolved mystery of TaqIB polymorphism. *J Mol Med* 2003 April;81(4):246-55.

(148) Thompson JF, Durham LK, Lira ME, Shear C, Milos PM. CETP polymorphisms associated with HDL cholesterol may differ from those associated with cardiovascular disease. *Atherosclerosis* 2005 July;181(1):45-53.

(149) Thompson A, Di AE, Sarwar N, Erqou S, Saleheen D, Dullaart RP, Keavney B, Ye Z, Danesh J. Association of cholesteryl ester transfer protein genotypes with CETP mass and activity, lipid levels, and coronary risk. *JAMA* 2008 June 18;299(23):2777-88.

(150) Lewis GF, Rader DJ. New insights into the regulation of HDL metabolism and reverse cholesterol transport. *Circ Res* 2005 June 24;96(12):1221-32.

(151) Blades B, Vega GL, Grundy SM. Activities of lipoprotein lipase and hepatic triglyceride lipase in postheparin plasma of patients with low concentrations of HDL cholesterol. *Arterioscler Thromb* 1993 August;13(8):1227-35.

(152) Merkel M, Eckel RH, Goldberg IJ. Lipoprotein lipase: genetics, lipid uptake, and regulation. *J Lipid Res* 2002 December;43(12):1997-2006.

(153) Wittrup HH, Andersen RV, Tybjaerg-Hansen A, Jensen GB, Nordestgaard BG. Combined analysis of six lipoprotein lipase genetic variants on triglycerides, high-density lipoprotein, and ischemic heart disease: cross-sectional, prospective, and case-control studies from the Copenhagen City Heart Study. *J Clin Endocrinol Metab* 2006 April;91(4):1438-45.

(154) Lanktree MB, Anand SS, Yusuf S, Hegele RA. Replication of genetic associations with plasma lipoproteins and triglycerides in a multi-ethnic

sample using a targeted cardiovascular SNP microarray. *J Lipid Res* 2009 March 18.

(155) Lambert G, Amar MJ, Martin P, Fruchart-Najib J, Foger B, Shamburek RD, Brewer HB, Jr., Santamarina-Fojo S. Hepatic lipase deficiency decreases the selective uptake of HDL-cholesteryl esters in vivo. *J Lipid Res* 2000 May;41(5):667-72.

(156) Hegele RA, Vezina C, Moorjani S, Lupien PJ, Gagne C, Brun LD, Little JA, Connelly PW. A hepatic lipase gene mutation associated with heritable lipolytic deficiency. *J Clin Endocrinol Metab* 1991 March;72(3):730-2.

(157) Isaacs A, Sayed-Tabatabaei FA, Njajou OT, Witteman JC, van Duijn CM. The -514 C->T hepatic lipase promoter region polymorphism and plasma lipids: a meta-analysis. *J Clin Endocrinol Metab* 2004 August;89(8):3858-63.

(158) Jaye M, Lynch KJ, Krawiec J, Marchadier D, Maugeais C, Doan K, South V, Amin D, Perrone M, Rader DJ. A novel endothelial-derived lipase that modulates HDL metabolism. *Nat Genet* 1999 April;21(4):424-8.

(159) Mank-Seymour AR, Durham KL, Thompson JF, Seymour AB, Milos PM. Association between single-nucleotide polymorphisms in the endothelial lipase (LIPG) gene and high-density lipoprotein cholesterol levels. *Biochim Biophys Acta* 2004 February 27;1636(1):40-6.

(160) Yamakawa-Kobayashi K, Yanagi H, Endo K, Arinami T, Hamaguchi H. Relationship between serum HDL-C levels and common genetic variants of the endothelial lipase gene in Japanese school-aged children. *Hum Genet* 2003 September;113(4):311-5.

(161) Hutter CM, Austin MA, Farin FM, Viernes HM, Edwards KL, Leonetti DL, McNeely MJ, Fujimoto WY. Association of endothelial lipase gene (LIPG) haplotypes with high-density lipoprotein cholesterol subfractions and apolipoprotein AI plasma levels in Japanese Americans. *Atherosclerosis* 2006 March;185(1):78-86.

(162) Ma K, Cilingiroglu M, Otvos JD, Ballantyne CM, Marian AJ, Chan L. Endothelial lipase is a major genetic determinant for high-density lipoprotein concentration, structure, and metabolism. *Proc Natl Acad Sci U S A* 2003 March 4;100(5):2748-53.

(163) Graf GA, Li WP, Gerard RD, Gelissen I, White A, Cohen JC, Hobbs HH. Coexpression of ATP-binding cassette proteins ABCG5 and ABCG8 permits their transport to the apical surface. *J Clin Invest* 2002 September;110(5):659-69.

(164) Berge KE, Tian H, Graf GA, Yu L, Grishin NV, Schultz J, Kwiterovich P, Shan B, Barnes R, Hobbs HH. Accumulation of dietary cholesterol in sitosterolemia caused by mutations in adjacent ABC transporters. *Science* 2000 December 1;290(5497):1771-5.

