Atheroprotective targets associated with high-density lipoprotein: phospholipid transfer protein, paraoxonase-1, and their effects on cerebrovascular disease

Daniel Seung Kim

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2015

Reading Committee:
Gail P. Jarvik, Chair
Brian L. Browning
Elizabeth M. Blue

Program Authorized to Offer Degree:
Genome Sciences

University of Washington

ABSTRACT

Atheroprotective targets associated with high-density lipoprotein: phospholipid transfer protein, paraoxonase-1, and their effects on cerebrovascular disease

Daniel Seung Kim

Chair of the Supervisory Committee:
Professor and Head, Gail P. Jarvik
Division of Medical Genetics, Department of Medicine
Department of Genome Sciences

The recent high-profile failures of clinical trials and Mendelian randomization studies to demonstrate a causal role of high-density lipoprotein cholesterol (HDL-C) in cardiovascular disease have shifted focus to other measures of HDL that may better reflect its cardioprotective properties. Here, I describe a series of work that first identified and validated that the smaller and denser HDL particles were more likely the source of HDL's cardioprotective properties. I then explore specific proteins that co-localize with these small, dense HDL particles: phospholipid transfer protein (PLTP) and paraoxonase-1 (PON1). First, I demonstrate that commercially available assays of PLTP activity are only weakly correlated with the low-throughput, but gold-standard radiometric PLTP assay. Using a large collection of participants with radiometrically-measured PLTP activity, I then show that PLTP activity measured by this method is protective against carotid artery disease. Finally, I use data from the NHLBI Exome Sequencing Project (ESP) to demonstrate Mendelian randomization for the role of *PON1* and cerebrovascular disease by identifying a rare variant that is associated with decreased PON1 activity and increased risk of ischemic stroke. These findings aid researchers and clinicians to narrow their focus onto specific functions and measures of HDL that may potentially be used as therapeutics.

ACKNOWLEDGEMENTS

Looking back on my education at the University of Washington, I can say with confidence that I have stood on the shoulders of giants. Without these people, named herein, I know that I would not have been able to do the work that is presented.

First and foremost, I am extremely grateful to my mentor, Gail P. Jarvik. Gail took me in on extremely short notice when I moved across the country due to family health concerns. Gail has always given me her precious time and filled it with sage advice, constructive criticism, and also enthusiasm for me pursuing my own passions. Gail has patiently taught me the scientific method over the past four years: coding, experimentation, interpretation, writing manuscripts, editing manuscripts, and writing grants – in short, how to be a scientist. All the while, Gail has also been a role model of the type of physician-scientist I hope to one day be a pale shadow of: calm and giving, whip-smart, and one who manages to maintain work-life balance. In short, my Ph.D. studies may have been my only time I could choose my own boss – but in Gail, I am certain that I made the best possible choice.

I am also appreciative of the faculty members on my dissertation committee: David A. Dichek, Debbie A. Nickerson, Brian L. Browning, and Elizabeth M. Blue. All have given selflessly of their time and energy to provide me with feedback and mold me into the fledgling scientist I am today.

More broadly, I have been fortunate to work with world-renowned leaders in the topics emphasized in this dissertation. For paraoxonase-1, I have worked closely with Clement E.

Furlong, from whom I gained a portion of his broad and thorough knowledge base on the paraoxonases and their importance in human health. For studies of high-density lipoprotein (HDL) and phosphopholipid transfer protein (PLTP), I have gained invaluable experience by working closely with John A. Albers, John D. Brunzell, Jay W. Heinecke, and Tomas Vaisar.

Within the collaborative Jarvik lab, I am indebted to the advice and help of Amber A. Burt, David R. Crosslin, Elisabeth A. Rosenthal, Adam S. Gordon, and Melody Rynerson Palmer. In addition, Jane E. Ranchalis, Martha J. Horike-Pyne, and Julieann K. Marshall both shepherded me through the Carotid Lesion Epidemiology And Risk (CLEAR) study data.

For funding, I am grateful for 2 years of support from the National Institutes of Mental Health for my F31. I am similarly grateful for training grant support from the Genome Training Grant and Cardiovascular Pathology Training Grant.

On the administrative side, I am grateful to the work of Sara Carlson, Brian Giebel, and Colleen Davis, without whom I would have struggled to accomplish simple tasks.

From my personal life, I am thankful for Gongzhu Yatong F. Li and Molly Kim-Li for their unconditional support throughout my education. I am also grateful for the unsolicited (but loving) advice from my father, Yongmin Kim, and for my mother, Eunai Kim, for being there to temper his critiques. Finally, I am grateful for my friends made during my education: Aaron H. McKenna, Alex K. Hu, Vanessa E. Gray, and Katherine A. Sitko – without your support through the ups and downs of graduate school, I am certain I would not have made it to the finish line.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1 Cardiovascular disease, stroke, and carotid artery disease

Cardiovascular disease (CVD) is one of the leading causes of death in the developed world, accounting for 1 out of every 3 deaths in the United States in 2009[1]. Approximately 1/2 of men and 1/3 of women will develop CVD in their lifetimes, with both the incidence and prevalence of CVD increasing with age[1].

The majority of CVD occurs from atherosclerosis in the major arteries that supply oxygen and nutrients to organs such as the heart and the brain. Atherosclerotic plaque growth can cause local tissue ischemia (e.g., transient ischemic attacks from atherosclerosis in the carotid arteries supplying the brain). As well, atherosclerotic plaques can rupture and cause a severe clotting response that blocks blood flow, causing a complete blockage of oxygen and nutrient delivery to tissue, which results in gross tissue necrosis (e.g., stroke). Stroke, considered independent of other CVD, is the third leading cause of death and leading cause of disability in the United States[1].

Atherosclerosis in the carotid arteries can be assessed non-invasively and quantitatively through use of ultrasound[2]. Significant atherosclerosis in the carotid arteries is diagnosed as carotid artery disease (CAAD)[2-4]. CAAD severity is directly related to risk of stroke[2-4]. Notably, CAAD of 80% stenosis (severe) is found in 4.2% of men and 2% of women in the 60-69 year age group[5]. Additionally, CAAD is predictive of death from CVD, and patients with symptomatic CAAD have a 1-year event rate of 14.5% for cardiovascular death, myocardial infarction, or stroke[6].

1.2 High-density lipoprotein: overview of functions

High-density lipoprotein (HDL) has been consistently associated with decreased cardiovascular events[7]. One of the previously hypothesized mechanisms through which HDL exerts its cardioprotective effects is through reverse cholesterol transport, through which cholesterol is ultimately excreted through the liver[8]. Recent longitudinal studies have demonstrated that HDL-mediated cholesterol efflux is significantly associated with decreased incidence of CVD events (defined as first myocardial infarction (MI), stroke, coronary revascularization, or CVD-related death). In this study, the highest quartile of cholesterol efflux was associated with a 67% reduction in hazard of CVD event as compared to the lowest quartile (Hazard Ratio (HR): 0.33)[9].

In addition to its cholesterol efflux functions, HDL and its associated proteins are antioxidant, with *in vitro* investigation demonstrating an HDL-mediated inhibition of LDL oxidation[10], one of the first steps in atherosclerotic plaque development. HDL also has been reported to be anti-inflammatory through its inhibition of monocyte adhesion to the vessel wall via downregulation of adhesion molecule expression[11,12] and through HDL-mediated inhibition of monocyte migration to local sites of vessel injury[13]. In addition to these anti-inflammatory effects, HDL is a part of the innate immunity, with the ability to carry complement proteins[14], bind to lipopolysaccharides[15], and regulate the response to viral infections[16]. Finally, HDL is endothelial protective and antithrombotic through its upregulation of nitric oxide production, which prevents atherogenesis [17].

1.3 HDL2 and HDL3

The HDL particle possesses a complex proteome, with a recent estimates of 67 associated proteins[14]. Broadly, HDL is composed of two sub-species, HDL2 and HDL3 that can be separated by ultracentrifugation and electrophoresis[18]. HDL2 and HDL3 each have different biochemical, physiologic, and metabolic functions[19]. The HDL subfraction that contributes to HDL's anti-inflammatory and endothelial protective functions is still in debate. HDL2 is considered to be closely associated with reverse cholesterol transport, and HDL3 is largely considered to be antioxidant[19].

HDL2 is the larger of the two HDL subfractions due to its higher concentration of cholesterol esters, and is closely associated with apolipoprotein AI (apoA1)[19]. ApoA1, in addition to cholesterol ester transport protein (CETP), is involved in macrophage reverse cholesterol transport, the process that is traditionally attributed the majority of HDL's cardioprotective effects. In the presence of triglyceride-rich lipoproteins, CETP transfers cholesterol from HDL to low-density lipoproteins (LDL)[20]. However, if triglyceride-rich lipoprotein concentrations are low, cholesterol is sent from HDL to the bile for excretion[20]. As the cholesterol-rich subfraction of HDL associated with apoA1 and reverse cholesterol transport, HDL2 has been hypothesized to be the subfraction most closely related with cardioprotection[21,22].

HDL3 is smaller and denser than HDL2, and possesses a different proteome. HDL3 particles are considered to be more antioxidant, likely due at least in part to a high association with paraoxonase 1 (PON1), an antioxidant enzyme[23]. As a result, HDL3 is less susceptible to oxidation than HDL2[24].

1.4 Paraoxonase-1 and Phospholipid Transfer Protein Activities on HDL3

PON1 enzyme activity is most strongly physically associated with HDL3[23]. PON1 activity has been shown to inhibit LDL cholesterol (LDL-C) oxidiation *in vitro*[10,25-27], one of the first steps in atherosclerosis. In addition, PON1 enzyme activity is considered cardioprotective. In numerous cross-sectional and case-control studies, we have reported a strong inverse relationship between PON1 activity and CAAD[28-30]. Moreover, longitudinal evidence has been reported from a prospective cohort of 1399 high-risk patients, which was followed for 3 years for cardiovascular events. At the time of analysis, it was found that lower PON1 activity was associated with an increased risk for cardiovascular events[31]. Moreover, when separating the 1399 patients into quartiles based on PON1 activity, there was a dose dependent decrease in risk for every increasing quartile of PON1 enzyme activity[31].

Similar to PON1, phospholipid transfer protein (PLTP) is also found more frequently on small, dense HDL particles. PLTP physically associates with apolipoprotein A1 (apoA1), and is involved in reverse cholesterol transport (see above):[32] PLTP is expressed in macrophages, is up-regulated in cholesterol-laden macrophage foam cells, and is present in atherosclerotic plaque[33,34]. Through this expression in atherosclerotic sites and cells, PLTP promotes the binding of high-density lipoprotein (HDL) to cholesterol-laden macrophages and fibroblasts and assists in HDL remodeling[35]. Only 10-50% of total plasma PLTP is "active", and can transfer triglyceride-rich phospholipids from very low-density lipoproteins (VLDLs) and chylomicrons to HDL particles[36-38]. Knockdown of this functional PLTP activity (PLTPa) has recently been

shown to decrease apoA1- and HDL3-mediated cholesterol efflux by 67% and 30%, respectively[39].

1.5 HDL cholesterol is not causally related to cardioprotective pathways

The amount of cholesterol carried by HDL (HDL-C) are an easily measured proxy of HDL content, and has been consistently inversely related in observational studies, with coronary heart disease (CHD) risk increasing by 25% for each 5 mg/dl decrease in HDL-C levels below the sex-specific median[7]. Additionally, the protective effects of HDL-C appear to be independent of LDL cholesterol levels (LDL-C)[40].

Despite the strong observational evidence in support of HDL-C being inversely associated with atherosclerosis, HDL-C has recently failed in multiple attempts to establish causality. In the first of these studies, HDL-C was investigated in 167 patients with CHD already taking a statin, with carotid intima media thickening as the endpoint (the ARBITER-2 trial)[41]. Those in the ARBITER-2 treatment group were given niacin daily in addition to their statin, and showed significant increases in HDL-C (21%) at 1-year follow-up. However, analyses of carotid intima media thickening showed no statistical difference between the treatment and placebo group for the ARBITER-2 study. In the second trial, HDL-C was studied in the setting of a randomized clinical trial of 3414 patients with established CHD, with the primary endpoint of any adverse cardiovascular event (the AIM-HIGH trial)[42]. In the AIM-HIGH randomized clinical trial, all patients received simvastatin at doses necessary to keep LDL-C < 100 mg/dl. Half of the patients were randomized to niacin at clinically effective doses to raise their levels of HDL-C, while the other half of patients were randomized to extremely low doses of niacin (placebo). At the two-

year follow-up, patients randomized to niacin and simvastatin had significantly higher levels of HDL-C (+7 mg/dl on average); however, the rate of adverse cardiovascular events was not significantly different than the group randomized to simvastatin and placebo in the AIM-HIGH study (p=0.79). Separately, a Mendelian randomization study of 20,913 cases of myocardial infarction (MI) and 95,407 controls studied the effects of a single nucleotide polymorphism (SNP) in the endothelial lipase gene (*LIPG*$_{Asn396Ser}$) known to increase HDL-C[43]. As expected, carriers of the *LIPG*$_{Asn396Ser}$ SNP had significantly higher levels of HDL-C, which would be expected from observational studies to decrease risk of MI by 13%. Yet, the *LIPG*$_{Asn396Ser}$ SNP was not associated with risk of MI (p=0.85).

1.6 CETP inhibitors in randomized clinical trials

Several other trials have focused on raising HDL-C through the specific HDL-associated process of reverse cholesterol efflux, through which HDL transports cholesterol to the liver where it is excreted in the bile. CETP is an HDL-associated enzyme that exchanges cholesterol esters on HDL2 for triglycerides stored on LDL particles, with the resulting decrease in HDL-C and increase in LDL-C. The hypothesis under investigation is that inhibition of CETP will prevent HDL from exchanging cholesterol with LDL (and other apolipoprotein B associated lipoproteins), thereby increasing HDL-C, and, likely, cholesterol efflux through the liver[8]. Despite the biological strength of this hypothesis, torcetrapib, a promising CETP inhibitor, recently failed to prevent adverse cardiovascular events in a randomized, double-blind clinical trial of 15,067 patients with baseline CHD (the ILLUMINATE study)[8]. After 1 year of follow-up in ILLUMINATE, torcetrapib increased HDL-C by 72.1% and decreased LDL-C by 24.9%. Despite the improvement in lipid profile, patients randomized to torcetrapib had a significantly

higher rate of adverse cardiovascular events, and a significantly higher rate of all-cause mortality. The likely culprit identified by ILLUMINATE investigators was the mean increase of 5.4 mmHg in systolic blood pressure in patients randomized to torcetrapib, which was a known side effect of the CETP inhibitor. Thus, despite the failure of this study, additional randomized clinical trials have continued to recruit and follow patients.

Three randomized clinical trials followed the ILLUMINATE study for the newer CETP inhibitors, dalcetrapib, anacetrapib, and evacetrapib. Dalcebtrapib showed no blood pressure side effects, and was tested in 15,871 patients who had all recently had an acute coronary syndrome (the Dal-Outcomes study). Despite a significant increase in HDL-C with dalcetrapib treatment, there was no change in rates of composite cardiovascular endpoint, and Dal-Outcomes trial was ended prematurely[44]. Similarly, a separate randomized clinical trial evaluated the effects of evacetrapib in approximately 400 patients with abnormal lipid profiles and found no blood pressure side effects and an improvement in HDL-C and LDL-C profiles[42]. However, Eli Lilly (the developer of evacetrapib) has recently discontinued development of evacetrapib after interim analyses of their phase 3 randomized clinical trial (ACCELERATE) found no evidence of effect after approximately 3 years of follow-up (no formal publication of these results is yet available). A final randomized clinical trial to evaluate the possible effects of CETP inhibition and CVD uses anacetrapib, a powerful CETP inhibitor that increases HDL-C while lowering LDL-C. In the DEFINE study, 1623 patients with CHD and taking a statin were randomized to anacetrapib or placebo. DEFINE investigators noted no blood pressure side effects, and a significant increase in HDL-C and a significant decrease in LDL-C in the anacetrapib and statin group compared to patients on placebo and statin therapy[45].

Given the failures of the ILLUMINATE, DEFINE, and ACCELERATE trials in conjunction with the findings from HDL-C Mendelian randomization and the AIM-HIGH trial, the hypothesis that HDL is cardioprotective has lost much of its prior support and enthusiasm from the medical community[46]. Together, the failures of randomized clinical trials investigating CETP inhibition in conjunction with the null results from HDL-C Mendelian randomization, ARBITER-2 and AIM-HIGH, have raised new doubt on the cardioprotective nature of HDL[46].

1.7 HDL particle concentration as a better predictor of cardioprotection

In contrast to the aforementioned clinical trials to raise HDL-C, more recent evidence from the Multi-Ethnic Study of Atherosclerosis (MESA) study has revealed a possible explanation for the negative results of HDL-C and cardiovascular events[47]. In this work, the MESA investigators followed a prospective cohort of 5,598 subjects without baseline CHD and not taking lipid-lowering medications for an average of 6 years and then evaluated the association of HDL-C and the total concentration of HDL particles in the plasma (HDL-P, measured by nuclear magnetic resonance spectroscopy) with incident CHD (n=227 events). As expected, HDL-C and HDL-P were highly correlated (Spearman's correlation coefficient of 0.69, $p<0.001$). However, in multivariate regression models, once HDL-P was in the model with LDL-C and other confounders (age, sex, ethnicity, hypertension, and smoking), HDL-C was no longer a significant predictor of CVD. This finding indicated that although HDL-C captured a large portion of HDL-P variation (as indicated by the high correlation between the two), HDL-C measurements alone did not reflect the individual elements of HDL that were primarily responsible for cardioprotection.

1.8 Research aims of this dissertation

In summary, recent work has demonstrated that measuring HDL-C alone likely does not reflect the individual cardioprotective elements of the complex HDL molecule and its vast number of associated proteins[47]. There is likely no one molecule or protein solely responsible for the cardioprotective effects of HDL: it is more likely that this effect is due to the contributions of numerous proteins, including apoA1, PON1, PLTP, and possibly through CETP inhibition.

As HDL-C measurements solely measure the cholesterol content of HDL and do not reflect its complex proteome, I hypothesized that one of the specific sub-fractions of HDL that better correlate with these specific HDL functions, HDL2 and HDL3, are superior predictors of CAAD. Investigation into this hypothesis comprises the first two research chapters of this dissertation, wherein I first use data from the Carotid Lesion Epidemiology And Risk (CLEAR) study to show that the smaller, denser HDL3 is a stronger predictor of CAAD[48]. In the second research chapter, I follow-up on these results in the MESA study and show that small, dense HDL-P concentration is robustly associated with carotid intima media thickness. In addition, I use data from the CLEAR study to demonstrate that possibly cardioprotective proteins PON1 and PLTP activities are more strongly associated with this smaller and denser HDL-P (in review).

I then delve into specific functions and enzyme activities associated with the small, dense HDL-P: specifically, PLTP activity and PON1 activity. In my third research chapter, I demonstrate that fluorescently measured PLTP activity is poorly correlated with gold-standard, radiometric PLTP activity. I then use data from the CLEAR study on 1,115 participants with gold-standard PLTP

activity measures to demonstrate that PLTP is protective against CAAD – a finding that is in conflict with numerous studies of fluorescently measured PLTP that have found it to be a risk factor of CVD[49].

In my final research chapter, I perform a Mendelian randomization study of rare variation in *PON1* from the National Heart, Blood, and Lung Institute Exome Sequencing Project (NHLBI ESP). In this work, I show that rare variants, and specifically the $PON1V_{109I}$ missense, are associated with decreased PON1 activity, and moreover with differential ischemic stroke risk[50].

Through these various projects focusing on HDL, I have narrowed research focus to the smaller, denser HDL particle, and have further elucidated PLTP and PON1 activities as important for cerebrovascular disease risk. These findings add to a body of knowledge that will aid future investigators to improve patient risk stratification and overall outcomes for prevalent and deadly cerebrovascular disease.