(165) Nishida HI, Nishida T. Phospholipid transfer protein mediates transfer of not only phosphatidylcholine but also cholesterol from phosphatidylcholine-cholesterol vesicles to high density lipoproteins. *J Biol Chem* 1997 March 14;272(11):6959-64.

(166) Rye KA, Barter PJ. Changes in the size and density of human high-density lipoproteins promoted by a plasma-conversion factor. *Biochim Biophys Acta* 1986 February 28;875(3):429-38.

(167) van HR, van TA, van GT, Scheek L, Visser P, van der Kamp A, Grosveld F, de CR. Increased risk of atherosclerosis by elevated plasma levels of phospholipid transfer protein. *J Biol Chem* 2002 December 13;277(50):48938-43.

(168) Schlitt A, Bickel C, Thumma P, Blankenberg S, Rupprecht HJ, Meyer J, Jiang XC. High plasma phospholipid transfer protein levels as a risk factor for coronary artery disease. *Arterioscler Thromb Vasc Biol* 2003 October 1;23(10):1857-62.

(169) Spirin V, Schmidt S, Pertsemlidis A, Cooper RS, Cohen JC, Sunyaev SR. Common Single-Nucleotide Polymorphisms Act in Concert to Affect Plasma Levels of High-Density Lipoprotein Cholesterol. *Am J Hum Genet* 2007 October 19;81(6).

(170) Engler MB, Pullinger CR, Malloy MJ, Natanzon Y, Kulkarni MV, Song J, Eng C, Huuskonen J, Rivera C, Poon A, Bensley M, Sehnert A, Zellner C, Kane J, Aouizerat BE. Genetic variation in phospholipid transfer protein modulates lipoprotein profiles in hyperalphalipoproteinemia. *Metabolism* 2008 December;57(12):1719-24.

(171) Kathiresan S, Willer CJ, Peloso GM, Demissie S, Musunuru K, Schadt EE, Kaplan L, Bennett D, Li Y, Tanaka T, Voight BF, Bonnycastle LL, Jackson AU, Crawford G, Surti A, Guiducci C, Burtt NP, Parish S, Clarke R, Zelenika D, Kubalanza KA, Morken MA, Scott LJ, Stringham HM, Galan P, Swift AJ, Kuusisto J, Bergman RN, Sundvall J, Laakso M, Ferrucci L, Scheet P, Sanna S, Uda M, Yang Q, Lunetta KL, Dupuis J, de Bakker PI, O'Donnell CJ, Chambers JC, Kooner JS, Hercberg S, Meneton P, Lakatta EG, Scuteri A, Schlessinger D, Tuomilehto J, Collins FS, Groop L, Altshuler D, Collins R, Lathrop GM, Melander O, Salomaa V, Peltonen L, Orho-Melander M, Ordovas JM, Boehnke M, Abecasis GR, Mohlke KL, Cupples LA. Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat Genet* 2009 January;41(1):56-65.

(172) McGovern MM, Pohl-Worgall T, Deckelbaum RJ, Simpson W, Mendelson D, Desnick RJ, Schuchman EH, Wasserstein MP. Lipid

abnormalities in children with types A and B Niemann Pick disease. *J Pediatr* 2004 July;145(1):77-81.

(173) Lee CY, Krimbou L, Vincent J, Bernard C, Larramee P, Genest J, Jr., Marcil M. Compound heterozygosity at the sphingomyelin phosphodiesterase-1 (SMPD1) gene is associated with low HDL cholesterol. *Hum Genet* 2003 May;112(5-6):552-62.

(174) Lee CY, Lesimple A, Denis M, Vincent J, Larsen A, Mamer O, Krimbou L, Genest J, Marcil M. Increased sphingomyelin content impairs HDL biogenesis and maturation in human Niemann-Pick disease type B. *J Lipid Res* 2006 March;47(3):622-32.

(175) Pajukanta P, Allayee H, Krass KL, Kuraishy A, Soro A, Lilja HE, Mar R, Taskinen MR, Nuotio I, Laakso M, Rotter JI, de Bruin TW, Cantor RM, Lusis AJ, Peltonen L. Combined analysis of genome scans of dutch and finnish families reveals a susceptibility locus for high-density lipoprotein cholesterol on chromosome 16q. *Am J Hum Genet* 2003 April;72(4):903-17.

(176) Soro A, Pajukanta P, Lilja HE, Ylitalo K, Hiekkalinna T, Perola M, Cantor RM, Viikari JS, Taskinen MR, Peltonen L. Genome scans provide evidence for low-HDL-C loci on chromosomes 8q23, 16q24.1-24.2, and 20q13.11 in Finnish families. *Am J Hum Genet* 2002 May;70(5):1333-40.

(177) Aouizerat BE, Allayee H, Cantor RM, Dallinga-Thie GM, Lanning CD, de Bruin TW, Lusis AJ, Rotter JI. Linkage of a candidate gene locus to familial combined hyperlipidemia: lecithin:cholesterol acyltransferase on 16q. *Arterioscler Thromb Vasc Biol* 1999 November;19(11):2730-6.

(178) Ng MC, So WY, Lam VK, Cockram CS, Bell GI, Cox NJ, Chan JC. Genome-wide scan for metabolic syndrome and related quantitative

traits in Hong Kong Chinese and confirmation of a susceptibility locus on chromosome 1q21-q25. *Diabetes* 2004 October;53(10):2676-83.

(179) Bu X, Warden CH, Xia YR, De Meester C, Puppione DL, Teruya S, Lokensgard B, Daneshmand S, Brown J, Gray RJ, . Linkage analysis of the genetic determinants of high density lipoprotein concentrations and composition: evidence for involvement of the apolipoprotein A-II and cholesteryl ester transfer protein loci. *Hum Genet* 1994 June;93(6):639-48.

(180) Shearman AM, Ordovas JM, Cupples LA, Schaefer EJ, Harmon MD, Shao Y, Keen JD, DeStefano AL, Joost O, Wilson PW, Housman DE, Myers RH. Evidence for a gene influencing the TG/HDL-C ratio on chromosome 7q32.3-qter: a genome-wide scan in the Framingham study. *Hum Mol Genet* 2000 May 22;9(9):1315-20.

(181) Mahaney MC, Almasy L, Rainwater DL, VandeBerg JL, Cole SA, Hixson JE, Blangero J, MacCluer JW. A quantitative trait locus on chromosome 16q influences variation in plasma HDL-C levels in Mexican Americans. *Arterioscler Thromb Vasc Biol* 2003 February 1;23(2):339-45.

(182) Gagnon F, Jarvik GP, Motulsky AG, Deeb SS, Brunzell JD, Wijsman EM. Evidence of linkage of HDL level variation to APOC3 in two samples with different ascertainment. *Hum Genet* 2003 November;113(6):522-33.

(183) Yip AG, Ma Q, Wilcox M, Panhuysen CI, Farrell J, Farrer LA, Wyszynski DF. Search for genetic factors predisposing to atherogenic dyslipidemia. *BMC Genet* 2003;4 Suppl 1:S100.

(184) Altshuler D, Brooks LD, Chakravarti A, Collins FS, Daly MJ, Donnelly P. A haplotype map of the human genome. *Nature* 2005 October 27;437(7063):1299-320.

(185) Willer CJ, Sanna S, Jackson AU, Scuteri A, Bonnycastle LL, Clarke R, Heath SC, Timpson NJ, Najjar SS, Stringham HM, Strait J, Duren WL, Maschio A, Busonero F, Mulas A, Albai G, Swift AJ, Morken MA, Narisu N, Bennett D, Parish S, Shen H, Galan P, Meneton P, Hercberg S, Zelenika D, Chen WM, Li Y, Scott LJ, Scheet PA, Sundvall J, Watanabe RM, Nagaraja R, Ebrahim S, Lawlor DA, Ben-Shlomo Y, vey-Smith G, Shuldiner AR, Collins R, Bergman RN, Uda M, Tuomilehto J, Cao A, Collins FS, Lakatta E, Lathrop GM, Boehnke M, Schlessinger D, Mohlke KL, Abecasis GR. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet* 2008 February;40(2):161-9.

(186) Kathiresan S, Melander O, Guiducci C, Surti A, Burtt NP, Rieder MJ, Cooper GM, Roos C, Voight BF, Havulinna AS, Wahlstrand B, Hedner T, Corella D, Tai ES, Ordovas JM, Berglund G, Vartiainen E, Jousilahti P, Hedblad B, Taskinen MR, Newton-Cheh C, Salomaa V, Peltonen L, Groop L, Altshuler DM, Orho-Melander M. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat Genet* 2008 February;40(2):189-97.

(187) Aulchenko YS, Ripatti S, Lindqvist I, Boomsma D, Heid IM, Pramstaller PP, Penninx BW, Janssens AC, Wilson JF, Spector T, Martin NG, Pedersen NL, Kyvik KO, Kaprio J, Hofman A, Freimer NB, Jarvelin MR, Gyllensten U, Campbell H, Rudan I, Johansson A, Marroni F, Hayward C, Vitart V, Jonasson I, Pattaro C, Wright A, Hastie N, Pichler I, Hicks AA, Falchi M, Willemsen G, Hottenga JJ, de Geus EJ, Montgomery GW, Whitfield J, Magnusson P, Saharinen J, Perola M, Silander K, Isaacs A, Sijbrands EJ, Uitterlinden AG, Witteman JC, Oostra BA, Elliott P, Ruokonen A, Sabatti C, Gieger C, Meitinger T, Kronenberg F, Doring A, Wichmann HE, Smit JH, McCarthy MI, van Duijn CM, Peltonen L. Loci influencing lipid levels

and coronary heart disease risk in 16 European population cohorts. *Nat Genet* 2009 January;41(1):47-55.