CHAPTER 2: HDL3 IS THE SUPERIOR PREDICTOR OF CAROTID ARTERY DISEASE

This chapter is based on the following peer-reviewed publication[48]:

Kim DS, Burt AA, Rosenthal EA, Ranchalis JE, Eintracht JF, Hatsukami TS, Furlong CE, Marcovina S, Albers JJ, Jarvik GP. HDL3 is a superior predictor of carotid artery disease in a case-control cohort of 1725 participants. *J Am Heart Assoc.* 2014 Jun 25;3(3):e000902. doi: 10.1161/JAHA.114.000902. PubMed PMID: 24965026.

2.1 SUMMARY

Purpose: Recent data suggests that high-density lipoprotein cholesterol levels (HDL-C) are likely not in the causative pathway of atheroprotection, shifting focus from HDL-C to its sub-fractions and associated proteins. This study's goal was to determine which HDL phenotype was the better predictor of carotid artery disease (CAAD).

Methods and Results: HDL2 and HDL3 were measured in 1,725 participants of European ancestry in a prevalent case-control cohort study of CAAD. Stratified analyses were conducted for males (n=1201) and females (n=524). Stepwise linear regression was used to determine whether HDL-C, HDL2, HDL3, or apolipoprotein A1 (apoA1) was the best predictor of CAAD, while adjusting for the confounders of censored age, diabetes, and current smoking status. In both males and females, HDL3 was negatively associated with CAAD (p=0.0011 and p=0.033 for males and females, respectively); once HDL3 was included in the model, no other HDL phenotype was significantly associated with CAAD. Addition of paraoxonase 1 (PON1) activity to the aforementioned regression model showed a significant and independent (of HDL3) association with CAAD in males (p=0.001), but not in the smaller female subgroup.

Conclusions: This study is the first to contrast the associations of HDL2 and HDL3 with CAAD. We found that HDL3 levels were more predictive of CAAD status than HDL2, HDL-C, or apoA1. Additionally, for males PON1 activity improved the overall model prediction for CAAD independently and additively with HDL3 levels. Further investigation into the molecular mechanisms through which HDL3 is associated with protection from CAAD is warranted.

2.2 BACKGROUND

Cardiovascular disease (CVD) is one of the leading causes of death in the developed world[51]. Stroke, considered independently of other CVD, is the fourth leading cause of death and the leading cause of disability in the United States[51], and the second leading cause of death globally[52]. Atherosclerosis in the carotid arteries can be assessed non-invasively and quantitatively through use of ultrasound[2]. Significant atherosclerosis in the carotid arteries is diagnosed as carotid artery disease (CAAD)[2-4]. CAAD severity is directly related to risk of stroke[2-4].

Despite strong observational evidence in support of high-density lipoprotein cholesterol levels (HDL-C) being cardioprotective[7], HDL-C has recently failed in numerous attempts to establish causality. Recent evidence from the Multi-Ethnic Study of Atherosclerosis (MESA) study has revealed a possible explanation for the negative results of HDL-C and cardiovascular events[47]. In this study, Mackey *et al.* evaluated the association of HDL-C and HDL particle concentration measured by nuclear magnetic resonance spectroscopy (HDL-P) with incident CHD (n=227 events). As expected, HDL-C and HDL-P were highly correlated (Spearman's correlation coefficient of 0.69, $p<0.001$). However, in multivariate regression models, HDL-P was the superior predictor of incident CHD in comparison to HDL-C. This finding indicated that although HDL-C captured a large portion of HDL-P variation, HDL-C measurements alone did not reflect the individual elements of HDL captured by HDL-P that were primarily responsible for cardioprotection.

Aspects of HDL that are not captured by HDL-C include its immensely complex proteome, with recent estimates of 64 associated proteins[14]. Broadly, HDL is composed of two sub-species, HDL2 and HDL3, which can be separated by ultracentrifugation and electrophoresis[19]. Both HDL2 and HDL3 have distinct biochemical, physiologic, and metabolic functions[19]. Investigations contrasting the cardioprotection of HDL2 or HDL3 have yielded conflicting results, with aspects of both sub-species being associated with decreased risk of CVD[19].

Given the recent evidence suggesting that HDL-C level does not have a direct causal role in cardioprotection, we hypothesized that specific aspects of HDL not completely correlated with HDL-C level were superior predictors of CAAD, as defined by greater than 50% stenosis in either carotid artery. Specifically, we sought to determine whether the level of one of the two subspecies of HDL (HDL2 versus HDL3) was a superior predictor of CAAD status in a prevalent case-control cohort from the Carotid Lesion Epidemiology and Risk (CLEAR) study, a Seattle-based repository comprised primarily of veterans, which was collected to identify risk factors for CAAD, CAAD progression, and other atherosclerotic disease end-points. Additionally, we attempted to elucidate whether functional aspects of HDL, such as its associated proteins apolipoprotein A1 (apoA1) and paraoxonase 1 (PON1), independently predicted further CAAD status.

2.3 METHODS

Ethics Statement

Institutional review boards at the University of Washington, Virginia Mason Medical Center, and Veterans Affairs Puget Sound approved the CLEAR study. Written, informed consent was obtained from each participant of the study.

Sample

The CLEAR study is a Seattle-based prevalent CAAD case-control study, composed primarily of veterans, with controls matched by age at diagnosis (for CAAD cases) and current unaffected age (for controls). Exclusion criteria included familial hypercholesterolemia, total fasting cholesterol greater than 400mg/dl, hypocoagulable state and/or the use of anticoagulant medication, post-organ transplant, or the inability to consent. All subjects underwent ultrasound assessment of their carotid arteries for the presence or absence of atherosclerotic plaque, except a small number of CAAD cases that had a prior carotid endarterectomy for symptomatic obstruction. CAAD case status was defined as >50% stenosis in either carotid artery as determined by ultrasound, while controls were also imaged and had <15% stenosis in both carotid arteries and absence of peripheral artery disease (PAD) or CHD. The cohort consisted of 688 CAAD cases and 1037 controls. The few participants with moderate carotid stenosis (15-49% obstruction in at least one carotid artery) were excluded from analysis. Censored age was the age at CAAD diagnosis for cases and age at enrollment and blood draw for controls. Current smoking status was obtained by self-report. Insulin use was determined via self-report matched to hospital pharmacy records.

The study population for this analysis consisted of 1,725 European-ancestry participants from the previously described CLEAR study[28-30,53-56]. To avoid population stratification, smaller numbers of non-European-ancestry subjects were excluded from all analyses presented here. European genetic ancestry was determined using STRUCTURE[57] and SNPs from the Illumina CVD chip[58,59] or 550k BeadChip data. Descriptive statistics of the cohort are presented in **Table 2.1**.

Table 2.1: Baseline characteristics of the studied European-ancestry subset of CLEAR, stratified by sex.

	Females (N=524)	Males (N=1201)	Combined (N=1725)
apoA1, IU	166 ± 28	138 ± 26	146 ± 29
HDL-C, mg/dl	63 ± 18	47 ± 15	52 ± 18
HDL-2, mg/dl	14.7 ± 7.8	8.5 ± 5.3	10.4 ± 6.8
HDL-3, mg/dl	49 ± 11	39 ± 11	42 ± 12
PON1 AREase activity, IU	167 ± 58	132 ± 49	143 ± 55
Current Smoker	6% (32)	17% (199)	13% (231)
Statin Use	24% (124)	40% (482)	35% (606)
Diabetic	9% (46)	22% (268)	18% (314)
Censored age, years*	63.6 ± 9.8	68.0 ± 9.4	66.6 ± 9.7
CAAD Status	19% (102)	49% (586)	40% (688)

Abbreviations: apoA1 = apolipoprotein A1; AREase = PON1 arylester hydrolysis rate; CAAD = carotid artery disease; HDL-C = high-density lipoprotein cholesterol.
Mean ± 1 SD.
Numbers after percents are counts.
*Censored age defined as the age at CAAD diagnosis (for CAAD cases) or the age at enrollment of controls.

Lipid Measurements

Standard methods were used to determine total cholesterol, triglycerides, and HDL in fasting whole plasma using an Abbott Spectrum analyzer. HDL fractions 2 and 3 were determined by precipitating HDL2 from isolated total HDL, measuring HDL3 in the supernatant, and subtracting this from total HDL to obtain HDL2. Apolipoprotein A-I was measured as previously described[60]. PON1 activity was measured by the rate of enzymatic degradation of phenylacetate (AREase) via a continuous spectrophotometric assay with lithium heparin plasma, as AREase is least affected by the functional $PON1_{Q192R}$ polymorphism and also is more closely related to

PON1 protein levels[61,62]. PON1 AREase activity was measured in triplicate and averaged. All lipid and associated protein measurements had approximate standard distributions. All data were generated blinded to CAAD status.

Statistical Analyses

Analyses were done in R (http://www.r-project.org/) using the available standard regression tools. All subjects had complete phenotype and covariate data for regression analyses. As there is a known sex-dependent difference in HDL levels, we chose to perform sex-stratified analyses. Sex-specific analyses comparing risk factors between CAAD cases and controls used either the Wilcoxon rank sum (for continuous variables) or Pearson's chi-square (for categorical variables) tests to determine significance. Correlation of covariates was summarized using Pearson's pairwise correlation coefficient.

Given the high correlation between the measurements, we performed stepwise logistic regression on the phenotype of CAAD, with HDL-C, HDL2, HDL3, and apoA1 entering the model. Model comparison was done using Akaike's information criterion (AIC), beginning with a base model comprised of the known confounders of CAAD status: censored age, diabetes status, and current smoking status. The measurement that best improved model prediction of CAAD via AIC was retained in the final model. A secondary analysis considered the addition of PON1 AREase activity to HDL3 levels and performed stepwise logistic regression, also considering censored age, current smoking, and diabetes status, to determine if PON1 AREase activity predicted additional CAAD variance independent of the other lipid measurements.

Statin drug use can also affect HDL levels. However, as CAAD is treated with statins, statin use was confounded for CAAD status and could not be included in a model predicting CAAD. We note that in 57 male CLEAR participants with repeat lipid measurements before and after statin initiation, although HDL levels do increase with statin use, the increase was not statistically significant (HDL-C, p=0.14; HDL2, p=0.57; HDL3, p=0.27; see **Table 2.2**).

Table 2.2: Effect of statin use on HDL-C, HDL-2, and HDL-3 concentration in a male-only subset of CLEAR with repeat lipid measures before and after statin use (n=57).

	Before Statin	After Statin	P-Value*
HDL-C	42.63 ± 13.03	46.44 ± 13.78	0.14
HDL-2	6.91 ± 3.99	7.28 ± 4.23	0.57
HDL-3	35.72 ± 9.47	37.38 ± 10.51	0.27

Mean ± 1 SD.
Tests used: *Two-sample, two-sided t-test without the assumption of equal variance.

2.4 RESULTS

Demographic, clinical, and lipid variables of the studied European-ancestry subset of the CLEAR study are presented in **Table 2.1**. The cohort was composed of 1201 males and 524 females, of which 586 (49%) and 102 (19%) participants were CAAD cases for men and women, respectively. Females had lower rates of smoking (6% vs. 19%), statin use (24% vs. 40%), and diabetes (9% vs. 22%). Additionally, females had higher levels of all HDL phenotypes (HDL-C, HDL2, HDL3, apoA1 and PON1). Due to these sex-dependent differences in lipid and clinical covariates, stratified analyses were conducted to determine which HDL measure was the better predictor of CAAD.

Table 2.3: Association of baseline lipid and clinical characteristics with CAAD status in men (N=1201).

	Controls (N=615)	CAAD Cases (N=586)	P-Value[†,‡]
apoA1, IU	142 ± 26	133 ± 24	<0.001[†]
HDL-C, mg/dl	50 ± 16	44 ± 13	<0.001[†]
HDL-2, mg/dl	9.2 ± 5.7	7.9 ± 4.8	<0.001[†]
HDL-3, mg/dl	40.6 ± 11.2	36.4 ± 9.4	<0.001[†]
PON1 AREase activity, IU	145 ± 49	119 ± 46	<0.001[†]
Current Smoker	9% (54)	25% (145)	<0.001[‡]
Statin Use	18% (112)	63% (367)	<0.001[‡]
Diabetic	13% (82)	32% (186)	<0.001[‡]
Censored age, years*	64.9 ± 9.0	71.2 ± 8.7	<0.001[†]

Abbreviations: apoA1 = apolipoprotein A1; AREase = PON1 arylester hydrolysis rate; CAAD = carotid artery disease; HDL-C = high-density lipoprotein cholesterol.
Mean ± 1 SD.
Numbers after percents are frequencies.
*Censored age defined as the age at CAAD diagnosis (for CAAD cases) or the age at enrollment of controls.
Tests used: [†]Wilcoxon rank sum test; [‡]Pearson chi-square test.

There were significant ($p < 0.05$) differences in all clinical and lipid covariates between CAAD cases and controls for both men (see **Table 2.3**) and women (see **Table 2.4**). CAAD cases had significantly lower levels of apoA1, HDL-C, HDL2, HDL3, and PON1 AREase activity compared to controls in both men and women, and also had significantly higher rates of current

smoking and diabetes. In addition, censored age was higher in both male and female CAAD cases when compared to controls.

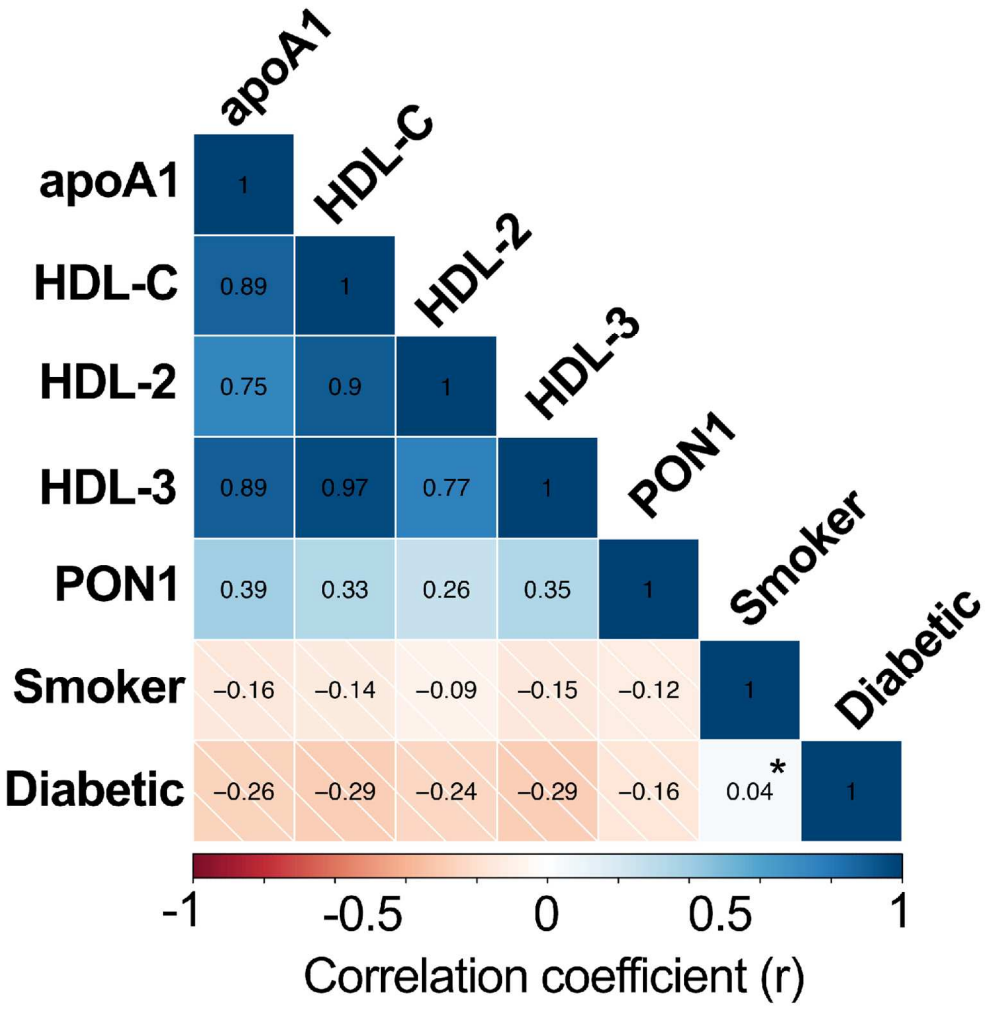
Table 2.4: Association of baseline lipid and clinical characteristics with CAAD status in women (N=524).

	Controls (N=422)	CAAD Cases (N=102)	P-Value[†,‡]
apoA1, IU	167 ± 27	160 ± 30	0.021[†]
HDL-C, mg/dl	65 ± 17	59 ± 19	<0.001[†]
HDL-2, mg/dl	15.1 ± 7.7	13.1 ± 7.7	0.003[†]
HDL-3, mg/dl	50 ± 11	45 ± 12	<0.001[†]
PON1 AREase activity, IU	171 ± 58	150 ± 57	<0.001[†]
Current Smoker	5% (20)	12% (12)	0.008[‡]
Statin Use	14% (57)	66% (67)	<0.001[‡]
Diabetic	5% (21)	25% (25)	<0.001[‡]
Censored age, years*	61.9 ± 9.1	70.7 ± 9.7	<0.001[†]

Abbreviations: apoA1 = apolipoprotein A1; AREase = PON1 arylester hydrolysis rate; CAAD = carotid artery disease; HDL-C = high-density lipoprotein cholesterol.
Mean ± 1 SD.
Numbers after percents are frequencies.
*Censored age defined as the age at CAAD diagnosis (for CAAD cases) or the age at enrollment of controls.
Tests used: [†]Wilcoxon rank sum test; [‡]Pearson chi-square test.

The pairwise correlation between each of the studied clinical and lipid covariates was considerable (see **Figure 2.1**). ApoA1, HDL-C, HDL2, and HDL3 were all strongly and positively correlated with each other (pairwise correlation coefficients, rho (r) $\geq$ 0.75). PON1 AREase activity was also positively correlated with the other lipid measurements, though not as strongly (r $\geq$ 0.26). Current smoking status and diabetes were each negatively correlated with each of the aforementioned lipid measurements, though they were not highly correlated with each other (r = 0.04).

Figure 2.1: Correlation matrix for plasma lipid measurements. Values inside each box represent r, the pairwise correlation coefficient, unadjusted for covariates.



In males, HDL3 was the lipid measurement that explained the greatest amount of CAAD variation (1.6%) and was negatively associated with CAAD (odds ratio (OR) with 95% confidence interval (CI) = 0.97 (0.95-0.98), p = 0.00011, see **Table 2.5**). In addition, PON1 AREase activity was negatively associated with and improved model prediction of CAAD (0.7% of CAAD variation, OR with 95% CI = 0.99 (0.98-0.99), p < 0.001) in males. A 10 mg/dl increase in HDL3 was calculated to decrease the odds of CAAD by approximately 26% (OR = 0.74 (0.63-0.86)), while a 10 IU increase in PON1 AREase activity was estimated to

independently decrease the odds of CAAD by approximately 7% (OR = 0.93 (0.89-0.96)). In the smaller female group, HDL3 was the only HDL phenotype included in the final regression model (0.7% of CAAD variation, OR with 95% CI = 0.97 (0.95-0.99), p = 0.042, see **Table 2.6**). In females, a 10 mg/dl increase in HDL3 was calculated to decrease the odds of CAAD by approximately 24% (OR = 0.76 (0.58-0.99)). Although underpowered, post-hoc analysis found no evidence of interaction (p > 0.05) between either HDL3 levels or PON1 AREase activity with any demographic (censored age) or clinical covariate (smoking and diabetes status).

Table 2.5: Best-fit model from stepwise linear regression predicting CAAD status in men using lipid and clinical covariates (N=1201).

	Odds Ratio (95% CI)	% CAAD Variation	P-Value
Censored age, years*	1.10 (1.09-1.13)	11.18%	<0.001
Current Smoker	5.63 (3.74-8.62)	7.31%	<0.001
Diabetic	3.41 (2.38-4.95)	4.44%	<0.001
HDL-3, mg/dl	0.97 (0.95-0.98)	1.59%	0.00011
PON1 AREase, IU	0.99 (0.98-0.99)	0.68%	<0.001

Abbreviations: AREase = PON1 arylester hydrolysis rate; CAAD = carotid artery disease.
Mean ± 1 SD.
*Censored age defined as the age at CAAD diagnosis (for CAAD cases) or the age at enrollment of controls.