(188) Genest JJ, Jr., Martin-Munley SS, McNamara JR, Ordovas JM, Jenner J, Myers RH, Silberman SR, Wilson PW, Salem DN, Schaefer EJ. Familial lipoprotein disorders in patients with premature coronary artery disease. *Circulation* 1992 June;85(6):2025-33.

(189) Lee CY, Krimbou L, Vincent J, Bernard C, Larramee P, Genest J, Jr., Marcil M. Compound heterozygosity at the sphingomyelin phosphodiesterase-1 (SMPD1) gene is associated with low HDL cholesterol. *Hum Genet* 2003 May;112(5-6):552-62.

(190) Genest JJ, Jr., Martin-Munley SS, McNamara JR, Ordovas JM, Jenner J, Myers RH, Silberman SR, Wilson PW, Salem DN, Schaefer EJ. Familial lipoprotein disorders in patients with premature coronary artery disease. *Circulation* 1992 June;85(6):2025-33.

(191) Yusuf S, Reddy S, Ounpuu S, Anand S. Global burden of cardiovascular diseases: Part II: variations in cardiovascular disease by specific ethnic groups and geographic regions and prevention strategies. *Circulation* 2001 December 4;104(23):2855-64.

(192) Inazu A, Jiang XC, Haraki T, Yagi K, Kamon N, Koizumi J, Mabuchi H, Takeda R, Takata K, Moriyama Y, . Genetic cholesteryl ester transfer protein deficiency caused by two prevalent mutations as a major determinant of increased levels of high density lipoprotein cholesterol. *J Clin Invest* 1994 November;94(5):1872-82.

(193) Fielding CJ, Fielding PE. Molecular physiology of reverse cholesterol transport. *J Lipid Res* 1995 February;36(2):211-28.

(194) Genest J, Pedersen TR. Prevention of cardiovascular ischemic events: high-risk and secondary prevention. *Circulation* 2003 April 22;107(15):2059-65.

(195) O'Connell BJ, Genest J, Jr. High-density lipoproteins and endothelial function. *Circulation* 2001 October 16;104(16):1978-83.

(196) Assmann G, Gotto AM, Jr. HDL cholesterol and protective factors in atherosclerosis. *Circulation* 2004 June 15;109(23 Suppl 1):III8-14.

(197) Wang N, Weng W, Breslow JL, Tall AR. Scavenger receptor BI (SR-BI) is up-regulated in adrenal gland in apolipoprotein A-I and hepatic lipase knock-out mice as a response to depletion of cholesterol stores. In vivo evidence that SR-BI is a functional high density lipoprotein receptor under feedback control. *J Biol Chem* 1996 August 30;271(35):21001-4.

(198) Ng DS, Leiter LA, Vezina C, Connelly PW, Hegele RA. Apolipoprotein A-I Q[-2]X causing isolated apolipoprotein A-I deficiency in a family with analphalipoproteinemia. *J Clin Invest* 1994 January;93(1):223-9.

(199) Takata K, Saku K, Ohta T, Takata M, Bai H, Jimi S, Liu R, Sato H, Kajiyama G, Arakawa K. A new case of apoA-I deficiency showing codon 8 nonsense mutation of the apoA-I gene without evidence of coronary heart disease. *Arterioscler Thromb Vasc Biol* 1995 November;15(11):1866-74.

(200) Romling R, von EA, Funke H, Motti C, Fragiacomo GC, Noseda G, Assmann G. A nonsense mutation in the apolipoprotein A-I gene is associated with high-density lipoprotein deficiency and periorbital xanthelasmas. *Arterioscler Thromb* 1994 December;14(12):1915-22.

(201) Matsunaga T, Hiasa Y, Yanagi H, Maeda T, Hattori N, Yamakawa K, Yamanouchi Y, Tanaka I, Obara T, Hamaguchi H. Apolipoprotein A-I

deficiency due to a codon 84 nonsense mutation of the apolipoprotein A-I gene. *Proc Natl Acad Sci U S A* 1991 April 1;88(7):2793-7.

(202) Weber M, McNicoll S, Marcil M, Connelly P, Lussier-Cacan S, Davignon J, Latour Y, Genest J, Jr. Metabolic factors clustering, lipoprotein cholesterol, apolipoprotein B, lipoprotein (a) and apolipoprotein E phenotypes in premature coronary artery disease in French Canadians. *Can J Cardiol* 1997 March;13(3):253-60.

(203) Marcil M, Bissonnette R, Vincent J, Krimbou L, Genest J. Cellular phospholipid and cholesterol efflux in high-density lipoprotein deficiency. *Circulation* 2003 March 18;107(10):1366-71.

(204) Krimbou L, Marcil M, Davignon J, Genest J, Jr. Interaction of lecithin:cholesterol acyltransferase (LCAT).alpha 2-macroglobulin complex with low density lipoprotein receptor-related protein (LRP). Evidence for an alpha 2-macroglobulin/LRP receptor-mediated system participating in LCAT clearance. *J Biol Chem* 2001 August 31;276(35):33241-8.

(205) Pisciotta L, Miccoli R, Cantafora A, Calabresi L, Tarugi P, Alessandrini P, Bittolo BG, Franceschini G, Cortese C, Calandra S, Bertolini S. Recurrent mutations of the apolipoprotein A-I gene in three kindreds with severe HDL deficiency. *Atherosclerosis* 2003 April;167(2):335-45.

(206) Grundy SM, Cleeman JI, Merz CN, Brewer HB, Jr., Clark LT, Hunninghake DB, Pasternak RC, Smith SC, Jr., Stone NJ. Implications of recent clinical trials for the National Cholesterol Education Program Adult Treatment Panel III guidelines. *Circulation* 2004 July 13;110(2):227-39.

(207) McPherson R, Frohlich J, Fodor G, Genest J, Canadian Cardiovascular Society. Canadian Cardiovascular Society position statement--recommendations for the diagnosis and treatment of dyslipidemia and

prevention of cardiovascular disease. *Can J Cardiol* 2006 September;22(11):913-27.

(208) Genest J, Jr., Bard JM, Fruchart JC, Ordovas JM, Schaefer EJ. Familial hypoalphalipoproteinemia in premature coronary artery disease. *Arterioscler Thromb* 1993 December;13(12):1728-37.

(209) Dastani Z, Engert JC, Genest J, Marcil M. Genetics of high-density lipoproteins. *Curr Opin Cardiol* 2006 July;21(4):329-35.

(210) McGovern MM, Pohl-Worgall T, Deckelbaum RJ, Simpson W, Mendelson D, Desnick RJ, Schuchman EH, Wasserstein MP. Lipid abnormalities in children with types A and B Niemann Pick disease. *J Pediatr* 2004 July;145(1):77-81.

(211) Lee CY, Lesimple A, Denis M, Vincent J, Larsen A, Mamer O, Krimbou L, Genest J, Marcil M. Increased sphingomyelin content impairs HDL biogenesis and maturation in human Niemann-Pick disease type B. *J Lipid Res* 2005 November 30.

(212) Karolchik D, Baertsch R, Diekhans M, Furey TS, Hinrichs A, Lu YT, Roskin KM, Schwartz M, Sugnet CW, Thomas DJ, Weber RJ, Haussler D, Kent WJ. The UCSC Genome Browser Database. *Nucleic Acids Res* 2003 January 1;31(1):51-4.

(213) *National Institutes of Health: Lipid Research Clinics Population Studies Databook*. 1 ed. 1980.

(214) Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin Chem* 1972 June;18(6):499-502.

(215) Alenezi MY, Marcil M, Blank D, Sherman M, Genest J, Jr. Is the decreased high-density lipoprotein cholesterol in the metabolic

syndrome due to cellular lipid efflux defect? *J Clin Endocrinol Metab* 2004 February;89(2):761-4.

(216) Schuchman EH, Levran O, Suchi M, Desnick RJ. An MspI polymorphism in the human acid sphingomyelinase gene (SMPD1). *Nucleic Acids Res* 1991 June 11;19(11):3160.

(217) Wan Q, Schuchman EH. A novel polymorphism in the human acid sphingomyelinase gene due to size variation of the signal peptide region. *Biochim Biophys Acta* 1995 April 24;1270(2-3):207-10.

(218) Sham PC, Curtis D. Monte Carlo tests for associations between disease and alleles at highly polymorphic loci. *Ann Hum Genet* 1995 January;59(Pt 1):97-105.

(219) Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 2001 April;68(4):978-89.

(220) Viana MB, Giugliani R, Leite VH, Barth ML, Lekhwani C, Slade CM, Fensom A. Very low levels of high density lipoprotein cholesterol in four sibs of a family with non-neuropathic Niemann-Pick disease and sea-blue histiocytosis. *J Med Genet* 1990 August;27(8):499-504.

(221) Subbaiah PV, Liu M. Role of sphingomyelin in the regulation of cholesterol esterification in the plasma lipoproteins. Inhibition of lecithin-cholesterol acyltransferase reaction. *J Biol Chem* 1993 September 25;268(27):20156-63.

(222) Rye KA, Hime NJ, Barter PJ. The influence of sphingomyelin on the structure and function of reconstituted high density lipoproteins. *J Biol Chem* 1996 February 23;271(8):4243-50.

(223) Bolin DJ, Jonas A. Sphingomyelin inhibits the lecithin-cholesterol acyltransferase reaction with reconstituted high density lipoproteins by

decreasing enzyme binding. *J Biol Chem* 1996 August 9;271(32):19152-8.

(224) Rye KA, Barter PJ. Formation and metabolism of prebeta-migrating, lipid-poor apolipoprotein A-I. *Arterioscler Thromb Vasc Biol* 2004 March;24(3):421-8.

(225) Simonaro CM, Desnick RJ, McGovern MM, Wasserstein MP, Schuchman EH. The demographics and distribution of type B Niemann-Pick disease: novel mutations lead to new genotype/phenotype correlations. *Am J Hum Genet* 2002 December;71(6):1413-9.