Table 2.6: Best-fit model from stepwise linear regression predicting CAAD status in women using lipid and clinical covariates (N=524).

	Odds Ratio (95% CI)	% CAAD Variation	P-Value
(Intercept)	0.00063 (0.00005-0.007)	-	<0.001
Censored age, years*	1.11 (1.08-1.15)	12.62%	<0.001
Current Smoker	3.56 (1.48-8.31)	5.05%	0.0036
Diabetic	3.89 (1.82-8.35)	1.51%	<0.001
HDL-3, mg/dl	0.97 (0.95-0.99)	0.70%	0.042

Abbreviations: CAAD = carotid artery disease; CI = confidence interval.
Mean ± 1 SD.
*Censored age defined as the age at CAAD diagnosis (for CAAD cases) or the age at enrollment of controls.

2.5 DISCUSSION

In the current study, we have used a CAAD case-control study of European-ancestry subjects to evaluate the effects of apoA1, HDL-C, HDL2 HDL3, and PON1 AREase activity on CAAD risk separately in men (n=1201) and women (n=524). Using stepwise regression to find the best predictor of CAAD from the highly correlated lipid measurements of apoA1, HDL-C, HDL2, and HDL3, we have identified HDL3 as the HDL phenotype that captures the most CAAD status variance in both men and women. With HDL3 in the regression model, none of the remaining HDL-related measurements, except PON1 AREase level, improved the model for CAAD. We noted no potential confounding of age on HDL lipoprotein levels that could negate our findings. PON1 AREase activity had a significant impact on CAAD risk that was independent of HDL3 levels in males only. To the best of our knowledge, this represents the first CAAD case-control study with population-based controls to evaluate the effects of HDL sub-species on CAAD risk.

Prior work on the associations of HDL2 and HDL3 with CVD has yielded conflicting results. In a recent literature review, it was reported that 45% of 37 total case-control studies found a statistically significant decrease in CVD cases for both HDL2 and HDL3 levels, 26% found a significant decrease in HDL2 only, 11% found a significant decrease in HDL3 only, and 17% found no statistically significant decrease in either HDL sub-fraction[637]. The vast majority of these studies collected CHD cases; to the best of our knowledge, only one prior study has looked at CAAD[28]: Atger et al. examined 181 asymptomatic hypercholesterolemic men by ultrasound to determine the presence of CAAD, in addition to femoral and abdominal aorta stenosis. No significant difference in either HDL2 or HDL3 levels was found between 43 men with CAAD compared to the 138 other men[28]. However, we note that their case sample size was only 7% of

our size (586 here vs. 43 in their study); thus, they lacked power to detect the effects of HDL3 we identified.

The majority of prior work on the subject of HDL sub-fractions and CVD was done in studies of coronary heart disease (CHD) and found a stronger association of HDL2 with cardioprotection[21,22,63-652]. One hypothesis for this correlation related to the much higher density of apoA1 on the larger HDL2 molecule as compared to HDL3. We evaluated this hypothesis in the current study through inclusion of apoA1 as one of the possible HDL-related covariates in the stepwise regression model. However in our analyses, HDL3 was the best predictor of CAAD variance and with HDL3 in the regression model, none of the other HDL-related covariates (HDL-C, HDL2, or apoA1) were able to improve prediction of CAAD status. We note that HDL3 is associated with other apolipoproteins; namely, A2, A3, A4, and pre-B1[19]. However, as these were not measured in the current study, we were unable to evaluate their individual effects on CAAD risk.

HDL3 is the smaller, denser, and more lipid-poor of the two sub-fractions of HDL. HDL3 is strongly antioxidant, and in prior work it has been demonstrated that the antioxidant capability of HDL increases with density[233]. In addition, HDL3 is closely associated with the glycoprotein enzyme, PON1[664]. PON1 is itself atheroprotective[28,31,67,68] and can prevent LDL[25,2639] and HDL oxidation[69] (other functions of PON1 are summarized in a recent review article[701]). To address the possibility that our association of HDL3 with CAAD was due to the effects of its association with PON1, we included PON1 AREase activity as a covariate in the stepwise regression model. Interestingly, in men both HDL3 and PON1 AREase activity were retained in the final model,

suggesting that the atheroprotective effects of HDL3 are independent and additive of PON1 enzyme activity. As noted previously, the molecular mechanisms for this association of HDL3 with CAAD could be due to unmeasured apolipoproteins with which HDL3 is correlated. It is notable that HDL3 likely has antioxidant properties that are independent of PON1[23]. The lack of detection of an additional PON1 effect in women may have been due to insufficient statistical power in that smaller group.

Our finding of HDL3 being the best predictor of CAAD status is incongruent with many studies that have used similar multivariate regression methods and found HDL2 as being more significantly associated. However, as noted above, the majority of past work regarding HDL2 and HDL3 has focused on CHD rather than CAAD. Though both disease states are driven by atherosclerosis, the underlying pathogenesis is likely different[71], as evidenced by the divergence of genes that are associated with myocardial infarction[72,73] vs. ischemic stroke[74,756] or carotid intimal media thickening[76] (No genome-wide association study has been performed for CAAD to date). In this context our finding of HDL3 being the most strongly associated, independent of PON1, may represent further evidence that the pathology of CAAD is distinct.

Several limitations of the current study must be considered. First, this cohort was comprised only of individuals of European-ancestry, limiting inferences from our data to participants of other races. Second, males consisted the majority of both the cohort and of the CAAD cases (102 vs. 586 for males), leaving our female-only analyses statistically less powered. Third, due to confounding with CAAD status our analyses could not adjust for the effects of statin use on HDL-C, HDL2, and HDL3 levels. Although under-powered, there was no statistically

significant evidence for an increase in HDL levels in 57 male subjects pre- and post-statin initiation data. Similarly, data from the COMPELL study, which studied the effects of 4 different statins on lipid profiles in 292 participants (50% female) did not show a statistically significant increase in HDL3 levels at 8 or 12 week follow-up. Regardless, the excess of statin use in cases would result in more conservative testing of our hypothesis of HDL phenotype differences between CAAD cases and controls if statin use does increase HDL levels as found elsewhere[48].

In conclusion, our analysis of HDL sub-fraction data and CAAD has found that of the sub-phenotypes apoA1, HDL-C, HDL2, and HDL3, HDL3 best predicts CAAD risk and the remaining phenotypes do not add significant predictive power. The effects of HDL3 were independent of and additive with PON1 enzyme activity. Given the importance of CAAD as a risk factor for ischemic stroke and also as a marker of CVD, further work elucidating the molecular mechanisms through which HDL3 is cardioprotective for CAAD is warranted.

Disclosures: The authors declare they have no conflicts of interest.

Acknowledgements: We would like to thank all participants of the CLEAR study. This work was funded in part by National Institutes of Health RO1 HL67406 and a State of Washington Life Sciences Discovery Award (265508) to the Northwest Institute of Genetic Medicine. DSK was supported in part by the Benjamin and Margaret Hall Endowed Fellowship in Genome Sciences, a Markey Foundation award, and National Institutes of Health 5T31HG000035-18 and 1F31MH101905-01.

CHAPTER 3: SMALL/MEDIUM HDL-P AND ITS ROBUST ASSOCIATION WITH CAROTID INTIMA MEDIA THICKNESS

This chapter is based on the following peer-reviewed manuscript (in review as of 2015-12-18):

Kim DS, Li YF, Bell GA, Burt AA, Vaisar T, Hutchins PM, Furlong CE, Otvos JD, Polak JF, Arnan MK, Kaufman JD, McClelland RL, Longstreth WT, Jarvik GP. Concentration of smaller high-density lipoprotein particle (HDL-P) is inversely correlated with carotid intima media thickness after confounder adjustment: the Multi Ethnic Study of Atherosclerosis (MESA). *Under review at the Journal of the American Heart Association*.

3.1 SUMMARY

Purpose: Recent studies have failed to establish a causal relationship between high-density lipoprotein cholesterol levels (HDL-C) and cardiovascular disease (CVD), shifting focus to other measures of HDL. We previously reported that smaller and denser HDL levels are protective against cerebrovascular disease. This study sought to determine which of small/medium HDL particle concentration (HDL-P) or large HDL-P was more strongly associated with carotid intima-media thickness (cIMT) in an ethnically diverse cohort.

Methods: In cross-sectional analyses of participants from the Multi Ethnic Study of Atherosclerosis (MESA), we evaluated the associations of nuclear magnetic resonance spectroscopy-measured small/medium vs. large HDL-P with cIMT measured in the common and internal carotid arteries, through linear regression.

Results: After adjustment for CVD confounders, low-density lipoprotein cholesterol (LDL-C), HDL-C, and large HDL-P, small/medium HDL-P remained significantly and inversely associated with common (coefficient=-1.68 μm, p=0.00042, n=6512) and internal cIMT (coefficient=-4.26 μm, p=0.0073, n=6418) after Bonferroni correction for 4 independent tests (threshold for significance=0.0125, α=0.05/4). In contrast, the association of large HDL-P with both common (p=0.36, n=6512) and internal cIMT (p=0.59, n=6418) was attenuated after adjustment for small/medium HDL-P, HDL-C and other CVD confounders. In a separate sample of 64 men, small/medium HDL-P was more strongly correlated with paraoxonase-1 activity and phospholipid transfer protein (PLTP) activity (r=0.32 and r=0.19, respectively) as compared to both total HDL-P (r=0.29 and r=0.18) and large HDL-P (r=0.18 and r=0.08) measures.

Conclusions: Small/medium HDL-P is significantly and inversely correlated with cIMT measurements. Correlation of small/medium HDL-P with cardioprotective paraoxonase-1 activity and PLTP activity may reflect functional aspects of HDL responsible for this finding.

3.2 BACKGROUND

Several large randomized clinical trials[42,44,77] and Mendelian randomization studies[43] have failed to show a causal role for high-density lipoprotein cholesterol levels (HDL-C) in cardioprotection, decreasing interest in HDL as a therapeutic target. However, recent work from the Multi Ethnic Study of Atherosclerosis (MESA) found that total HDL particle (HDL-P) concentration is a significant predictor of incident cardiovascular events and carotid intima media thickness (cIMT), even when adjusting for the effects of HDL-C and other cardiovascular disease (CVD) confounders[47]. This study suggested that unmeasured aspects of HDL not reflected by HDL-C might be responsible for the long-held and consistent cardioprotective associations of HDL[7].

As previously described, HDL is a heterogeneous and complex particle, with over 60 associated proteins[14]. Despite its complexity, HDL can be sub-divided by size and density. In the previous chapter, the smaller and denser HDL3 sub-fraction was found to be the best predictor of reduced CAAD risk, demonstrating the utility of the HDL3 measure vs. HDL-C, HDL2, or apoA1 [48]. Follow-up work in a subset of this CAAD case-control study found that differences in small and medium HDL-P concentration were the primary determinants of HDL-P concentration differences between cases and controls[78]. In a separate study of approximately 1000 stroke-free subjects, HDL3 was negatively associated with carotid plaque area, while HDL2 was positively associated[79]. These results support the hypothesis that cholesterol-poor HDL subspecies, which are underrepresented in measures of HDL-C, may be important protective factors against cerebrovascular disease.

Given these findings, we sought to validate the association of HDL3 with cerebrovascular disease phenotypes in an independent cohort. As HDL3 is closely approximated by the sum of small and medium HDL-P concentration, our primary goal in the present study was to determine whether small/medium HDL-P concentration was associated with decreased common and internal cIMT in a cross-sectional analysis of MESA participants. We simultaneously sought to evaluate whether large HDL-P concentration was similarly cardioprotective in cross-sectional analyses after covariate adjustment. Finally, as a follow-up analysis we evaluated the correlation between the studied HDL-P subspecies and paraoxonase-1 (PON1) and phospholipid transfer protein activities (for detailed explanations of these activities and their functions, please refer to **Chapters 4** and **5** for PLTP and PON1 activities, respectively), which both have previously been reported as protective against CAAD[28-30,49] in the CLEAR study. The MESA cohort did not have these PON1 and PLTP activity data available.

3.3 METHODS

Ethics Statement

Informed, written consent was obtained from each participant of MESA at the time of recruitment between July 2000 and August 2002. Institutional review boards at each participating center of MESA approved of the study. The University of Washington institutional review board approved analyses presented in this work.

MESA Cohort Sample

MESA is a multicenter, longitudinal cohort study of the prevalence and progression of subclinical atherosclerotic disease phenotypes. Between July 2000 and August 2002, MESA recruited 6814 participants between 45 and 84 years of age from 6 field sites: Baltimore City and Baltimore County, Maryland; Chicago, Illinois; Forsyth County, North Carolina; Los Angeles County, California; northern Manhattan and the Bronx, New York; and St Paul, Minnesota. Participants were of African American, Hispanic, white, and Chinese-American self-reported ancestry. Baseline exclusion criteria for recruitment were: self-reported cardiovascular disease (CVD) history, pregnancy, cancer, cognitive impairment, or weight >136 kg. At the time of recruitment, each participant underwent examination including blood draw, anthropometric measurements, and collection of questionnaire data.

Lipid Measurements

In MESA, all lipid and covariate data were collected at physical examination during study enrollment. Blood was drawn after a 12 hour fast, and samples were stored at -70^{o}C with EDTA. Plasma HDL-P concentration was measured at LipoScience, Inc. (Raleigh, North Carolina) using

NMR spectroscopy and the LipoProfile-3 algorithm[80,81]. All other lipids and fasting glucose were measured at a separate central laboratory (Collaborative Studies Clinical Laboratory at Fairview University Medical Center, Minneapolis, Minnesota). HDL-C was measured after precipitation of non-HDL cholesterol through exposure to magnesium/dextran, using the cholesterol oxidase method (Roche Diagnostics, Indianapolis, Indiana). LDL-C was calculated using the Friedwahl equation[82]. The mean HDL-C and LDL-C were 50.9 and 117.2 mg/dl, respectively. Large HDL-P represented the minority of the total HDL-P concentration, with a mean of 6.02 μmol/L as compared to 28.01 μmol/L for small/medium HDL-P.

Covariates

Demographic information such as body mass index (BMI), blood pressure, smoking history, and medication use were all collected at the baseline MESA physical examination. Smoking was defined as never, former (smoked ≥100 cigarettes in a lifetime), or current. Pack-years of smoking were also calculated from this survey information. Hypertension was defined as systolic blood pressure ≥ 140 mmHg, diastolic blood pressure ≥ 90 mmHg, self-reported history of hypertension, or use of antihypertensive medication. Diabetes status was defined as a fasting glucose ≥ 126 mg/dl or use of an antidiabetic medication.

Carotid Intima Media Thickening Measurement

High-resolution B-mode ultrasound was used to measure carotid atherosclerosis as previously described[83]. Total cIMT in micrometers (μm) was calculated separately from maximal carotid artery thickness at 4 sites (right and left; near and far walls) for the common and internal carotid

arteries, respectively[84]. Common and internal artery cIMT had means of 0.87 and 1.07 mm, respectively.

CLEAR Study Sample

The CLEAR study, described in **Chapter 2**, is a Seattle-based prevalent CAAD case-control study, comprised primarily of veterans, with controls distribution matched by sex and age at diagnosis of CAAD cases. The analyzed subset of CLEAR described here is composed of 64 men without CVD not on statin pharmacotherapy. PLTP activity was measured by transfer of ^{14}C-phosphatidylcholine to an HDL acceptor, as previously described (and will be described in greater detail in **Chapter 4**)[49,85].

Statistical Analysis

All analyses and graphics were performed in R (http://r-project.org).

For our primary analyses, the outcomes of total common and internal cIMT were separately analyzed for association with the sum of small/medium HDL-P and large HDL-P concentrations. We used a linear regression model adjusting for the confounders of age, sex, race (coded as a dummy variable with whites, the largest subgroup, as the reference group and blacks, Hispanics, and Asians as the comparison groups), BMI, cigarette smoking status (former or current, with the never group as reference), pack-years of smoking, systolic blood pressure, anti-hypertensive medication use, fasting glucose, diabetes status, lipid-lowering medication use, LDL-C, and HDL-C. A Bonferroni-corrected threshold of significance of 0.0125 (α=0.05/4) was applied due

to the 4 tests in our primary hypothesis: common cIMT with small/medium HDL-P and large HDL-P and internal cIMT with small/medium HDL-P and large HDL-P.

To aid in interpretation, we present results in a staged regression model, adjusting first for demographic (age, sex, race), then adding cardiovascular risk factors (BMI, cigarette smoking status, pack-years of smoking, systolic blood pressure, anti-hypertensive medication, fasting glucose, diabetes status, and lipid-lowering medication use), subsequently adding LDL-C, and finally adding HDL-C to the linear regression model. We note that staged regression models were not used in the primary analyses but are presented in table form to illustrate the effect of HDL-C on the association of small/medium HDL-P vs. large HDL-P on the cIMT outcomes.

To illustrate multivariable regression results, we separately plotted covariate-adjusted mean internal and common cIMT for quartiles of small/medium HDL-P and large HDL-P, respectively. Two lines are presented in each plot: one for covariate-adjusted internal or common cIMT with all covariates except HDL-C in the model and a second line with all covariates in the final regression model.

Sensitivity analyses were performed by race group and sex for the association of small/medium HDL-P with both common and internal cIMT. Significant differences in association of small/medium HDL-P with either outcome were followed-up with addition of a formal interaction term to the regression model (e.g., male sex * small/medium HDL-P). Presence of effect modification was concluded if the interaction term was significantly different than zero ($P \leq 0.05$).

3.4 RESULTS

MESA participants analyzed in this work were multiethnic men and women without baseline CVD with a mean age of 62.2 years (see **Table 3.1**). Men composed 47.2% of the study population. Self-reported whites were the largest ethnic group, representing 38.5% of the participants, followed by blacks (27.8%), Hispanics (21.9%), and Asians (11.8%). Baseline clinical and demographic characteristics of the MESA participants are summarized in **Table 3.1**.

Table 3.1: Baseline MESA cohort demographic and clinical characteristics.

	N	Mean (± SD) or N (%)
Age, years	6814	62.2 (± 10.2)
Race	6814	
White, %		2622 (38.5%)
Asian, %		804 (11.8%)
Black, %		1892 (27.8%)
Hispanic, %		1496 (21.9%)
Men, %	6814	3213 (47.2%)
Body mass index, kg/m^2	6814	28.3 (± 5.48)
Smoking status	6792	
Never, %		3418 (50.3%)
Former, %		2487 (33.6%)
Current, %		887 (13.1%)
Cigarette pack-years	6720	11.3 (± 20.9)
Systolic blood pressure, mmHg	6811	126.6 (± 21.5)
Antihypertensive Rx, %	6811	2536 (37.2%)
Fasting glucose, mg/dl	6789	104.5 (± 30.9)
Diabetes status, %	6814	859 (12.6%)
Lipid-lowering Rx, %	6811	1100 (16.2%)
LDL-C, mg/dl	6701	117.2 (± 31.5)
HDL-C, mg/dl	6701	50.9 (± 14.8)
Large HDL-P, μmol/L	6786	6.02 (± 3.46)
Small/Medium HDL-P, μmol/L	6786	28.01 (± 5.48)
Common cIMT, μm	6726	870.51 (± 193.43)
Internal cIMT, μm	6629	1071.52 (± 603.44)

Abbreviations: cIMT = carotid intima media thickening; HDL-C = high-density lipoprotein cholesterol; HDL-P = high-density lipoprotein particle concentration; LDL-C = low-density lipoprotein cholesterol; Rx = medication.

After full adjustment, including LDL-C, HDL-C, and large HDL-P, small/medium HDL-P concentration was significantly and inversely associated common cIMT at a Bonferroni-

corrected threshold of α=0.0125 (*P*=0.00042, see **Table 3.2**), with a final covariate-adjusted beta coefficient (aBeta) of -1.68 μm. In contrast, after full adjustment, large HDL-P concentration was not significantly associated with common cIMT (*P*=0.36). Full regression coefficients for all covariates are presented for the common cIMT multivariable regression model in **Supplemental Table A1**.