(226) Prenger VL, Beaty TH, Kwiterovich PO. Genetic determination of high-density lipoprotein-cholesterol and apolipoprotein A-1 plasma levels in a family study of cardiac catheterization patients. *Am J Hum Genet* 1992 November;51(5):1047-57.

(227) Lander ES. The new genomics: global views of biology. *Science* 1996 October 25;274(5287):536-9.

(228) Collins FS, Guyer MS, Charkravarti A. Variations on a theme: cataloging human DNA sequence variation. *Science* 1997 November 28;278(5343):1580-1.

(229) Wang N, Lan D, Chen W, Matsuura F, Tall AR. ATP-binding cassette transporters G1 and G4 mediate cellular cholesterol efflux to high-density lipoproteins. *Proc Natl Acad Sci U S A* 2004 June 29;101(26):9774-9.

(230) Assmann G, von EA, Funke H. High density lipoproteins, reverse transport of cholesterol, and coronary artery disease. Insights from mutations. *Circulation* 1993 April;87(4 Suppl):III28-III34.

(231) Cohen JC, Wang Z, Grundy SM, Stoesz MR, Guerra R. 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* 1994 December;94(6):2377-84.

(232) Davignon J, Genest J, Jr. Genetics of lipoprotein disorders. *Endocrinol Metab Clin North Am* 1998 September;27(3):521-50.

(233) Arya R, Duggirala R, Almasy L, Rainwater DL, Mahaney MC, Cole S, Dyer TD, Williams K, Leach RJ, Hixson JE, MacCluer JW, O'Connell P, Stern MP, Blangero J. Linkage of high-density lipoprotein-cholesterol concentrations to a locus on chromosome 9p in Mexican Americans. *Nat Genet* 2002 January;30(1):102-5.

(234) Kort EN, Ballinger DG, Ding W, Hunt SC, Bowen BR, Abkevich V, Bulka K, Campbell B, Capener C, Gutin A, Harshman K, McDermott M, Thorne T, Wang H, Wardell B, Wong J, Hopkins PN, Skolnick M, Samuels M. Evidence of linkage of familial hypoalphalipoproteinemia to a novel locus on chromosome 11q23. *Am J Hum Genet* 2000 June;66(6):1845-56.

(235) NIH Washington DC. Lipid Research Clinics population studies databook. *Department of health and human services, Public Health Services, Washington DC*.NIH publication 80-1527: 1980. p. 28-81.

(236) Marcil M, Yu L, Krimbou L, Boucher B, Oram JF, Cohn JS, Genest J, Jr. Cellular cholesterol transport and efflux in fibroblasts are abnormal in subjects with familial HDL deficiency. *Arterioscler Thromb Vasc Biol* 1999 January;19(1):159-69.

(237) Lange K, Weeks D, Boehnke M. Programs for Pedigree Analysis: MENDEL, FISHER, and dGENE. *Genet Epidemiol* 1988;5(6):471-2.

(238) Schaid DJ, McDonnell SK, Thibodeau SN. Regression models for linkage heterogeneity applied to familial prostate cancer. *Am J Hum Genet* 2001 May;68(5):1189-96.

(239) Sobel E, Lange K. Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. *Am J Hum Genet* 1996 June;58(6):1323-37.

(240) Myerowitz R, Hogikyan ND. Different mutations in Ashkenazi Jewish and non-Jewish French Canadians with Tay-Sachs disease. *Science* 1986 June 27;232(4758):1646-8.

(241) Elbein SC, Hasstedt SJ. Quantitative trait linkage analysis of lipid-related traits in familial type 2 diabetes: evidence for linkage of triglyceride levels to chromosome 19q. *Diabetes* 2002 February;51(2):528-35.

(242) Lilja HE, Soro A, Ylitalo K, Nuotio I, Viikari JS, Salomaa V, Vartiainen E, Taskinen MR, Peltonen L, Pajukanta P. A candidate gene study in low HDL-cholesterol families provides evidence for the involvement of the APOA2 gene and the APOA1C3A4 gene cluster. *Atherosclerosis* 2002 September;164(1):103-11.

(243) Murray CJ, Lopez AD. Alternative projections of mortality and disability by cause 1990-2020: Global Burden of Disease Study. *Lancet* 1997 May 24;349(9064):1498-504.

(244) Despres JP, Lemieux I, Dagenais GR, Cantin B, Lamarche B. HDL-cholesterol as a marker of coronary heart disease risk: the Quebec cardiovascular study. *Atherosclerosis* 2000 December;153(2):263-72.

(245) Assmann G, Schulte H, von Eckardstein A, Huang Y. High-density lipoprotein cholesterol as a predictor of coronary heart disease risk. The PROCAM experience and pathophysiological implications for reverse cholesterol transport. *Atherosclerosis* 1996 July;124 Suppl:S11-S20.

(246) Breslow JL. Familial disorders of high density lipoprotein metabolism. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. *The Molecular and Metabolic Basis of Inherited Diseases*.New York: McGraw-Hill; 2001.

(247) Dastani Z, Dangoisse C, Boucher B, Desbiens K, Krimbou L, Dufour R, Hegele RA, Pajukanta P, Engert JC, Genest J, Marcil M. A novel nonsense apolipoprotein A-I mutation (apoA-I(E136X)) causes low HDL cholesterol in French Canadians. *Atherosclerosis* 2006 March;185(1):127-36.

(248) Rioux JD, Stone VA, Daly MJ, Cargill M, Green T, Nguyen H, Nutman T, Zimmerman PA, Tucker MA, Hudson T, Goldstein AM, Lander E, Lin AY. Familial eosinophilia maps to the cytokine gene cluster on human chromosomal region 5q31-q33. *Am J Hum Genet* 1998 October;63(4):1086-94.

(249) Laitinen T, Daly MJ, Rioux JD, Kauppi P, Laprise C, Petays T, Green T, Cargill M, Haahtela T, Lander ES, Laitinen LA, Hudson TJ, Kere J. A susceptibility locus for asthma-related traits on chromosome 7 revealed by genome-wide scan in a founder population. *Nat Genet* 2001 May;28(1):87-91.

(250) Abecasis GR, Cherny SS, Cookson WO, Cardon LR. GRR: graphical representation of relationship errors. *Bioinformatics* 2001 August;17(8):742-3.

(251) Lee JC, Weissglas-Volkov D, Kyttala M, Dastani Z, Cantor RM, Sobel EM, Plaisier CL, Engert JC, van Greevenbroek MM, Kane JP, Malloy MJ, Pullinger CR, Huertas-Vazquez A, guilar-Salinas CA, Tusie-Luna T, de Bruin TW, Aouizerat BE, van der Kallen CC, Croce CM, Aqeilan RI, Marcil M, Viikari JS, Lehtimaki T, Raitakari OT, Kuusisto J, Laakso M, Taskinen MR, Genest J, Pajukanta P. WW-domain-

containing oxidoreductase is associated with low plasma HDL-C levels. *Am J Hum Genet* 2008 August;83(2):180-92.

(252) Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 2002 January;30(1):97-101.

(253) Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005 January 15;21(2):263-5.

(254) Heath SC. Markov chain Monte Carlo segregation and linkage analysis for oligogenic models. *Am J Hum Genet* 1997 September;61(3):748-60.

(255) Lange K, Cantor RM, Horvath S, Perola M, Sabatti C, Sinsheimer J, Sobel E. Mendel version 4.0: A complete package for the exact genetic analysis of discrete traits in pedigree and population data sets. *Amer J Hum Genetics* 2001;(69(Suplement)):504.

(256) Kong A, Gudbjartsson DF, Sainz J, Jonsdottir GM, Gudjonsson SA, Richardsson B, Sigurdardottir S, Barnard J, Hallbeck B, Masson G, Shlien A, Palsson ST, Frigge ML, Thorgeirsson TE, Gulcher JR, Stefansson K. A high-resolution recombination map of the human genome. *Nat Genet* 2002 July;31(3):241-7.

(257) Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 1995 November;11(3):241-7.

(258) Shearman AM, Ordovas JM, Cupples LA, Schaefer EJ, Harmon MD, Shao YJ, Keen JD, DeStefano AL, Wilson PWF, Myers RH. Evidence for a gene influencing the TG/HDL-C ratio on chromosome 7q32.3-qter: a genome-wide scan in the Framingham Study. *Human Molecular Genetics* 2000 May 22;9(9):1315-20.

(259) Gagnon F, Jarvik GP, Motulsky AG, Deeb SS, Brunzell JD, Wijsman EM. Evidence of linkage of HDL level variation to APOC3 in two samples with different ascertainment. *Human Genetics* 2003 November;113(6):522-33.

(260) Mehrabian M, Castellani LW, Wen PZ, Wong J, Rithaporn T, Hama SY, Hough GP, Johnson D, Albers JJ, Mottino GA, Frank JS, Navab M, Fogelman AM, Lusis AJ. Genetic control of HDL levels and composition in an interspecific mouse cross (CAST/Ei x C57BL/6J). *J Lipid Res* 2000 December;41(12):1936-46.

(261) Ishimori N, Li R, Kelmenson PM, Korstanje R, Walsh KA, Churchill GA, Forsman-Semb K, Paigen B. Quantitative trait loci analysis for plasma HDL-cholesterol concentrations and atherosclerosis susceptibility between inbred mouse strains C57BL/6J and 129S1/SvImJ. *Arterioscler Thromb Vasc Biol* 2004 January;24(1):161-6.

(262) Thompson JF, Durham LK, Lira ME, Shear C, Milos PM. CETP polymorphisms associated with HDL cholesterol may differ from those associated with cardiovascular disease. *Atherosclerosis* 2005 July;181(1):45-53.

(263) Kondo I, Berg K, Drayna D, Lawn R. Dna Polymorphism at the Locus for Human Cholesteryl Ester Transfer Protein (Cetp) Is Associated with High-Density Lipoprotein Cholesterol and Apolipoprotein Levels. *Clinical Genetics* 1989 January;35(1):49-56.