A similar pattern was observed for associations with cIMT of the internal carotid artery. After full adjustment, small/medium HDL-P concentration was significantly and inversely associated at a Bonferroni-corrected threshold of α=0.0125 with internal cIMT (*P*=0.0073, aBeta=-4.26 μm, see **Table 3.3**). In contrast, after full adjustment, large HDL-P concentration was not significantly associated with internal cIMT (*P*=0.59). Full regression coefficients for all covariates are presented for the internal cIMT multivariable regression model in **Supplemental Table A1**.

Notably, in *post-hoc* analyses, large HDL-P was strongly and inversely associated with common and internal cIMT after adjustment for demographic covariates, CVD risk factors, and LDL-C. However, addition of HDL-C caused both of these associations to attenuate (see **Tables 3.2** and **3.3**). A similar trend was observed with full adjustment in a multivariable linear regression model including both small/medium HDL-P and large HDL-P, with large HDL-P being not significantly associated with either cIMT outcome. In contrast, the inverse association of small/medium HDL-P with both common and internal cIMT remained robust even after the addition of HDL-C and large HDL-P to the model (see **Tables 3.2** and **3.3**). Graphical representations of the effect of HDL-C and either small/medium or large HDL-P on regression

Table 3.2: Staged linear regression model association of small/medium HDL-P and large concentration with carotid intima media thickening in the common carotid arteries (n=6512).

	Small/Medium HDL-P		*Large HDL-P*	
	aBeta ± SE	**P-Value**	**aBeta ± SE**	**P-Value**
Unadjusted effect of small/medium HDL-P	-1.63 ± 0.43	0.00024	-4.15 ± 0.69	<0.0001
+Demographic characteristics*	-0.95 ± 0.41	0.019	-5.45 ± 0.67	<0.0001
+Above and CVD Risk Factors†	-1.94 ± 0.40	<0.0001	-2.95 ± 0.69	<0.0001
+Above and LDL-C	-1.76 ± 0.39	<0.0001	-2.28 ± 0.69	0.0011
+Above and HDL-C	-1.46 ± 0.41	0.00037	-1.55 ± 1.50	0.30
+Above and small/medium and large HDL-P	-1.68 ± 0.48	0.00042	-1.61 ± 1.74	0.36

Final regression model coefficients/p-values for association with common cIMT available in Supplemental Tables S1 (small/medium HDL-P) and S2 (large HDL-P).
Abbreviations: cIMT = carotid intima media thickening; HDL-C = high-density lipoprotein cholesterol; HDL-P = high-density lipoprotein particle concentration; LDL-C = low-density lipoprotein cholesterol.
*****Demographic characteristics are defined as age, sex, and race (Asian, black, or Hispanic, with white as the reference group).
†Cardiovascular disease (CVD) risk factors defined as body mass index, cigarette smoking status (former or current, with never group as reference), pack-years of smoking, systolic blood pressure, anti-hypertensive medication, fasting glucose, diabetes status, and lipid-lowering medication use.

Table 3.3: Staged linear regression model association of small/medium HDL-P and large concentration with carotid intima media thickening in the internal carotid arteries (n=6418).

	Small/Medium HDL-P		*Large HDL-P*	
	aBeta ± SE	**P-Value**	**aBeta ± SE**	**P-Value**
Unadjusted effect of small/medium HDL-P	-2.88 ± 1.38	0.037	-10.91 ± 2.18	<0.0001
+Demographic characteristics*	-2.15 ± 1.34	0.11	-12.17 ± 2.22	<0.0001
+Above and CVD Risk Factors†	-4.91 ± 1.33	0.00036	-6.36 ± 2.28	0.0054
+Above and LDL-C	-4.45 ± 1.33	0.00081	-4.62 ± 2.31	0.045
+Above and HDL-C	-3.82 ± 1.36	0.0051	-4.84 ± 4.99	0.33
+Above and small/medium and large HDL-P	-4.23 ± 1.59	0.0073	-3.13 ± 5.80	0.59

Final regression model coefficients/p-values for association with internal cIMT are available in Supplemental Tables S3 (small/medium HDL-P) and S4 (large HDL-P).
Abbreviations: cIMT = carotid intima media thickening; HDL-C = high-density lipoprotein cholesterol; HDL-P = high-density lipoprotein particle concentration; LDL-C = low-density lipoprotein cholesterol.
*****Demographic characteristics are defined as age, sex, and race (Asian, black, or Hispanic, with white as the reference group).
†Cardiovascular disease (CVD) risk factors defined as body mass index, cigarette smoking status (former or current, with never group as reference), pack-years of smoking, systolic blood pressure, anti-hypertensive medication, fasting glucose, diabetes status, and lipid-lowering medication use.

model prediction of common and internal cIMT are presented in **Figures 3.1** and **3.2**, respectively for both small/medium and large HDL-P. Correlation coefficients among small/medium HDL-P, large HDL-P, total HDL-P, and HDL-C for 6786 MESA participants are presented in **Figure 3.3**.

Sensitivity analyses were performed for the association of small/medium HDL-P with common cIMT (**Supplemental Tables A2**) and internal cIMT (**Supplemental Tables A3**) across racial groups. Notably, after full adjustment, small/medium HDL-P concentration was consistently associated with decreased common and internal cIMT for whites (n=2622, aBeta=-1.33 and *P*=0.075 with common cIMT and aBeta=-5.51 and *P*=0.032 with internal cIMT). In blacks and Hispanics, small/medium HDL-P was associated with decreased common cIMT (Blacks n=1892, aBeta=-1.58 and *P*=0.089, Hispanics (n=1496), aBeta=-2.77 and *P*=0.0089), but not with internal cIMT (aBeta=-3.54 and *P*=0.26 and aBeta=-2.66 and *P*=0.44 for blacks and Hispanics, respectively; see **Supplemental Tables A2-A3**). Finally, after full adjustment, small/medium HDL-P was not associated with either cIMT outcome in Asians (n=804, aBeta=-0.88 and *P*=0.53 with common cIMT and aBeta=-3.59 and *P*=0.37 with internal cIMT; see **Supplemental Tables A2-A3**). Statistical interaction tests of race group and small/medium HDL-P were conducted on full regression models predicting both common and internal cIMT. No race-by-small/medium HDL-P interactions were statistically significant (*P*>0.05) for either common (black interaction *P*=0.56, Hispanic interaction *P*=0.21, Asian interaction *P*=0.59) and internal cIMT (black interaction *P*=0.71, Hispanic interaction *P*=0.64, Asian interaction *P*=0.78) outcomes.

Separate sensitivity analyses were performed for the association of small/medium HDL-P with common cIMT and internal cIMT across sex (see **Supplemental Table A4**). After full adjustment, small/medium HDL-P concentration was consistently associated with decreased common cIMT in men (n=3173, aBeta=-3.54 and $P<0.0001$) but not in women (n=3553, aBeta=-0.70 and $P=0.22$). In sensitivity analyses of internal cIMT, both women and men had marginally significant associations of small/medium HDL-P correlating with a decrease in internal cIMT (Females n=3492, aBeta=-3.69 and $P=0.055$; male n=3137, aBeta=-4.91, and $P=0.076$; results summarized in **Supplemental Table A4**). Statistical interaction tests of gender (with females as the reference) and small/medium HDL-P were conducted on full regression models predicting both common and internal cIMT. The male gender-by-small/medium HDL-P interaction for a full linear regression model of internal cIMT was not significant ($P=0.74$); however, the interaction for common cIMT was significant ($P=0.0065$).

To determine whether PON1 or PLTP activities could potentially mediate the inverse association of small/medium HDL-P with common and internal cIMT, we performed a follow-up analysis in the CLEAR study where PON1 and PLTP measures were available. For this analysis, 64 men without CVD and not taking statins were evaluated for HDL-P concentrations, HDL2, HDL3, PON1 AREase activity, and PLTP activity. Pearson's pairwise correlation coefficients were then calculated for each subgroup of HDL-P and PON1 AREase activity (see **Figure 3.4**). PON1 AREase activity was most strongly correlated with small/medium HDL-P (r=0.32), followed by total HDL-P (r=0.29), and large HDL-P (r=0.18). PLTP activity was similarly most correlated with small/medium HDL-P (r=0.19), followed by total HDL-P (r=0.18), and with a weak

association with large HDL-P (r=0.08). Total HDL-P was strongly correlated with small/medium HDL-P (r=0.83), but not with large HDL-P (r=0.09).

Figure 3.1: Mean covariate-adjusted common cIMT for Small/Medium HDL-P (A) and Large HDL-P (B) before and after adjustment for HDL-C and the other HDL-P variable. See Table 3.2 for list of covariates.

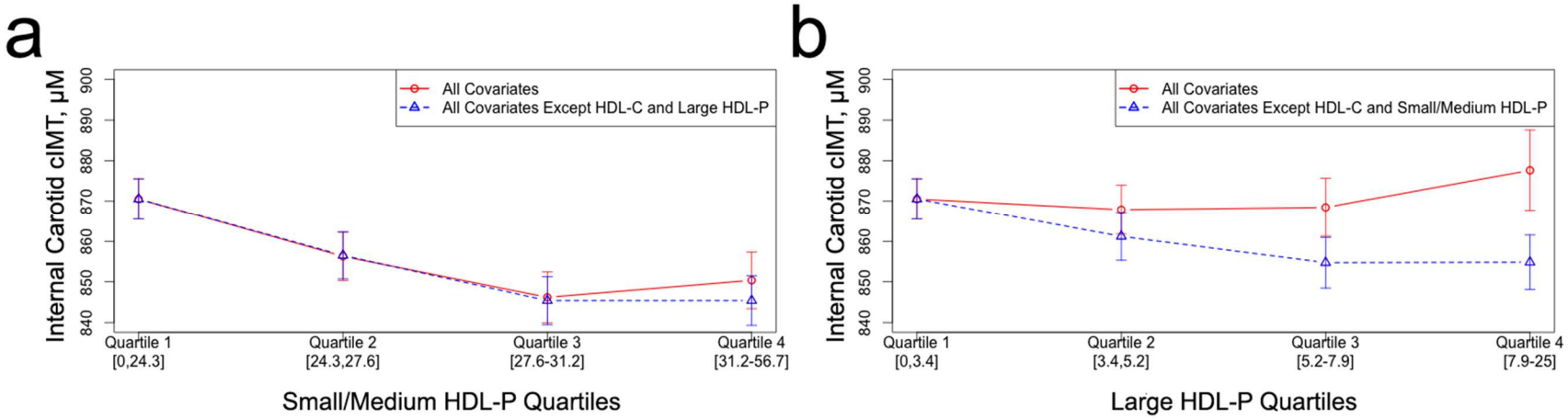


Figure 3.2: Mean covariate-adjusted internal cIMT for Small/Medium HDL-P (A) and Large HDL-P (B) before and after adjustment for HDL-C and the other HDL-P variable. See Table 3.3 for list of covariates.

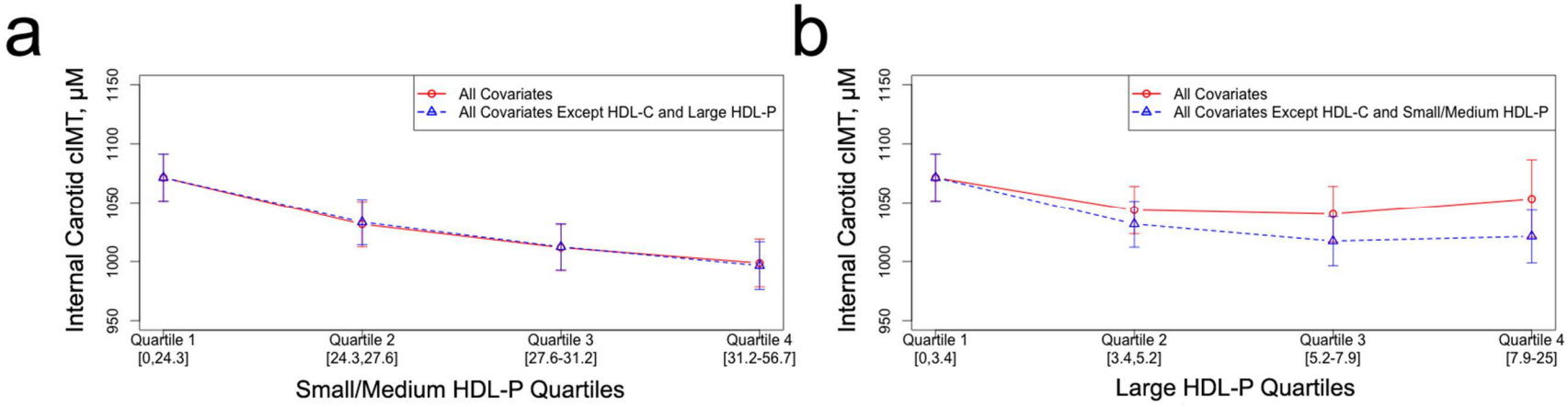


Figure 3.3: Small/medium HDL-P concentration is strongly correlated with Total HDL-P, while Large HDL-P is associated with HDL-C in MESA data (n=6786).

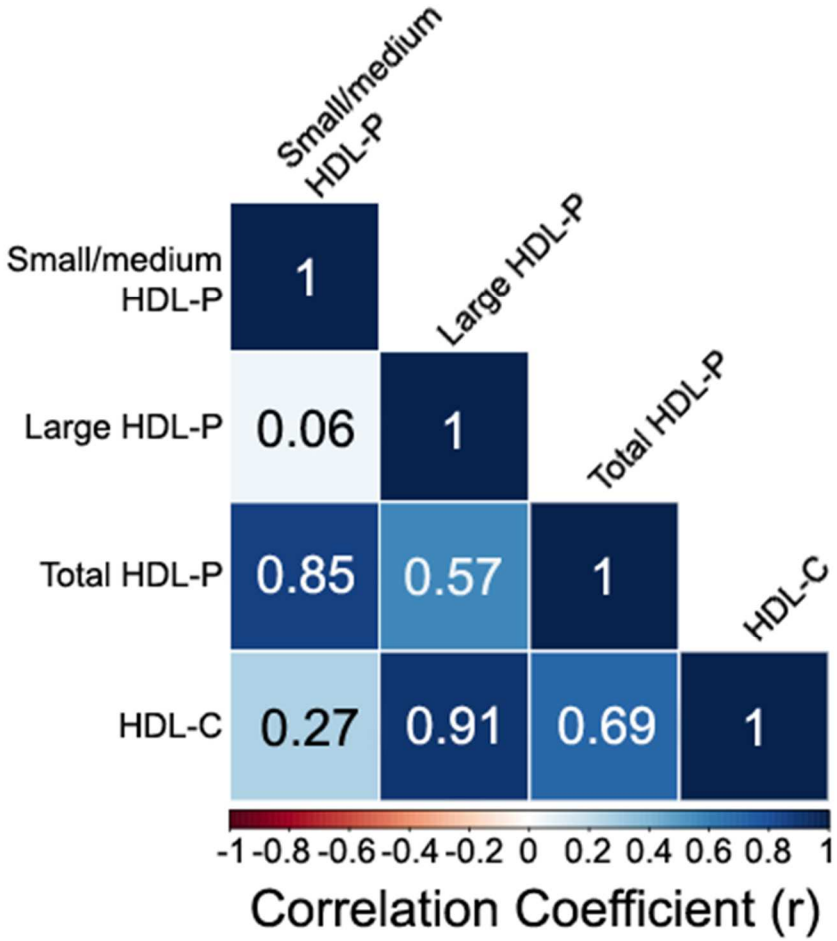
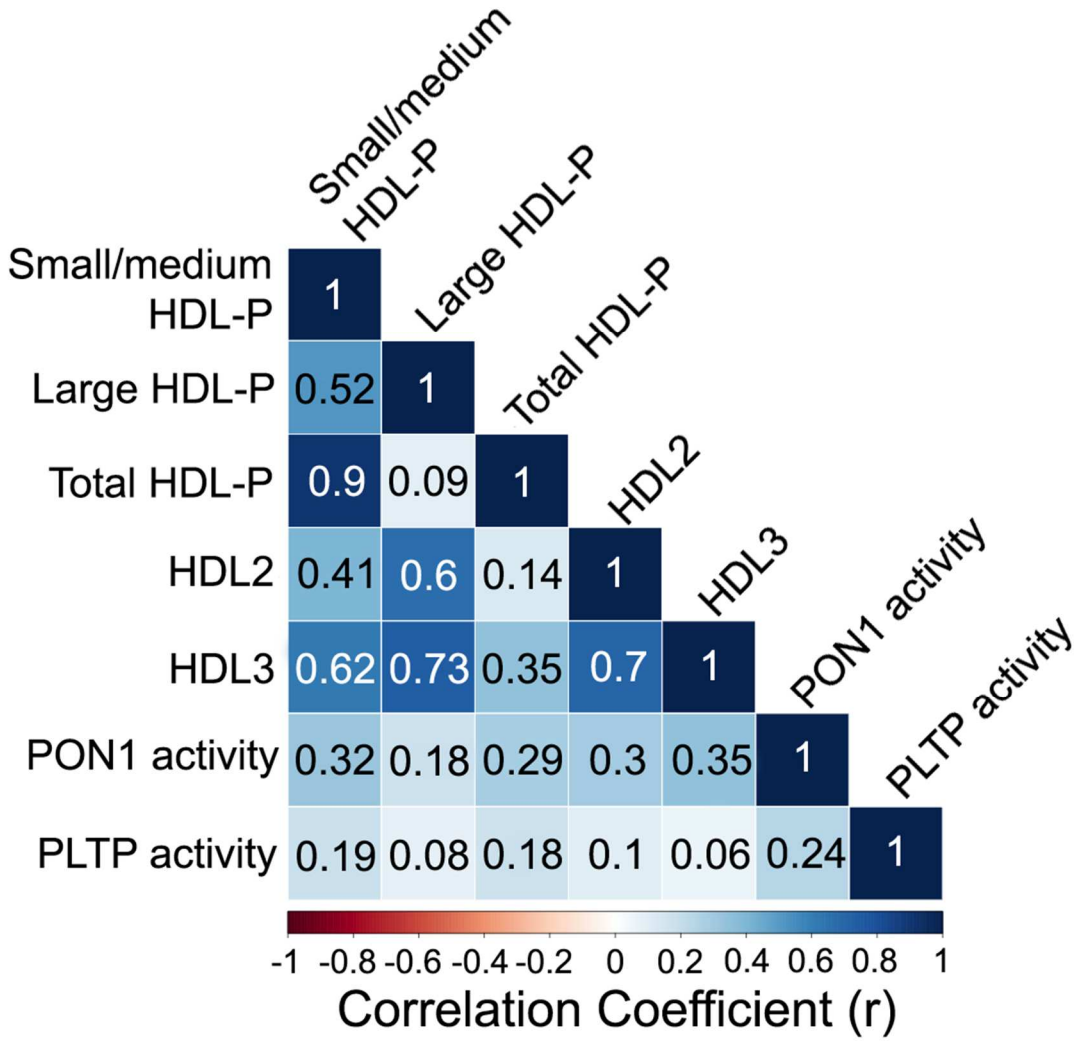


Figure 3.4: PON1 Arylesterase hydrolysis activity (AREase) and PLTP activity are associated with Small/Medium HDL-P concentration, but not Large HDL-P concentration in the Carotid Lesion Epidemiology And Risk (CLEAR) study (n=64).



3.5 DISCUSSION

Functional aspects of HDL not reflected by HDL-C measurements may be responsible for the repeated observation that HDL is protective against cardiovascular disease[46,86-88]. In the current work we present further evidence that proteins carried by HDL—rather than the cholesterol load—are more likely to be protective against cerebrovascular disease. Specifically, we have used linear regression in a large cross-sectional analysis of approximately 6500 MESA participants to demonstrate that cholesterol-poor small/medium HDL-P concentration is significantly and inversely associated with common and internal cIMT, while large HDL-P was not. Finally, using 64 CLEAR participants, we report that PON1 AREase activity and PLTP activity are more strongly associated with small/medium HDL-P than large HDL-P concentration.

Large HDL-P is strongly associated with HDL2, the cholesterol-rich subfraction of HDL[19,89]. In these data the association of large HDL-P with both common and internal cIMT is attenuated by *post-hoc,* stepwise addition of HDL-C to the regression model. We note that large HDL-P is strongly correlated with HDL-C in our data (pairwise correlation coefficient of 0.91, see **Figure 3.3**), which likely explains the attenuation of the large HDL-P association with both cIMT measures after adjustment for HDL-C. These findings add to prior evidence that cholesterol carried by HDL is likely not cardioprotective[42,43,47,48]. In contrast, the association of small/medium HDL-P was robust to *post-hoc* stepwise addition of HDL-C to the regression model, and also the adjustment of large HDL-P. In addition, small/medium HDL-P was only weakly correlated with HDL-C (r=0.27, see **Figure 3.3**). These findings suggest that unmeasured

functional aspects of HDL, better captured by small/medium HDL-P than large HDL-P and not reflected by cholesterol content of HDL, are the more likely source of cardioprotection.

PON1 is a glycoprotein enzyme produced in the liver with extremely broad substrate specificity, including toxic organophosphate pesticides[68,90], statin drug adducts[91], fluoroquinolones[92], and the quorum sensing factor of *Pseudomonas Aeruginosa*[93] (for further in-depth discussion of PON1, please refer to a recent review[70]). Rare mutations in the *PON1* gene have previously been reported to be predictive of ischemic stroke[50]. Additionally, PON1 prevents the oxidation of LDL[10,25,26] and HDL[69]. In human population studies, PON1 enzyme activity has consistently been associated with decreased prevalence[28-30] and incidence[31] of CVD. Notably, PON1 is strongly associated with the smaller/denser HDL particles[23]. Thus, in this context our findings that small/medium HDL-P is strongly associated with decreased common and internal cIMT may, at least in part, reflect the increased cardioprotective enzyme activity of PON1 in these particles[23] as compared to large HDL-P.

Reverse cholesterol transport (RCT), whereby cholesterol is removed from peripheral cells and atherosclerotic plaque by HDL-borne proteins such as phospholipid transfer protein (PLTP) [35,39], represents another possible function of HDL that may underlie its cardioprotective effects. Notably, PLTP is enriched on cholesterol poor HDL and pre-HDL lipoprotein particles[94] – a trend that is similarly reflected in our data, with PLTP activity having highest correlation with small/medium HDL-P as compared to large HDL-P (see **Figure 3.4**).

Sensitivity analyses of the relationship of small/medium HDL-P with common and internal cIMT measures across racial groups have limited power. Notably the direction of the relationship of the HDL measures with both common and internal cIMT is negative across all racial groups and both sex groups, if not always statistically significant. However, we note that Asians (n=804) and women (n=3553) have a near zero association of small/medium HDL and common CIMT. While a formal interaction test was not significant for effect modification by race group (including Asians) of small/medium HDL-P association with either common or internal cIMT, we did note a significant interaction between male gender and small/medium HDL-P for the association with common cIMT. Further work is required to elucidate whether these findings represent true population differences in associations of small/medium HDL-P concentration, or whether they result from random population variance.

Limitations of this study should be considered. First, this study is cross-sectional and observational. Second, although this current work implies that unmeasured functional aspects of HDL are responsible for its cardioprotective nature, it is not currently possible to distinguish whether the effects are due to PON1 activity, RCT and PLTP, both, or possibly other proteins. Similarly, PON1 enzyme activity requires calcium ions for proper function. Therefore, plasma stored with EDTA or other calcium chelating agents cannot be measured for their PON1 activity; consequently, PON1 activity could not be measured in the MESA cohort. Strengths of this study include its multi-ethnic and gender-balanced design, which allow for greater sex and genetic ancestry-specific generalizations regarding the associations of HDL-P subclasses with cIMT outcomes.

In summary, we have performed analyses on participants in the large and well-characterized MESA cohort and report that small/medium HDL-P is strongly and inversely associated with cIMT in both the common and internal carotid arteries. Large HDL-P is not associated with cIMT in either carotid artery branch after adjustment for HDL-C, suggesting that cholesterol carried by HDL is likely not important in prevention of cardiovascular and cerebrovascular disease. Finally, we report from separate data in the CLEAR study that cardioprotective PON1 AREase and PLTP activities are most strongly correlated with small/medium HDL-P, as compared to both total HDL-P and large HDL-P measures. Future work should evaluate the specific functional aspects of HDL, including PON1 activity, PLTP activity, and overall cholesterol efflux capacity of the HDL molecule to identify biomarkers for clinical intervention.

Disclosures: The authors declare that they have no conflicts of interest.

Acknowledgements: This research was supported by contracts N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168 and N01-HC-95169 from the National Heart, Lung, and Blood Institute and by grants UL1-TR-000040 and UL1-TR-001079 from NCRR. The authors thank the other investigators, the staff, and the participants of the MESA study for their valuable contributions. A full list of participating MESA investigators and institutions can be found at http://www.mesa-nhlbi.org. The CLEAR study was supported in part by RO1 HL67406 and a State of Washington Life Sciences Discovery Award (265508) to the Northwest Institute of Genetic Medicine. DSK was supported by NIH 1F31MH101905-01 and T32HL007312.

CHAPTER 4: PLTP ACTIVITY IS PROTECTIVE AGAINST CAROTID ARTERY DISEASE

This chapter is based on the following peer-reviewed manuscript[49]:

Kim DS, Burt AA, Ranchalis JE, Vuletic S, Vaisar T, Li WF, Rosenthal EA, Dong W, Eintracht JF, Motulsky AG, Brunzell JD, Albers JJ, Furlong CE, Jarvik GP. Plasma phospholipid transfer protein (PLTP) activity inversely correlates with carotid artery disease: Effects of paraoxonase 1 enzyme activity and genetic variants on PLTP activity. J Lipid Res. 2015 May 25. pii: jlr.P058032. PMID: 26009633.

4.1 SUMMARY

Purpose: Recent studies have failed to demonstrate a causal cardioprotective effect of high-density lipoprotein cholesterol levels (HDL-C), shifting focus to the functional aspects of HDL. Phospholipid transfer protein is an HDL-associated protein involved in reverse cholesterol transport. This study sought to determine the genetic and non-genetic predictors of plasma PLTP activity (PLTPa), and separately, to determine whether PLTPa predicted carotid artery disease (CAAD).

Methods: PLTPa was measured in 1,115 European ancestry participants from a case-control study of CAAD. A multivariate logistic regression model was used to elucidate the relationship between PLTPa and CAAD. Separately, a stepwise linear regression determined the non-genetic clinical and laboratory characteristics that best predicted PLTPa. A final stepwise regression considering both non-genetic and genetic variables identified the combination of covariates that explained maximal PLTPa variance.

Results: PLTPa was significantly associated with CAAD (7.90×10^{-9}), with a 9% decrease in odds of CAAD per 1 unit increase in PLTPa (OR=0.91). Triglyceride levels (p=0.0042), diabetes ($p=7.28 \times 10^{-5}$), paraoxonase 1 (PON1) activity (p=0.019), statin use (p=0.026), *PLTP* SNP rs4810479 ($p=6.38 \times 10^{-7}$), and *PCIF1* SNP rs181914932 (p=0.041) were all significantly associated with PLTPa.

Conclusions: PLTPa is significantly inversely correlated with CAAD. Furthermore, we report a novel association between PLTPa and PON1 activity, a known predictor of CAAD.

4.2 BACKGROUND

Recent randomized clinical trials and Mendelian Randomization studies have failed to demonstrate a causal role for HDL-C in atherosclerosis, thereby shifting research to the specific proteins carried on the HDL particle and their functions. We have previously reported the association of phospholipid transfer protein (PLTP) activity with *PLTP* locus single nucleotide polymorphisms (SNPs) and non-genetic covariates including HbA1c in a male non-CAAD subset of this present study (n=87) and a 210 person replication dataset[85]. Due to the relatively low throughput of our standardized PLTP activity assay, similar data have not been reported in a larger human dataset, nor has the association of PLTP activity with other plasma lipid traits been evaluated in a large cohort. Thus, the goal of this larger study was three-fold: first, to determine the correlation between fluorescence-based PLTP activity assay measures and our standardized PLTP activity assay; second, to examine the PLTP activity associations with plasma lipid traits, regional SNPs, and other clinical or demographic covariates; and third, to determine the relationship between PLTP activity and CAAD, adjusting for potential confounders.

Phospholipid transfer protein (PLTP) is involved in the transport of cholesterol from the periphery to the liver (reverse cholesterol transport, RCT), which represents another potentially cardioprotective aspect of HDL that is not reflected by HDL-C. PLTP is ubiquitously expressed[33,34] and physically associates with apolipoprotein A1 (apoA1)[32], which facilitates cholesterol and phospholipid transfer from peripheral cells to HDL. Two forms of functionally defined PLTP are found in the plasma[36]: an "active" form that can transfer triglyceride-rich phospholipids from very-low density lipoprotein (VLDL) and chylomicrons to HDL[37,38], and an "inactive" form that lacks this capability. Between 50-90% of total plasma PLTP mass is

reported to be in the "inactive" form[36,95], which is distinguished from active plasma PLTP mass by mean molecular weight (160 to 210 kDa for active PLTP vs. 340 to >600 kDA for inactive PLTP) and Stokes diameter (7.6 to 12.0 nM vs. 12 to >17 nM for active and inactive PLTP, respectively). For the remainder of this chapter and dissertation, "PLTPa" will refer to the active form of PLTP in the plasma.

PLTP facilitates the efflux of cholesterol from peripheral cells to HDL for RCT through various mechanisms: first, PLTP is expressed in macrophages, is upregulated in cholesterol-laden macrophage foam cells, and is present in atherosclerotic plaque[33,34]. On peripheral cells or in atherosclerotic plaque, PLTP promotes the binding of HDL to cholesterol-laden macrophages and fibroblasts and assists in HDL remodeling, which improves cholesterol and phospholipid removal[35]. Recent evidence demonstrated that *in vitro* RNA interference silencing of PLTP decreased apoA1-mediated and HDL3-mediated cholesterol efflux by 67% and 30%, respectively, highlighting the potential role of PLTP in RCT[39].

In contrast to these findings, mouse and other animal models have generally supported a proatherogenic role of PLTP (see review in [96]). Mice deficient in *PLTP* have decreased atherosclerosis when combined with apolipoprotein E (apoE)-deficient or apolipoprotein B (apoB)-transgenic backgrounds[97]. Increased *PLTP* expression in mice has been positively associated with atherosclerosis[98,99] and decreased HDL, low-density lipoprotein (LDL), and VLDL levels[98]. Moreover, increased *PLTP* gene expression has been reported to decrease RCT in mice through measurement of the amount of intraperitoneally injected radioactive-labeled cholesterol that reaches the feces[100]. In humans genetic variants in *PLTP* are associated with

decreased HDL-C and increased triglyceride levels[101]. Separately, a gene score based on two *PLTP* SNPs, rs6065904 and rs378114, predicted lower PLTP activity (PLTPa), decreased hepatic *PLTP* expression, increased HDL3 concentration, and decreased CVD risk in humans[102].

In contrast to these findings that suggest PLTP is atherogenic, it is notable that macrophage *PLTP* gene expression is cardioprotective in mouse models. First, in transgenic *PLTP -/-* and *APOE -/-* mice receiving a bone marrow transplant resulting in *APOE* expression in macrophages only, the lack of *PLTP* gene expression resulted in decreased APOE secretion from macrophages, increased cholesterol buildup, and accelerated atherosclerosis development[103]. Similarly, in transgenic *PLTP -/-* and *LDLR -/-* mice receiving a bone marrow transport resulting in *PLTP* gene expression in macrophages only, the mice expressing *PLTP* had significantly lower LDL cholesterol levels, higher HDL cholesterol levels, and decreased size of atherosclerotic plaque[104]. Finally, even in transgenic *LDLR -/-* mice with normal systemic *PLTP* and *APOAI* gene expression, it was found that PLTP-mediated macrophage atheroprotection could still be derived following bone marrow transplant[105]. Together, these findings suggest that the physiologic context of *PLTP* expression is an important factor in determining its effects on atherosclerosis – with atherosclerotic lesion-based PLTPa likely being atheroprotective and systemic PLTPa possibly being atherogenic[106], though conflicting results have also been reported with macrophage PLTPa not being associated with an increase in RCT[100].

Further investigation on the effects of PLTPa and vascular disease in humans is necessary. Several studies have associated PLTPa with increased risk of coronary heart disease (CHD)[102,107-109]. However, cardioprotective effects of PLTP activity have also been reported for

peripheral artery disease[110]. Given this finding in combination with the report that *PLTP* expression in macrophages is associated with decreased atherosclerosis[104], we hypothesized that PLTPa could potentially be protective against cerebrovascular disease.

4.3 METHODS

Ethics statement

Institutional review boards at the University of Washington, Virginia Mason Medical Center, and Veterans Affairs Puget Sound Health Care approved this study. Written, informed consent was obtained from all participants.

Sample

The study population is a subset (n=1115) of the previously described CLEAR study[28-30,48,53-56,85,111-113] composed of 493 CAAD cases and 622 controls with data on PLTPa, plasma lipid measures, baseline demographic/clinical characteristics, and genetic variants in the *PLTP* region. Of the studied 1115 participants, 462 (41%) were on statin pharmacotherapy; this rate was 72% in CAAD cases and 18% in controls (see **Supplemental Table B1**). Censored age, for analyses with CAAD as the outcome, is defined as the age at CAAD diagnosis for cases and age at enrollment and blood draw for controls. Participants with milder carotid artery stenosis (between 15-49% in one or both internal carotid artery) were not phenotyped for PLTPa and where therefore excluded from these analyses.

Genotyping and Imputation of SNPs

To maintain genetic homogeneity, only European ancestry subjects, as confirmed by use of Illumina HumanCVD BeadChip SNP data[58] and the program STRUCTURE[57], were included in these analyses. Genotyping of 48,742 SNPs relevant to CVD was performed using the Illumina HumanCVD BeadChip[58]. Duplicate genotyping in 34 participants showed 99.7% consistency in genotype calls. Participants were filtered for an individual call rate of < 97%. SNPs were filtered

for a SNP-specific call rate of < 97%.

Imputation of the Illumina HumanCVD BeadChip data (post-quality control measures outlined above) was performed for all autosomes (chromosomes 1-22) using IMPUTE2[114] with phased haplotype data from the 1000 Genomes Project (March 2012 release) as a reference panel[115]. CLEAR participant genotype data was pre-phased prior to imputation using SHAPEIT[116]. Only imputed genotypes with a probability > 0.9 were included. Imputed SNPs with > 5% missing genotypes or whose genotype distributions were not consistent with Hardy-Weinberg equilibrium at the $p < 10^{-6}$ level were filtered out of the genetic dataset. SNPs within 50kb of the *PLTP* gene and with a minor allele frequency > 1% were considered for analyses, a total of 124 SNPs measured in 1,591 European Ancestry participants.

PLTP activity measurement

Radiometric Gold-Standard Liposome Assay

Plasma PLTPa was measured from whole fasting plasma using the standardized radiometric method as previously reported[37,117], blinded to all other participant data, at the Northwest Lipid Metabolism and Diabetes Research Laboratories. Each plasma sample was stored at -70°C and only thawed once prior to assay. In brief, the transfer of ^{14}C-phosphatidylcholine to an HDL acceptor was measured in the presence of 1µl of human plasma. All samples were incubated at 37°C for 15 minutes; the rate of transfer from cultured liposomes to HDL is linear within this period[85,118]. The analysis was performed in three technical replicates, counted four times, from each participant's sample to reduce measurement variation. All assays were performed using the same liposome acceptor preparation to reduce assay variability. Quality control (QC) values

varied on average by 4-8%; assays with QC values exceeding 10% variability were repeated. Assay validation was also established by repeating randomly chosen samples in another, separate assay, from which 3-5% intra-assay variability in assay yields for triplicate samples was calculated. Similar to the prior step, samples with greater than 10% variability in the intra-assay validation step were re-assayed. After all samples had been assayed, the values of PLTPa were adjusted for the QC variability to remove inter-assay differences.

Fluorescence-based Liposome Assay

PLTPa was measured in a subset of 20 male CLEAR participants without CAAD and not taking statin pharmacotherapy using two methods: the previously described radiometric assay and secondly, using a commercial fluorometric assay from Roar Biomedical, Inc, previously used in numerous studies of PLTPa and CVD[107-109]. In brief, 3 µl of sample plasma was incubated with kit-provided donor and acceptor particles resulting in the PLTP-mediated transfer of fluorescent phospholipid. Each assay was performed in triplicate and repeated if inter-assay variation exceeded 10%.

Statistical Analysis

Determination of non-genetic and genetic (full model) sources of PLTP activity:

Using R (http://www.r-project.org) with standard regression packages, stepwise linear regression was used to determine the non-genetic covariates that predicted the highest PLTPa variance (n=1056). Non-genetic covariates entering the model were: BMI, diabetes status, statin use, current smoking status, total cholesterol levels, ln(triglycerides), ln(VLDL), apoA1, HDL-C,

HDL2, HDL3, and PON1 AREase activity. Model comparison was performed using Akaike's information criterion (AIC), beginning with a base model considering age and sex.

After identification of the non-genetic covariates and *PLTP* region SNPs that significantly associated with PLTPa, a final stepwise linear regression model considering both genetic and non-genetic covariates was used to determine the regression model that predicted the maximal amount of PLTPa variance (n=995). All non-genetic covariates and *PLTP* region SNPs that improved model prediction via AIC were included in the final stepwise linear regression model, with model comparison beginning with a base model with age and sex. SNP genotypes were coded additively.

PLTP *region SNP associations with PLTP activity*

SNPs with $\geq$ 1% frequency identified in the *PLTP* region between 50kb upstream of the first codon or downstream of the last (n=124 SNPs) were included in our analyses of the genetic predictors of PLTPa. Due to a highly dispersed pattern of linkage disequilibrium (LD) among these SNPs (see **Supplementary Figure B1**), attempts to reduce the number of tests through use of LD patterns to cluster correlated SNPs into a single test [119] were unproductive. PLINK [120] was used to perform linear regression with the 124 identified *PLTP* region SNPs (with SNP genotypes coded additively) on the outcome of PLTPa, adjusted for the effects of age, sex, diabetes status, statin use, ln(triglyceride levels), current smoking status, and PON1 AREase activity. Only SNPs significantly individually associated with adjusted-PLTPa in univariate analyses ($p \leq 0.05$) were included in the final non-genetic and genetic predictors of PLTPa stepwise regression model (see above).

Association of PLTP activity and potential confounders with CAAD status:

Using R with standard regression packages available, we performed multivariate logistic regression to determine whether PLTPa was associated with CAAD case status. In this model, age, sex, and all covariates we previously determined to be predictive of PLTPa variance by decreasing model AIC (ln(triglycerides), diabetes status, PON1 AREase activity, current smoking status, and apoA1 – all of which are potential confounders of CAAD status) were included in addition to PLTPa.

As CAAD is treated with statins, statin use was confounded for CAAD status and could not be directly included in the model predicting CAAD. Therefore, in the model predicting CAAD status, we adjusted PLTPa for the effects of statin use by subtracting the covariate-adjusted difference in PLTPa from to statin use in controls (-1.02, n=601) using a previously described method[55,121,122].

4.4 RESULTS

Demographic, clinical, and lipid variables of the studied subset of the CLEAR study are presented in **Table 4.1**. The studied subset of CLEAR (n=1115) was comprised of two groups: those who had been previously analyzed by Jarvik *et al.* [85] (n=87) and those who were newly phenotyped for PLTPa for this study (n=1028). The previous analysis examined only non-CAAD control males who were not on statins, while the newly PLTPa phenotyped subset included females (15%), statin users (41%), and CAAD cases (44%). In addition, the newly PLTPa phenotyped subset tended to be older (69.3 years vs. 65.4 years), more likely to be diabetic (16% vs. 8%), and more likely to smoke (12% vs. 7%). With regard to lipid covariates, the newly PLTPa phenotyped subset had significantly lower total cholesterol (mean 187 mg/dl vs. 197 mg/dl) and PON1 AREase activity (mean 137 IU vs. 151 IU), but significantly higher levels of ln(triglycerides) (mean 4.9 vs. 4.7) and ln(VLDL) (mean 3.3 mg/dl vs. 3.2 mg/dl). HDL-related measures (apoA1, HDL-C, HDL2, HDL3) were not significantly different between the two subsets of the current data. Finally, the newly PLTPa phenotyped subset of CLEAR had significantly lower mean PLTPa when compared to the previously analyzed subset (13.4 μM/hr vs. 15.3 μM/hr). Descriptive statistics of the current subset of CLEAR are also presented stratified by CAAD status in **Supplemental Table B1.**

As a separate quality-control analysis, we performed additional PLTPa assays on 20 male control subjects not on statin pharmacotherapy using the commercial fluorometric PLTPa assay. We found that the Pearson's pairwise correlation coefficient for this subset of 20 participants was equal to 0.4799 between the gold-standard and low-throughput PLTPa assay[37,117] used in this study and the commercial fluorescence-based PLTPa assay (see **Figure 4.1**).