(264) de Grooth GJ, Klerkx AH, Stroes ES, Stalenhoef AF, Kastelein JJ, Kuivenhoven JA. A review of CETP and its relation to atherosclerosis. *J Lipid Res* 2004 November;45(11):1967-74.

(265) Thompson JF, Lira ME, Durham LK, Clark RW, Bamberger MJ, Milos PM. Polymorphisms in the CETP gene and association with CETP mass and HDL levels. *Atherosclerosis* 2003 April;167(2):195-204.

(266) Hasegawa N, Torii T, Kato T, Miyajima H, Furuhata A, Nakayasu K, Kanai A, Habuchi O. Decreased GlcNAc 6-O-sulfotransferase activity in the cornea with macular corneal dystrophy. *Invest Ophthalmol Vis Sci* 2000 November;41(12):3670-7.

(267) Rezaee F, Casetta B, Levels JH, Speijer D, Meijers JC. Proteomic analysis of high-density lipoprotein. *Proteomics* 2006 January;6(2):721-30.

(268) Kathiresan S, Melander O, Guiducci C, Surti A, Burtt NP, Rieder MJ, Cooper GM, Roos C, Voight BF, Havulinna AS, Wahlstrand B, Hedner T, Corella D, Tai ES, Ordovas JM, Berglund G, Vartiainen E, Jousilahti P, Hedblad B, Taskinen MR, Newton-Cheh C, Salomaa V, Peltonen L, Groop L, Altshuler DM, Orho-Melander M. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nature Genetics* 2008 February;40(2):189-97.

(269) Yusuf S, Hawken S, Ounpuu S, Dans T, Avezum A, Lanas F, McQueen M, Budaj A, Pais P, Varigos J, Lisheng L. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet* 2004 September 11;364(9438):937-52.

(270) Assmann G, Nofer JR. Atheroprotective effects of high-density lipoproteins. *Annu Rev Med* 2003;54:321-41.

(271) Clee SM, Kastelein JJ, van DM, Marcil M, Roomp K, Zwarts KY, Collins JA, Roelants R, Tamasawa N, Stulc T, Suda T, Ceska R, Boucher B, Rondeau C, DeSouich C, Brooks-Wilson A, Molhuizen HO,

Frohlich J, Genest J, Jr., Hayden MR. Age and residual cholesterol efflux affect HDL cholesterol levels and coronary artery disease in ABCA1 heterozygotes. *J Clin Invest* 2000 November;106(10):1263-70.

(272) Goldstein JL, Brown MS. Molecular medicine. The cholesterol quartet. *Science* 2001 May 18;292(5520):1310-2.

(273) Reich DE, Lander ES. On the allelic spectrum of human disease. *Trends Genet* 2001 September;17(9):502-10.

(274) Priore OC, Pisciotta L, Li VG, Sambataro MP, Cantafora A, Bellocchio A, Catapano A, Tarugi P, Bertolini S, Calandra S. Inherited apolipoprotein A-V deficiency in severe hypertriglyceridemia. *Arterioscler Thromb Vasc Biol* 2005 February;25(2):411-7.

(275) Rader DJ, Cohen J, Hobbs HH. Monogenic hypercholesterolemia: new insights in pathogenesis and treatment. *J Clin Invest* 2003 June;111(12):1795-803.

(276) Brown ML, Inazu A, Hesler CB, Agellon LB, Mann C, Whitlock ME, Marcel YL, Milne RW, Koizumi J, Mabuchi H, . Molecular basis of lipid transfer protein deficiency in a family with increased high-density lipoproteins. *Nature* 1989 November 23;342(6248):448-51.

(277) Hegele RA, Little JA, Vezina C, Maguire GF, Tu L, Wolever TS, Jenkins DJ, Connelly PW. Hepatic lipase deficiency. Clinical, biochemical, and molecular genetic characteristics. *Arterioscler Thromb* 1993 May;13(5):720-8.

(278) Santamarina-Fojo S, Brewer HB, Jr. The familial hyperchylomicronemia syndrome. New insights into underlying genetic defects. *JAMA* 1991 February 20;265(7):904-8.

(279) Manolio TA. Cohort studies and the genetics of complex disease. *Nat Genet* 2009 January;41(1):5-6.

(280) Wright AF, Carothers AD, Pirastu M. Population choice in mapping genes for complex diseases. *Nat Genet* 1999 December;23(4):397-404.

(281) Varilo T, Peltonen L. Isolates and their potential use in complex gene mapping efforts. *Curr Opin Genet Dev* 2004 June;14(3):316-23.

(282) Harrap SB, Wong ZY, Scurrah KJ, Lamantia A. Genome-wide linkage analysis of population variation in high-density lipoprotein cholesterol. *Hum Genet* 2006 June;119(5):541-6.

APPENDIX

Copyright waivers, co-author permissions and research compliance certificates have been attached to the library waivers.