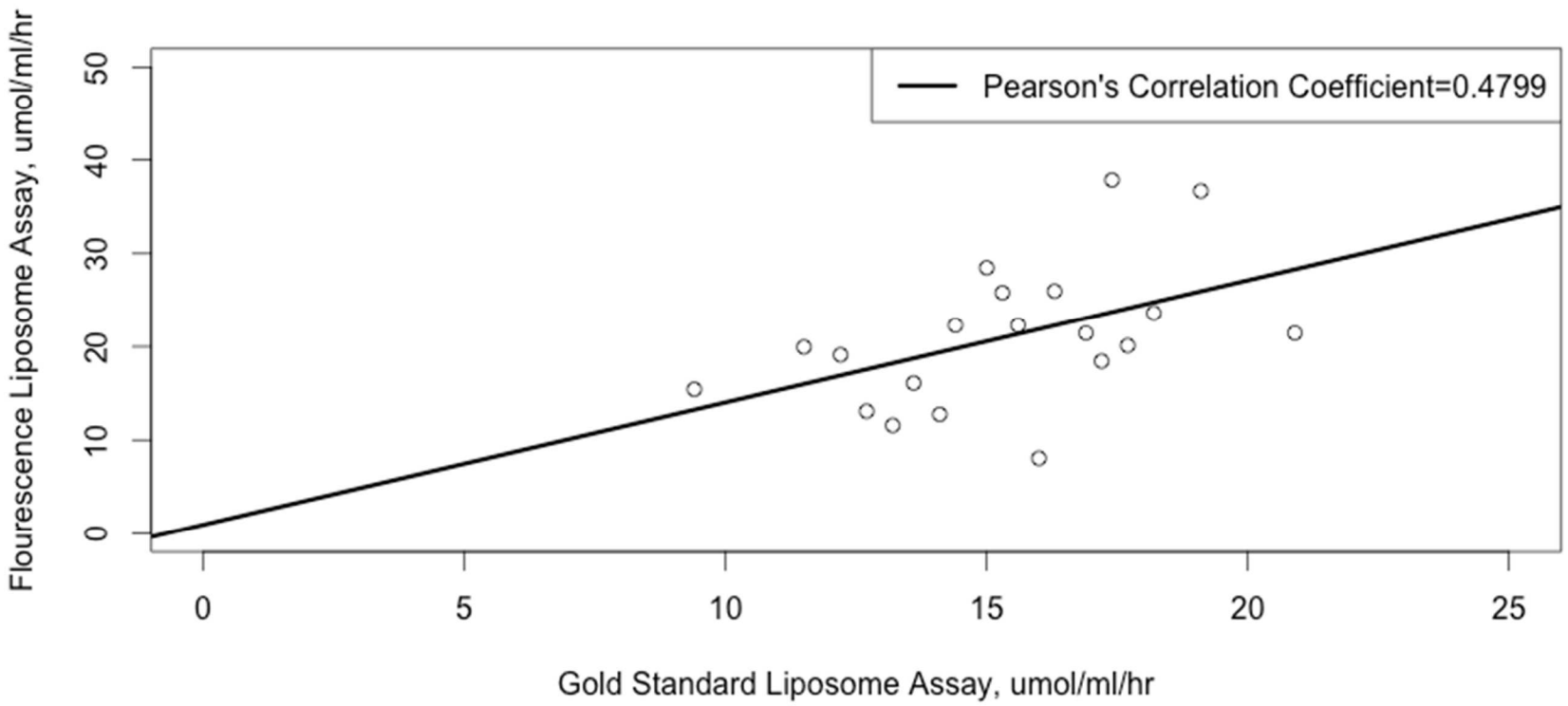
Table 4.1: Demographic and clinical characteristics of the studied CLEAR subset.

Variable	N	Current subset (n=1028)	Prior subset (n=87)	Total (n=1115)	P-Value
Current age, y	1115	69.3 ± 8.6	65.4 ± 9.4	69.0 ± 8.7	< 0.001[a]
Female, %	1115	166 (16%)	0 (0%)	166 (15%)	< 0.001[b]
Diabetic, %	1115	169 (16%)	7 (8%)	176 (16%)	0.03[b]
BMI, kg/m^2	1115	28.1 ± 4.9	28.3 ± 4.4	28.1 ± 4.9	0.88[a]
Statin use, %	1115	462 (45%)	0 (0%)	462 (41%)	< 0.001[b]
Current smoking status, %	1115	123 (12%)	6 (7%)	129 (11%)	< 0.001[b]
CAAD case, %	1115	493 (48%)	0 (0%)	493 (44%)	< 0.001[b]
Total cholesterol, mg/dl	1100	187 ± 41	198 ± 37	187 ± 41	0.004[a]
apoA1, mg/dl	1110	144 ± 28	143 ± 27	144 ± 28	0.81[a]
HDL-C, mg/dl	1100	50 ± 17	51 ± 17	50 ± 17	0.59[a]
HDL-2, mg/dl	1098	9.6 ± 6.3	9.7 ± 6.1	9.6 ± 6.3	0.37[a]
HDL-3, mg/dl	1099	41 ± 11	41 ± 12	41 ± 11	0.75[a]
Ln(Triglycerides)	1100	4.9 ± 0.55	4.7 ± 0.55	4.9 ± 0.54	0.004[a]
Ln(VLDL)	1100	3.3 ± 0.55	3.2 ± 0.54	3.3 ± 0.54	0.003[a]
PON1 AREase activity, IU	1077	137 ± 51	151 ± 46	139 ± 50	< 0.001[a]
PLTPa, μM/hr	1115	13.4 ± 4.5	15.2 ± 2.7	13.5 ± 4.4	< 0.001[a]

Abbreviations: apoA1 = apolipoprotein A1; BMI = body mass index; CAAD = carotid artery disease; CLEAR = Carotid Lesion Epidemiology and Risk cohort; HDL = high-density lipoprotein; PON1 AREase activity = Paraoxonase 1 arylester hydrolysis activity; PLTPa = phospholipid transfer protein activity; VLDL = very-low-density lipoprotein.

Tests used for p-value calculations for differences between subsets: [a] = Wilcoxon rank sum test; [b] = Pearson chi-square test.

Figure 4.1: Correlation between PLTPa measurements from the gold-standard radiometric liposome assay used in the current study and the widely used commercial fluorometric liposome assay. The black line represents the correlation coefficient from a classical linear regression model between the two PLTPa assays.



To determine the predictors of PLTPa in an unbiased method, stepwise linear regression beginning with a base model of age and sex was used in a subset of the cohort (n=1056) with full covariate data. Several lipid-related measures and clinical characteristics were found to improve model prediction of PLTPa variance (by decreasing model AIC, see the "*Determination of non-genetic and genetic (full model) sources of PLTP activity"* subsection of the **Methods***).* Triglyceride levels (β coefficient = 0.88 PLTPa increase per 1 unit increase in ln(triglycerides), p = 0.002) and diabetes status (β coefficient = 1.41 increase in PLTPa if diabetic, p = 0.0003) were positively associated with PLTPa and explained 1.60% and 0.96% of PLTPa variance, respectively. PON1 AREase activity was also associated with an increase in PLTPa (β coefficient = 0.0082 increase in PLTPa per one IU increase in PON1 AREase activity, p = 0.01) and explained an additional 0.98% of PLTPa variance. Finally, current smoking status, apoA1, and statin use all met the AIC threshold to remain in the final regression model for PLTPa, but were not significantly associated (p > 0.05) with PLTPa. These results are summarized in **Table 4.2**.

Table 4.2: Predictors of PTLPa from multivariate stepwise linear regression (N=1056)*.

Variable	Coefficient ± SE	% PLTPa	P-Value
Intercept	*7.34 ± 2.32*	-	-
Current age	-0.014 ± 0.018	0.59%	0.44
Female gender	0.32 ± 0.44	0.00043%	0.47
Ln(Triglycerides)	0.88 ± 0.28	1.60%	0.0017
Diabetes status	1.41 ± 0.38	0.96%	0.0003
Statin use	-0.85 ± 0.30	0.29%	0.005
PON1 AREase activity	0.0082 ± 0.0032	0.98%	0.010
Current smoking status	0.58 ± 0.44	0.19%	0.19
apoA1	0.0090 ± 0.0062	0.16%	0.14

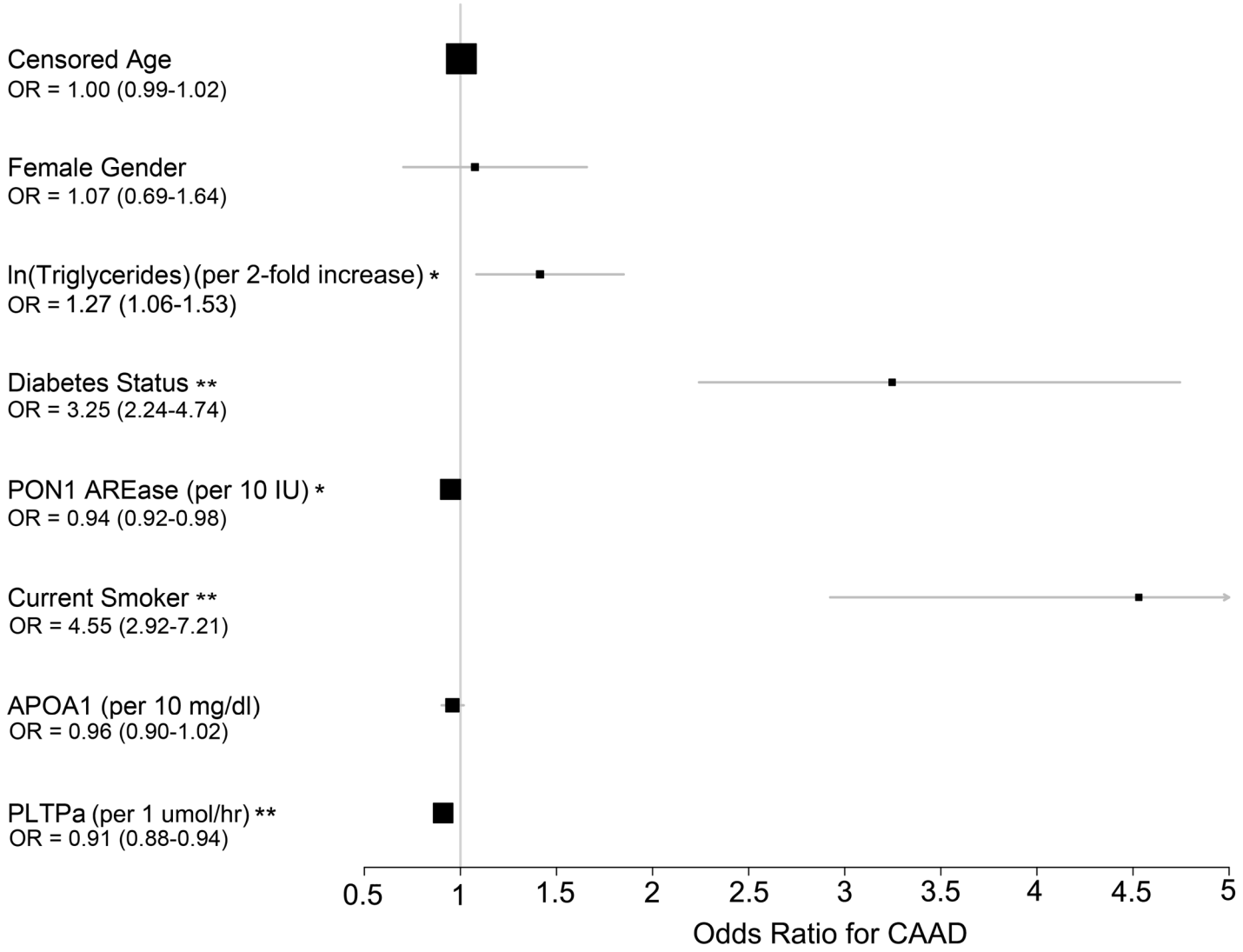
Abbreviations: apoA1 = apolipoprotein A1; PON1 AREase activity = paraoxonase 1 arylester hydrolysis enzymatic activity; PLTPa = phospholipid transfer protein activity.

*1056 participants had complete data on all parameters considered in this stepwise linear regression model. All covariates included in the final model improved model prediction of PLTP activity as measured by AIC. Other covariates considered but not retained in the final model were: BMI, total cholesterol levels, ln(VLDL), HDL-C, HDL2, and HDL3. Please see Methods subsection on "*Determination of non-genetic and genetic (full model) sources of PLTP activity" for further information.*

Sensitivity analysis evaluating if the PLTPa relationship with other factors differed for cases vs. controls was performed by analyzing each of these groups separately; controls (n=601, see **Supplemental Table B2**) and CAAD cases (n=455, see **Supplemental Table B3**). With the exception of statin use, all variables significantly associated with PLTPa in **Table 4.2** had the same direction of effect in both CAAD cases and controls. However, statin use was significantly associated with a decrease in PLTPa of 1.02 μM/hr in CAAD controls (p = 0.037, see **Supplemental Table B2**) while it was not associated with a difference in PLTPa in CAAD cases.

To evaluate the association of PLTPa with CAAD, multivariate logistic regression was used to predict CAAD considering PLTPa, censored age, female gender, ln(triglycerides), diabetes status, PON1 AREase activity, current smoking status, and apoA1 levels (see **Supplemental Table B4**) (n=1056 with complete covariate and phenotype information). As discussed in the methods, PLTPa was pre-adjusted for statin use, due to confounding of statin use and CAAD status. From this analysis, statin-adjusted PLTPa was strongly inversely associated with CAAD status (OR = 0.91, $p=7.90 \times 10^{-9}$), with an approximate 9% decrease in odds of CAAD per 1 μM/hr increase in statin-adjusted PLTPa. Notably, of the predictors of PLTPa that were also included in the regression model, ln(triglycerides) (OR=1.27 per two-fold increase in triglyceride levels, p=0.0051), diabetes status (OR=3.25, $p=6.31 \times 10^{-9}$), PON1 AREase activity (OR=0.94 per 10 IU increase in activity, p=0.0011), and current smoking status (OR=4.55, $p=1.38 \times 10^{-12}$) were all significantly associated with CAAD status. These effects sizes are summarized in **Figure 4.2**.

Figure 4.2: Odds ratio association with CAAD status of PLTP activity and predictors from multivariate logistic regression (n=1056). * = 0.05 > *P* > 0.001. ** = *P* < 0.001. See **Supplemental Table B4** for full regression model coefficients.



To identify the genetic predictors of PLTPa while maximizing statistical power, we used linear regression including PLTPa-associated covariates: age, sex, diabetes status, statin use, ln(triglyceride levels), current smoking status, apoA1 levels, and PON1 AREase activity on 124 SNPs within 50kb of the *PLTP* gene on chromosome 20. The analysis included 1025 CLEAR study participants with genetic and covariate data. Of these 124 *PLTP* region SNPs, 31 were nominally associated with covariate-adjusted PLTPa in univariate association testing ($p \leq 0.05$, **Supplemental Table B5**). An additional 16 of these 31 PLTPa-associated SNPs were also nominally predictive of CAAD status in univariate analyses after adjustment for censored age, sex, diabetes status, ln(triglyceride levels), current smoking status, apoA1 levels, and PON1 AREase activity ($p \leq 0.05$, **Supplemental Table B5**). We then analyzed a subset of 9 SNPs previously reported by Jarvik *et al.* to be significantly associated with PLTPa in the subsample of 87 participants (see **Table 4.3**); 5 of the 9 SNPs replicated (significantly associated with PLTPa and same direction of effect) in this larger dataset. The association of these 9 SNPs with CAAD was previously reported[85] and is summarized in **Table 4.3**. Notably, rs6065904 remains predictive of PLTPa in the larger sample and is also predictive of CAAD status.

To determine the optimal combination of non-genetic factors (from **Table 4.2**) and genetic variants (from **Supplemental Table B5**) that best predict PLTPa variance in our data, we used a final stepwise linear regression model in 995 participants with complete genetic and covariate data. In this model, the 31 *PLTP* region SNPs from **Supplemental Table B5** were added to the model and compared using AIC to a base model containing all previously identified predictors of PLTPa (from **Table 4.2**), such that only the SNPs that further improved model prediction of PLTPa variance were retained. From this process, two SNPs that further improved PLTPa were

Table 4.3: Comparison of *PLTP* SNP association with covariate-adjusted PLTPa and in previously reported vs. current data.

SNP	Function	MAF[a]	2010 N	2010 Beta[b]	2010 *P*[b]	2010 CAAD[c]	2015 N[d]	2015 Beta[e]	2015 *P*[e]	Replicate[f]
rs1736493	Intron 4	0.067	87	0.09	0.011	+	1025	0.26	0.53	-
rs553359	Intron 4	0.39	87	-0.05	0.0083	-	1025	-0.18	0.36	-
rs11086985	Intron 6	0.37	87	0.07	0.0011	-	1025	0.23	0.27	-
rs6065904	Intron 8	0.22	87	-0.08	2.55×10^{-5}	+	1025	-1.05	1.35×10^{-5}	+
rs6073952	Intron 11	0.19	87	-0.10	3.39×10^{-6}	-	1025	-0.89	0.00033	+
rs378114	Intron 12	0.27	87	0.10	5.29×10^{-5}	-	1025	0.51	0.024	+
rs11569668	5' of *PLTP*	0.026	87	0.10	0.016	+	1025	0.51	0.39	-
rs4810479	5' of *PLTP*	0.25	87	-0.07	4.93×10^{-4}	-	1025	-1.18	2.63×10^{-7}	+
rs7679	5' of *PLTP*	0.18	87	-0.11	5.39×10^{-7}	-	987	-1.10	2.84×10^{-5}	+

Abbreviations: MAF = minor allele frequency; PLTPa = phospholipid transfer protein activity; SNP = single nucleotide polymorphism.

[a] = minor allele frequency within the entire CLEAR cohort.

[b] = analyses adjusted by age, BMI, and HBA1c levels (only men analyzed).

[c] = Column indicates whether previously reported *PLTP* SNP associations with covariate-adjusted CAAD (n=939) by Jarvik et al[85] are replicated in the current, slightly larger dataset (n=1118). A (+) symbol indicates that the tested SNP has both $P < 0.05$ and consistent direction of effect for the odds ratios in both datasets. We replicate previously reported associations for rs1736493, rs6065904, and rs11569668 and CAAD. A (-) symbol indicates that either the tested *PLTP* SNP was not significantly associated or the direction of effect on the outcome of CAAD was inconsistent between the two datasets.

[d] = Number available for linear regression analysis of *PLTP* region SNP effects on PLTP activity dependent on the number of CLEAR participants with complete parameter data (PLTP activity, covariates, and genotype data).

[e] = analyses adjusted by age, sex, diabetes status, statin use, ln(triglyceride levels), current smoking status and PON1 AREase activity.

[f] = Replication refers to the *PLTP* region SNP association with covariate-adjusted PLTPa. To have positive replication (+) a SNP must both be significantly associated with PLTPa ($P < 0.05$) and have a consistent direction of effect (increase or decrease PLTPa) in both datasets. Negative replication (-) indicates that either the tested *PLTP* region SNP was not significantly associated with PLTPa in the larger dataset or that the direction of effect on the outcome of PLTPa was not consistent in the two datasets.

identified from the list of 31: rs4810479 and rs181914932 (see **Table 4.4**). *PLTP* SNP rs4810479 was associated with a decrease in PLTPa (β coefficient = -1.08, p = 6.38×10^{-7}) per each minor allele, C, and explained an additional 2.40% PLTPa variance. *PCIF1* SNP rs181914932 was positively associated with PLTPa (β coefficient = 1.18, p = 0.041) per each minor allele, C, and explained 0.31% PLTPa variance.

Table 4.4: Genetic and non-genetic predictors of PTLP activity from stepwise linear regression (N=995)*.

Variable	Coefficient ± SE	% PLTPa	P-Value
Intercept	*13.01 ± 1.42*	-	-
Current age	-0.018 ± 0.017	1.10%	0.31
Female gender	-0.38 ± 0.40	0.0019%	0.35
Statin use	-0.80 ± 0.33	0.29%	0.026
Ln(Triglycerides)	0.89 ± 0.26	1.38%	0.0042
Diabetes status	1.55 ± 0.39	0.72%	7.28×10^{-5}
PON1 AREase activity	0.0072 ± 0.0029	1.06%	0.019
Current smoking status	0.76 ± 0.43	0.23%	0.086
apoA1	0.0093 ± 0.0055	0.21%	0.038
rs4810479	-1.08 ± 0.22	2.40%	6.38×10^{-7}
rs181914932	1.18 ± 0.66	0.31%	0.041

Abbreviations: apoA1 = apolipoprotein A1; PON1 AREase activity = paraoxonase 1 arylester hydrolysis enzymatic activity; PLTPa = phospholipid transfer protein activity.

*995 participants had complete data on all parameters considered in this stepwise linear regression model. All covariates included in the final model improved model prediction of PLTP activity as measured by AIC. Other covariates considered but not retained in the final model were: BMI, total cholesterol levels, ln(VLDL), HDL-C, HDL2, HDL3, and genotypes (coded additively) of the remaining 29 *PLTP* region SNPs significantly associated with PLTP activity (see **Supplemental Table S5** for full list). Please see Methods subsection on "*Determination of non-genetic and genetic (full model) sources of PLTP activity"* for further information.

4.5 DISCUSSION

Functional aspects of the HDL particle not captured by HDL-C measures may explain the recent failures of interventions to raise HDL-C to prevent CVD mortality[48,88]. Of the many functional aspects of HDL, one of the most promising is reverse cholesterol transport whereby lipids are removed from peripheral cells and atherosclerotic plaque by HDL-borne proteins, such as PLTP and apoA1, and excreted through the liver[35].

Within this context, we have provided further evidence that reverse cholesterol transport is cardioprotective through our finding that plasma PLTPa is associated with decreased risk of CAAD (OR = 0.91 per 1 μM/hr increase in PLTPa). In addition, we have performed the largest to-date analysis of the non-genetic and genetic predictors of plasma PLTPa. For the non-genetic predictors we confirmed known associations of triglycerides, diabetes, smoking status, apoA1, and statin use while also discovering the novel and strong association of PON1 AREase activity with PLTPa. With regard to the genetic predictors of PLTPa, we have identified numerous associated *PLTP* region SNPs, confirmed prior associations reported by Jarvik et al [85], and also identified *PLTP* SNP rs4810479 and *PCIF1* SNP rs181914932 as the best genetic predictors of PLTPa in these data.

In this work we also report on the relatively poor concordance (Pearson's pairwise correlation coefficient (r) of 0.4799) between the gold-standard radiometric PLTPa assay used in the current study and the commercially available fluorometric PLTPa assay previously used for numerous studies of PLTPa and CVD. This finding of poor correlation in male controls not on statin pharmacotherapy is in contrast with prior reports of high concordance (r = 0.90) between the two

methods[107-109]. Similar to the high correlation between HDL-C and HDL-P in MESA (r=0.69) but the lack of HDL-C to predict cIMT in a multivariate model including HDL-P [47], the relative lack of correlation observed between the two PLTPa assays (see **Figure 4.1**) in the current work may underlie the differences in conclusions drawn from our study and others in the published literature. To begin, in contrast to our findings of an inverse association of PLTPa and CAAD, cross-sectional human epidemiologic CVD studies that have reported increased PLTPa associated with coronary heart disease (CHD)[108] and carotid intima media thickness[123]. Additionally, one prospective study of 1,085 participants with angiographically-documented CHD at baseline has reported a positive association between PLTPa and the composite outcome of incident myocardial infarction and death[109]. A separate prospective study of 2,679 participants without known CVD at baseline from the Framingham Heart Study found that PLTP higher than the median was a risk factor for incident CVD in males only[107].

However, the cross-sectional study of CHD[108] and the prospective studies of CVD outcomes[107,109] both used fluorescence-based assay kits to measure PLTPa rather than the radiometric-based, standardized direct transfer method of assessing PLTPa. As previously discussed, we have found very poor correlation between the standardized method of PLTPa assessment and commercially available fluorescence PLTPa assay (see **Figure 4.1**). Thus, assay differences and the unexplained variation not captured by the fluorescence-based PLTPa measures (r=0.4799) may explain inconsistency in the direction of the PLTPa—CVD association. Additionally, as PLTPa is upregulated in atherosclerosis[33,34], it is possible that reports of positive association between PLTPa and acute CVD[108,123] reflect sampling of subjects already with significant atherosclerotic disease, and therefore, higher PLTPa, which may be a

consequence of the underlying pathophysiological processes or a compensatory change. In addition, it may be that the cardioprotective effects of PLTPa have increased importance in primary prevention (i.e., before development of sizable atherosclerotic plaque) due to its key role in reverse cholesterol transport. Notably, prospective work by Robins *et al.* did find that higher levels of PLTPa were associated with incident CVD; however, the measures of PLTPa were from the fluorometric-based assay and did not specifically assess cerebrovascular disease, instead using the composite outcome (CHD or hemorrhagic stroke or ischemic stroke) of all incident CVD[107]. As a result, we believe that further prospective studies beginning with disease-free participants and measuring PLTPa with the radiometric-based, standardized direct transfer method are needed to firmly establish whether PLTPa is causally protective against cerebrovascular disease (including CAAD) in particular.

In this work we present what is, to the best of our knowledge, the first report of the association of PON1 AREase activity with plasma PLTPa. In our data, PON1 AREase activity and PLTPa were positively associated with each other and were each inversely correlated with increased CAAD risk. In contrast to the inconsistent direction of the PLTPa—CVD associations, PON1 activity has been consistently reported to be cardioprotective[28,29,as reviewed in 70], including in a well-described and large prospective cohort for incident heart disease[31].

PON1 is an antioxidant glycoprotein enzyme physically associated with the HDL particle [14] and is enriched in the smaller HDL3 subfraction[23]. PON1 has previously been reported to prevent both HDL[69] and LDL oxidation[26], leading to theories that LDL-oxidized phospholipids are one of the numerous targets of PON1 enzymatic action[10]. Use of transgenic mice demonstrated a key

role for PON1 in atherosclerosis, as mice lacking *PON1* had accelerated atherosclerosis when fed a high fat diet[68], a process that was further exacerbated in double knock-out *PON1/APOE* mice[124]. More recent evidence has shown that rare coding variation in *PON1* is associated with ischemic stroke in a sample of approximately 5000 human participants[50]. With regard to PLTP, it is known that PON1 protein is one of 24 different proteins that associates with PLTP in plasma[94]. It is not known if PON1 directly interacts with PLTP in this plasma macromolecular complex, but its association with this complex may relate to its known interaction with apoA1 and clusterin, two major components of PLTP-containing plasma complexes[118]. It is notable that PLTP interacts with both HDL2 and HDL3 and is critically involved in the maturation processes of the HDL particle[35]. Furthermore, PLTP has been shown to reduce secretion of the pro-inflammatory markers and mediators in human primary macrophages and differentiated THP1 cells *in vitro* by modulation of the signal transduction pathways[125], and PLTPa was inversely correlated with the extent of tissue damage in a model of chronic inflammation in humans[126]. Given the enrichment of acute phase reactants in PLTP complexes[94], further molecular work could elucidate a unique role for PON1 and PLTP relationship in inflammation.

Triglyceride and apoA1 levels were both positively associated with plasma PLTPa in our cohort. For triglyceride levels, this finding validates numerous prior reports[118,127] that have previously proposed that this is due to the enhanced ability of triglyceride-rich HDL to accept additional phospholipids via the actions of PLTP[128]. PLTP physically associates with apoA1 in large protein complexes[94]; moreover, apoA1 stabilizes recombinant PLTP lipid transfer activity[94]. It should be noted that both PLTP and apoA1 are able to elicit ABCA1-dependent lipid efflux for reverse cholesterol transport, but with different lipoprotein efflux acceptors: PLTP-mediated

ABCA1 lipid efflux requires the presence of mature HDL particles, while apoA1 uses lipid-poor HDL particles[35,129]. Moreover, it is notable that in this process, PLTP binds to ABCA1 and can competitively inhibit apoA1 binding, indicating a shared protein-interaction region[129]. Finally PLTP is ubiquitously expressed[33,34], while apoA1 is largely restricted to the liver and intestines[130], thereby highlighting the importance of plasma PLTPa in reverse cholesterol transport involving macrophage foam cells at peripheral sites with atherosclerotic plaque[131].

In this work we report that 31 total regional SNPs significantly associated with PLTPa when evaluated separately; moreover, 5 of 9 previously reported SNPs[85] replicated in this dataset. Of the 31 significantly associated SNPs, 2 were significantly and additively predictive of plasma PLTPa, when considering other covariates, indicating that there were likely 2 distinct regions independently associated with PLTPa. The first of these SNPs is a 5' *PLTP* SNP, rs4810479. Our report that rs4810479 is negatively associated with plasma PLTPa replicates prior findings in a smaller subset of the current data[85]. Moreover, rs4810479 has previously been associated with decreased mean HDL particle size[132], increased small HDL particle concentration[133], and decreased large low-density lipoprotein (LDL-C) concentrations[133]. As rs4810479 is located 5' of *PLTP* and is not protein coding, it likely represent a regulatory region for PLTP gene; it should be noted that it is not in strong linkage disequilibrium (LD, $r2 < 0.8$) with rs7679, a *PLTP* SNP previously reported to affect *PLTP* expression[134]. However, in our data rs4810479 was the *PLTP* region SNP most strongly associated with plasma PLTPa and once it was in the final regression model, no other *PLTP* SNP (including rs7679) explained additional PLTPa variance. The second unique genetic signal associating with PLTPa was *PCIF1* 5' SNP rs181914932, which was positively associated with PLTPa. We were unable to ascribe this association to LD

with a *PLTP* SNP. Comparatively little is known about *PCIF1* or rs181914932. *PCIF1* encodes a WW-domain interacting protein that has been shown to affect RNA Polymerase II activity[135] and *MODY4* expression[136], a gene that has been implicated in diabetes pathogenesis. Further work is needed to elaborate the role that *PCIF1* plays in PLTPa.

We also report in this work that 16 of the 31 SNPs significantly associated with PLTPa are also predictive of CAAD status (see **Supplemental Table B5**), replicating 3 SNPs from prior report[85] and addressing Mendelian randomization evidence[43] of the potential causal role of PLTPa in atherogenesis. We note that of the two SNPs that were retained in the final model predicting PLTPa (see **Table 4.4**), *PLTP* SNP rs4810479 was only marginally predictive of CAAD (P=0.097), while *PCIF1* SNP rs181914932 was not (P=0.62). Moreover, it is notable that those SNPs that significantly predict both decreased PLTPa and CAAD in our univariate data are associated with decreased CAAD risk. For example, rs6065904 is associated with a decrease in PLTPa of 1.054 uM/hr per minor allele (P =1.35x10^{-5}) and is also associated with decreased CAAD odds (OR=0.74, 0.59-0.94, P=0.012). Such findings are not consistent with PLTPa being cardioprotective and suggest a complex relationship between genotype, PLTPa, and atherosclerotic end-organ damage. As a potential explanation, it is possible that the *PLTP* SNPs associated with CAAD in this work are more strongly correlated with increased atherogenic *PLTP* gene expression that is hepatic or systemic[98,99], resulting in the finding that *PLTP* SNPs are predictive of CAAD. In contrast, our gold standard assay measure of PLTPa may better reflect peripheral tissue and atherosclerotic plaque PLTPa – thereby informing our finding that PLTPa measured by the radiometric assay used in this study is atheroprotective[104,105]. As a whole, such a complex relationship with high correlation between systemic and peripheral *PLTP*

gene expression that is differentially reflected in our PLTPa measurements could result in conflicting findings, which are overall reflective of the immense complexity of PLTP and its role in atherosclerosis. Further work on the tissue-specific gene expression changes in *PLTP* with varying genotype are needed – both in peripheral tissue and atherosclerotic lesions where *PLTP* expression is expected to be protective against atherosclerosis, in addition to measurements of systemic and hepatic elevations of *PLTP* expression that are likely proatherogenic – to firmly elucidate the role of PLTPa and *PLTP* genetic variants on risk of CAAD.

Several limitations of our study should be considered. First, this subset of the CLEAR study was composed entirely of European ancestry subjects, limiting inferences from our data to participants of other races. Second, due to the cross-sectional nature of this study, no inferences can be made on causality, although the association of *PLTP* SNPs with CAAD suggests a causal role. Future follow-up with prospective design, *PLTP* genotype data, and baseline plasma PLTPa measured using the gold standard assay is required to better understand the etiological role that PLTPa has in CVD. Third, we only examined SNPs in the *PLTP* region. Other genetic predictors of PLTPa likely exist. Finally, we could not directly adjust for the effects of statins in our analysis of CAAD, as statin use is confounded with CAAD status. Instead, we adjusted the levels of plasma PLTPa by the mean change in controls on statins (-1.02), a method previously used to adjust for the effects of statins on lipid levels[55,121], leading to a more conservative estimate of the effect of PLTPa on CAAD status.

In summary, we have performed the largest known study of the genetic and non-genetic factors predicting plasma PLTPa. We confirm in our data the associations of triglycerides, diabetes,

smoking status, apoA1, statin use, and numerous *PLTP* region SNPs on PLTPa. We also report the novel and positive association of PON1 AREase activity with PLTPa and secondly, the association of *PCIF1* SNPs. In these data PLTPa was a significant predictor of CAAD, independently of other covariates. Due to the emerging focus on the process of reverse cholesterol transport[9], an increased focus on PLTPa, which we have shown to be cardioprotective in this data, is warranted.

Disclosures: The authors declare no conflicts of interests.

Acknowledgements: We would like to thank all CLEAR participants. This work was funded in part by National Institutes of Health RO1 HL67406 and a State of Washington Life Sciences Discovery Award (265508) to the Northwest Institute of Genetic Medicine. DSK was supported in part by the Benjamin and Margaret Hall Endowed Fellowship in Genome Sciences, a Markey Foundation award, and National Institutes of Health 1F31MH101905-01. Dr. John Brunzell, our coauthor, colleague, and friend, passed away while this paper was under review; we wish to acknowledge all of his many years of extraordinary science and our longstanding valuable collaborations. He will be missed.

CHAPTER 5: RARE VARIATION IN THE *PON1* GENE PREDICTS ISCHEMIC STROKE

This chapter is based on the following peer-reviewed manuscript[50]:

Kim DS, Crosslin DR, Auer PL, Suzuki SM, Marsillach J, Burt AA, Gordon AS, Meschia JF, Nalls MA, Worrall BB, Longstreth WT Jr, Gottesman RF, Furlong CE, Peters U, Rich SS, Nickerson DA, Jarvik GP; on behalf of the NHLBI Exome Sequencing Project. Rare coding variation in paraoxonase-1 is associated with ischemic stroke in the NHLBI Exome Sequencing Project. *J Lipid Res*. 2014 Apr 7;55(6):1173-1178. PubMed PMID: 24711634; PubMed Central PMCID: PMC4031948.

5.1 SUMMARY

Purpose: HDL-associated paraoxonase-1 (PON1) is an enzyme whose activity is associated with cerebrovascular disease. Common *PON1* genetic variants have not been consistently associated with cerebrovascular disease. Rare coding variation that likely alters PON1 enzyme function may be more strongly associated with stroke.

Methods: The NHLBI Exome Sequencing Project (ESP) sequenced the coding regions (exomes) of the genome for heart, lung, and blood-related phenotypes (including ischemic stroke). In this sample of 4,204 unrelated participants, 496 had verified, non-cardioembolic ischemic stroke. After filtering, 28 non-synonymous *PON1* variants were identified.

Results: Analysis with the Sequence Kernel Association Test (SKAT), adjusted for covariates, identified significant associations between *PON1* variants and ischemic stroke ($p=3.01x10^{-3}$). Stratified analyses demonstrated a stronger association of *PON1* variants with ischemic stroke in African ancestry (AA) participants ($p=5.03x10^{-3}$). Ethnic differences in the association between *PON1* variants with stroke could be due to the effects of $PON1_{Val109Ile}$ (overall $p=7.88x10^{-3}$; AA $p=6.52x10^{-4}$), found at higher frequency in AA participants (1.16% vs. 0.02%) and whose protein is less stable than the common allele.

Conclusions: In summary, rare genetic variation in *PON1* was associated with ischemic stroke, with stronger associations identified in those of AA. Increased focus on PON1 enzyme function and its role in cerebrovascular disease is warranted.

5.2 BACKGROUND

Recent results from a large-scale Mendelian randomization study[43] and randomized clinical trial[42] investigating high density lipoprotein (HDL) have raised doubt on the long-held belief that total HDL cholesterol (HDL-C) is cardioprotective. In light of these findings, research has shifted to the individual components of HDL, whose activities are not reflected by usual measures of HDL-C. Paraoxonase 1 (PON1), encoded by the *PON1* gene, is a liver-produced glycoprotein enzyme whose enzyme activity is strongly cardioprotective, particularly for carotid artery disease[28], a risk factor for ischemic stroke.

Numerous single nucleotide variants (SNVs), including rare protein truncating[112] and promoter SNVs[137] that alter gene expression have been described for *PON1*. Three specific *PON1* variants (*PON1*$_{-108C/T}$, *PON1*$_{L55M}$, and *PON1*$_{Q192R}$) have been extensively studied for their strong effects on *PON1* expression, enzyme activity, or both. Despite the strong association between PON1 enzyme activity and cerebrovascular disease, common *PON1* SNVs (minor allele frequency (MAF) greater than 5%) have not been consistently associated with atherosclerotic end-organ damage[29,30]. Moreover, meta-analyses of *PON1*$_{Q192R}$ have found only a weak association with coronary artery disease (CAD), while *PON1*$_{-108C/T}$ and *PON1*$_{L55M}$ have no demonstrated evidence for CAD association[138,139].

Rare coding SNVs are often unique to an individual or family and likely alter protein function, possibly accounting for a greater portion of genetic risk and missing heritability than common SNVs[140]. By focusing on the putative deleterious coding SNVs in *PON1* that result in a change or loss in PON1 enzyme activity, a stronger association between *PON1* variation and

cerebrovascular disease may be revealed. The goal of this study was to determine whether the burden of rare coding variation in the *PON1* gene was associated with ischemic stroke in participants of the NHLBI Exome Sequencing Project (ESP) and to functionally characterize the most strongly associated rare variant with non-cardioembolic ischemic stroke.

5.3 METHODS

Ethics Statement:

Institutional review boards at each individual site involved in the ESP approved the study, and each study participant at each study site provided written, informed consent.

Participants:

The National Heart, Lung and Blood Institute (NHLBI) Exome Sequencing Project (ESP) is a multi-center study to deeply sequence the exomes of individuals with a variety of heart, lung, and blood disorders. The participants in the ESP were ascertained from different studies, defined by three consortia: WHISP (Women's Health Initiative Sequencing Project), HeartGO (Framingham Heart Study, Jackson Heart Study, Multi-Ethnic Study of Atherosclerosis, Atherosclerosis Risk in Communities, Coronary Artery Risk Development in Young Adults, and Cardiovascular Health Study), and LungGO (Genomic Research on Asthma in the African Diaspora, Lung Health Study, Pulmonary Arterial Hypertension population, Acute Lung Injury cohort, and the Cystic Fibrosis cohort). Additional participants with ischemic stroke were identified from two independent studies, Siblings With Ischemic Stroke Study (SWISS) and the Ischemic Stroke Genetics Study (ISGS). Other ischemic stroke participants represented by these analyses were identified within HeartGO and the WHISP. Previous analysis of a large subset (n=2,440) of the ESP found no systematic biases in patterns and characteristics of SNVs attributable to cohort or technical sources of variation[141].

Unrelated ischemic stroke cases (non-cardioembolic small/lacunar and large/atherosclerotic vessel subtypes), diagnosed <65 years of age or with a positive family history of stroke, were

selected for inclusion from HeartGo (n=250) and WHISP (n=250). Additional affected sib-pairs (n=50) with ischemic stroke were selected from SWISS. Subjects with hemorrhagic stroke were excluded from all analyses.

Exome Sequencing and Variant Calling:

Exome sequencing was performed at the University of Washington and the Broad Institute of MIT/Harvard University. Library construction, exome capture, sequencing, and mapping were performed as previously described[141].

SNVs were called using the UMAKE pipeline at the University of Michigan (http://genome.sph.umich.edu/wiki/UMAKE), which allows for all samples to be analyzed simultaneously for both variant calling and filtering processes. In brief, summaries of Burrows-Wheeler Aligner (BWA) Alignment[142] generated at the University of Washington and Broad Institute were stored as BAM files, which were used as input for the UMAKE pipeline. Clipping of overlapping ends in paired reads was performed to avoid PCR artifacts. All reads not confidently mapped (Phred-scaled mapping quality < 20) were removed. SAMtools[143] was used to calculate genotype likelihoods for exome targeted regions and 50 flanking bases while accounting for per-base alignment quality. Variable sites and their allele frequencies were identified using a maximum-likelihood model, implemented in glfMultiples (http://genome.sph.umich.edu/wiki/GlfMultiples). The support vector machine (SVM) machine learning algorithm was used to separate likely true positive and false positive SNVs. SVM collects a series of features related to the quality of each SNV: overall depth, fraction of samples with coverage, fraction of reference bases in heterozygous individuals (allele balance),

correlation of alternative alleles with strand and read position (strand and cycle bias), and inbreeding coefficient for each variant. For training dataset, SNVs that deviated significantly from expected values in three or more categories were flagged as likely false positives, whereas SNVs from the HapMap and 1000 Genomes project data were flagged as likely true positives. After examining this training set, the SVM classifier was used to identify all likely false positive sites, which were then excluded from the dataset. Using this SNV calling pipeline, a total of 1,908,614 SNVs passed the SVM filter, with an overall transversion to transition ratio (Ts/Tv) of 2.84.

Identification of Related Individuals:

In total, 6,823 subjects were collected for the ESP. The software KING[144] was used on 34,945 common (>5% minor allele frequency) and linkage disequilibrium-pruned SNVs to identify kinship (up to a 3rd degree, or first cousin relationship) through pairwise comparisons across all individuals. The KING kinship coefficient threshold used was 0.04419. From this analysis, clusters of related individuals were formed, with the majority of clusters comprising of two subjects. When multiple individuals were related to each other in a single cluster, subjects with the highest genotype missingness were preferentially removed. When clusters had partial relationships (e.g., when A is related to both B and C, but B and C are not related), preference was given to removing the subject that would leave the largest number of samples. Using these criteria, 242 subjects were removed for relatedness.

Single Nucleotide Variant Filtering:

Genetic variants within the *PON1* gene cluster were extracted from variant call format (VCF) files. SNVs were filtered for a minimum read-depth of 8x, 97% overall site call-rate, and a Hardy-Weinberg equilibrium rejection cut-off $P = 10^{-6}$. Only non-synonymous coding SNVs that are predicted to alter protein residues (missense), splicing of mRNA transcripts (splice), or prematurely truncate proteins (nonsense) were included for analyses. After applying these criteria, a total of 28 SNVs remained for *PON1*. Description of these SNVs can be found in **Table 5.1**. Conservation for single-base variants was assessed through genomic evolutionary rate profiling (GERP)[145] scores using SeattleSeq SNP annotation (http://snp.washington.edu/SeattleSeqAnnotation137).

Table 5.1: *PON1* variants present in ESP 6500 data set.

Base*	rsID	Type†	Protein Change	PhastCon	GERP	Alleles	Percent, All of ESP‡	Percent, EA ESP‡	Percent, AA ESP‡
Shared variants between EA/AA subjects									
7:94931624		M	268/356	0.952	3.79	T/C	0.0154	0.0116	0.0227
7:94937419	rs80019660	M	201/356	0.409	4.1	A/G	0.1999	0.2674	0.0681
7:94937446	rs662	M	192/356	0.001	-1.02	C/T	41.2194	28.5	33.9537
7:94940880	rs144390653	M	127/356	0.588	4.88	C/A	0.1	0.1395	0.0227
7:94944679	rs61736513	M	109/356	0	-6.93	T/C	0.3998	0.0233	1.1348
7:94944735	rs72552788	M	90/356	1	5.06	G/A	0.0538	0.0465	0.0681
7:94946084	rs854560	M	55/356	0.995	-1.76	T/A	30.5244	37.3023	17.2946
7:94953733	rs141948033	M	19/356	0.002	2.35	C/T	0.1615	0.2326	0.0227
EA-specific variants									
7:94931521		M	302/356	0.999	4.67	A/G	0.0077	0.0116	0
7:94931583		M	281/356	1	4.67	G/C	0.0077	0.0116	0
7:94937393		M	210/356	0.971	4.16	A/G	0.0077	0.0116	0
7:94937439	rs3917594	S	194/356	1	5.04	T/C	0.0077	0.0116	0
7:94944741		M	88/356	1	3.9	G/A	0.0077	0.0116	0
7:94944768		M	79/356	0.02	4.17	C/G	0.0077	0.0116	0
7:94947635	rs149100710	MNS	49/356	0.992	3.49	T/C	0.0077	0.0116	0
7:94947638	rs144612002	M	48/356	0.969	0.73	C/T	0.0077	0.0116	0
7:94947656	rs138512790	M	42/356	1	4.37	G/A	0.0077	0.0116	0
7:94947661	rs141665531	M	40/356	1	4.37	A/G	0.0077	0.0116	0
7:94953721	rs146211440	M	23/356	0.882	3.03	C/A	0.0077	0.0116	0
7:94953771	rs150657027	M	6/356	0	-4.3	A/G	0.0231	0.0349	0
AA-specific variants									
7:94928294		M	344/356	0.987	4.93	A/C	0.0154	0	0.0454
7:94928305	rs141598837	M	340/356	0.968	3.77	C/T	0.0077	0	0.0227
7:94928336	rs145997673	M	330/356	0.773	4.92	T/C	0.0077	0	0.0227
7:94931560	rs148911901	M	289/356	0.151	-0.39	T/A	0.1076	0	0.3177
7:94940782	rs13306698	M	160/356	1	3.55	C/T	0.0077	0	0.0227
7:94940884	rs148785172	M	126/356	0.472	0.76	T/C	0.0461	0	0.1362
7:94944642	rs147867887	M	121/356	0.999	3.33	A/G	0.0077	0	0.0227
7:94953727	rs148153353	S	21/356	0.003	2.37	A/G	0.0154	0	0.0454

AA = African Ancestry; EA = European Ancestry; ESP6500 = Latest release of the NHLBI Exome Sequencing Project data containing genetic variant information on 6503 participants. GERP = Genomic Evolutionary Rate Profiling.

* = Positional information based on NCBI 37 build.

† = 3' UTR = 3' untranslated region variant; 5' UTR = 5' untranslated region variant; M = missense variant; MNS = missense near splice variant; S = stop mutation gained.

‡ = Percent, All ESP, Percent EA ESP, and Percent AA ESP reflect frequency expressed in percent among 6503 exomes in the ESP dataset of which 4300 samples are of European Ancestry and 2203 samples are of African Ancestry.

Functional Characterization of Individual PON1 Variants

Expression of PON1 variants:

Recombinant PON1 variants were generated using the QuikChange Lightning Multi Site-Directed Mutagenesis kit (Agilent), with plasmids containing the previously cloned *PON1* gene[93]. The pSCodon plasmid system and SE1 expression cells (Eurogentec) were used for expression of PON1 proteins. Transformants with the PON1-containing plasmids were grown in cultures of 5 ml at 37°C overnight and culture was added at 1:100 to a 100 ml LB starter culture. The starter culture was grown overnight at 25°C; the next day, A_{600} was assessed and 2 L of LB were inoculated with a volume of starter culture to attain a starting A_{600} of ≈0.1. The 2 LB culture was grown at 25°C with 1% glycerol, 10 mM $CaCl_2$, trace metals added at 1 ml/L (sodium chloride 5 g/L, zinc sulfate heptahydrate 1 g/L, manganese chloride tetrahydrate 4 g/L, ferric chloride hexahydrate 4.75 g/L, cupric sulfate pentahydrate 0.4 g/L, boric acid 0.575 g/L, sodium molybdate dihydrate 0.5 g/L and 6N sulfuric acid 12.5 ml/L), and 50 µg/ml carbenicillin. Cells were grown to ≈0.6 A_{600} and induced with 0.5 mM IPTG. They were harvested by centrifugation ≈17 hours post-induction and stored at -20°C.

Purification of PON1 variants:

Lysis of PON1 variants: $PON1_{R192}$ and $PON1_{R192\text{-}V109I}$ were purified as previously described[93]. Frozen cells were thawed, then diluted with buffer (20 mM Tris-HCl pH 8.0, 1 mM $CaCl_2$, 0.1% Tween® 20) containing protease inhibitors (EDTA-free Protease Inhibitor Set III [EMD Millipore],1:10,000 dilution) and 2.5 U/ml Benzonase® nuclease (EMD Millipore) at 3 ml buffer/1 g cell pellet. Thawed, resuspended cell pellets were lysed on ice using an Ultrasonic Processor sonicator. The resuspended cells were sonicated at amplitude 30 with 10 s pulses for 1

min, followed by chilling the lysate on ice and another 1 min sonication. Lysate was centrifuged for 15 min at 7000 rpm at 4°C in a Sorvall® RC-5B Refrigerated Superspeed Centrifuge and the supernatant stored at 4°C. The remaining cell pellet was resuspended again in buffer and the sonication-centrifugation protocol was repeated. Supernatants from multiple sonications were pooled. This was repeated four times, with a sixth and final sonication-centrifugation using buffer with 1% Tergitol® non-ionic detergent (Sigma-Aldrich, Type NP-10).

$PON1_{R192}$ purification: $PON1_{R192}$ lysate from sonication (300 ml from 22.5 g cells) was loaded on a 40 ml diethylaminoethyl (DEAE [GE Healthcare Biosciences]) anion exchange column (1.7 cm diameter x 32 cm height), equilibrated with 20 mM Tris-HCl pH 8.0, 1 mM $CaCl_2$, 0.1% Tween® 20 (Sigma-Aldrich) buffer, washed with 5 resin bed volumes (rbv) of the same buffer, and eluted with a 200 rbv gradient of 0-500 mM NaCl in equilibration buffer. Fractions from all columns were assayed for phenyl acetate (arylesterase) activity using 3.26 mM phenyl acetate in a 9 mM Tris-HCl pH 8.0, 0.9 mM $CaCl_2$ buffer. Column fractions (5 µl of each fraction) were assayed using the SpectraMax Plus^{384} Microplate Reader (Molecular Devices) with transparent UV 96-well plates (Greiner Bio-One) for 4 min at 270 nm. Arylesterase-containing (AREase) fractions were pooled and the pool was desalted using Amicon Ultra Centrifugal Filters (30K, EMD Millipore). The PON1 pool was concentrated to ≈2.5 ml in a Beckman J-6B centrifuge at 2500 rpm and was diluted fivefold up to ≈15 ml with 20 mM Tris-HCl pH 8.0, 1 mM $CaCl_2$ buffer, then concentrated again by centrifugation to 2.5 ml. The fivefold dilution was repeated. Two more concentration-dilutions were performed, with the sample only being diluted twofold. The desalted pool was loaded onto a second DEAE column (10 ml, 0.5 x 20 cm) in 20 mM Tris-HCl pH 8.0, 1 mM $CaCl_2$ buffer, washed until the A_{280} reading reached baseline, then eluted with

a 40 rbv 0-500 mM NaCl gradient in equilibration buffer. The pool containing AREase activity was concentrated to 700 µl and loaded on a 75 ml Superdex-200 size exclusion column (1.8 x 45 cm [GE Healthcare]). The column was eluted with 20 mM Tris-HCl pH 8.0, 1 mM $CaCl_2$, 150 mM NaCl buffer. AREase containing fractions were pooled and loaded on a butyl-hydrophobic interaction column (HIC) (GE Healthcare) (3 ml, 0.9 x 3 cm). The HIC column was initially washed with 5 rbv of a no salt buffer (10 mM Tris-HCl pH 8.0, 1 mM $CaCl_2$), then equilibrated with high salt buffer (10 mM Tris-HCl pH 8.0, 1 mM $CaCl_2$, 3 M NaCl). Following sample loading, the column was washed with 5 rbv high salt buffer, and eluted with a 40 rbv reverse salt gradient, 3 M-0 M NaCl in the Tris-calcium buffer. Another elution with no salt buffer followed; protein elution was measured by A_{280} and the elution continued until absorbance returned to the initial levels. The column was finally eluted with 0.1% Tween® 20 in no salt buffer. Fractions were assayed for AREase activity and the Tween® 20-containing active fractions were pooled. The Tween® 20 elution pool was loaded on a final DEAE column (1.5 ml, 0.5 x 3 cm) equilibrated with 20 mM Tris-HCl pH 7.5,1 mM $CaCl_2$ buffer, washed with the same buffer until absorbance reached baseline, and eluted with a 40 rbv 0-750 mM NaCl gradient in equilibration buffer.

<u>$PON1_{R192\text{-}V109I}$ purification:</u> $PON1_{R192\text{-}V109I}$ lysate from sonication (225 ml from 9.6 g cells) was loaded onto a 30 ml DEAE column (1.7 cm diameter x 24 cm height), equilibrated with 20 mM Tris-HCl pH 8.0, 1 mM $CaCl_2$, 0.1% Tween® 20 buffer, washed with 4 rbv of equilibration buffer, and eluted with a 40 rbv gradient of 0-500 mM NaCl in equilibration buffer. AREase-containing fractions were pooled and the pool was desalted using Amicon Ultra Centrifugal Filters as described above. The desalted pool was loaded onto another DEAE resin (10 ml, 0.5 x

20 cm) in 20 mM Tris-HCl pH 8.0, 1 mM $CaCl_2$, 0.1% Tween® 20 buffer, washed with buffer until the A_{280} reached the baseline level, and eluted with a 40 rbv 0-500 mM NaCl gradient in equilibration buffer. The pool was desalted again with Amicon Ultra Centrifugal Filters and loaded onto a 10 ml non-detergent DEAE column (0.5 x 20 cm) equilibrated in 20 mM Tris-HCl pH 7.0, 1 mM $CaCl_2$ buffer, washed with buffer until the A_{280} reading reached zero, and eluted with a 40 rbv gradient of 0-750 mM NaCl in equilibration buffer. AREase containing fractions were pooled and loaded on a butyl-HIC column (3 ml, 0.9 x 3 cm) and purified as described above. The Tween® 20 elution pool was taken to a final DEAE column (1.5 ml, 0.5 x 3 cm) in 20 mM Tris-HCl pH 7.5,1 mM $CaCl_2$ buffer, washed with equilibration buffer until the A_{280} reading returned to initial levels, and eluted with a 40 rbv 0-750 mM NaCl gradient in equilibration buffer.

Activity assays:

AREase and paraoxonase (POase) activities of the purified PON1s were determined as described previously[146] Briefly, 20 ul of sample was assayed using the SpectraMax Plus[384] Microplate Reader (Molecular Devices) with transparent UV 96-well plates (Greiner Bio-One) for 4 min at 270 nm for AREase and transparent 96-well visible plates (Greiner Bio-One) for 4 min at 405 nm for POase, with assay run at 37°C for the POase assay. Kinetic sample absorbance readings were taken every 15 s for both assays. AREase was assayed using freshly prepared 3.26 mM phenyl acetate in a 9 mM Tris-HCl pH 8.0, 0.9 mM $CaCl_2$ buffer and the POase with freshly prepared 1.2 mM paraoxon in pre-warmed 100 mM Tris-HCl pH 8.5, 2 mM $CaCl_2$, and 2 M NaCl buffer. $PON1_{R192}$ was diluted 1:8.75 for the AREase assay and 1:2 for the POase assay. $PON1_{R192\text{-}V109I}$ was diluted 1:90 for the AREase assay and 1:10 for the POase assay. Activities

were converted to Units/ μg (AREase) or Units/mg (POase) using the molar extinction coefficients of 18 $mM^{-1}cm^{-1}$ or 1.310 $mM^{-1}cm^{-1}$ for AREase and POase, respectively.

Mass Spectrometric Peptide Confirmation of $PON1_{R192-V109I}$:

Both $PON1_{R192}$ and $PON1_{R192-V109I}$ were analyzed by mass spectrometry to confirm the V109I mutant PON1. Five μg of each purified PON1 were boiled for 5 min at 90ºC in the presence of 0.1% RapiGest™ (Waters Corporation, Milford, MA), reduced with 5 mM dithiothreitol (DTT) for 30 min at 50ºC and alkylated with 15 mM iodoacetic acid (IAA) for 30 min at RT in the dark. Then, 1 μg of porcine trypsin (Promega, Madison, WI) was added and samples were incubated at 37ºC for 2 h while rotating. Following digestion, RapiGest was hydrolyzed by adding 100 mM hydrochloric acid to each sample and re-incubating them at 37ºC for 45 min. The generated peptides were analyzed with an LTQ-Orbitrap instrument (Thermo Fisher Scientific) using nanoflow chromatography (Waters nanoACQUITY UPLC). About 200 ng of each sample were loaded onto a 20 cm long, 75 μm inner diameter (I.D.) silica column packed in-house with Jupiter 4 μm Proteo 90 Å C12 reverse-phase beads (Phenomenex, Torrance, CA). A 5 μm I.D. tip was pulled using a P-2000 CO_2 laser puller (Sutter Instrument Company, Novato, CA). Peptides were separated using a 70 min gradient of 2-32% acetonitrile in 0.1% formic acid, at 250 nL/min flow. The Sequest algorithm[147] was used for database searches of the spectral files. All cysteines were monitored as carbamidomethyl cysteines (57.021464 Da static modification), and the valines as isoleucines (14.01565 Da variable modification).

PON1 Protein Heat Resistance Assay:

Assays measuring purified PON1 resistance to heat were performed to assess the function of PON1 protein including the amino acid change most associated with ischemic stroke. These were

carried out at 55°C in a GeneMate Mini Dry Bath incubator block. Protein samples were diluted with a Tris-calcium buffer (20 mM Tris-HCl pH 8.0, 1 mM $CaCl_2$, 150 mM NaCl) to obtain nearly equivalent rates of AREase activity, or about 0.711 Units/ml AREase activity. The $PON1_{R192}$ protein was diluted 1:20 and the $PON1_{R192\text{-}V109I}$ sample was diluted 1:140. Total sample (750 µl) was mixed and 125 µl removed prior to heating. Samples were taken at the timepoints indicated and stored on ice. Samples (20 µl) were assayed in triplicate for each timepoint, with 6 replicates assayed for the starting zero min timepoint. Samples were assayed using a SpectraMax $Plus^{384}$ Microplate Reader (Molecular Devices) with transparent UV 96-well plates (Greiner Bio-One) for 4 min at 270 nm, with sample absorbance readings taken every 15 s. AREase activity was assayed with 3.26 mM phenyl acetate in a no salt buffer (9 mM Tris-HCl pH 8.0, 0.9 mM $CaCl_2$). The average activity at the time zero min reading (no heat) was assigned the 100% value and residual AREase activity was expressed as a percentage of the 100% value.

Analyses:

Subject Filtering:

Of the 6,823 participants in the ESP dataset, 4,224 were used for analyses of *PON* gene cluster variation association with ischemic stroke. Exclusion criteria for this specific study included: relatedness up to the 3rd degree (first cousins), sex mismatch, low concordance with prior genotype data, and individual genotype call rate < 90%. As SWISS recruited sibships with ischemic stroke that were then sequenced as part of the NHBLI ESP, only one sibling from each pair was used for analyses (n=49 cases). For ischemic stroke controls, additional phenotype exclusion criteria excluded participants with other cardiovascular or potentially confounding

phenotype (e.g., myocardial infarction, chronic obstructive pulmonary disease, and ventilator use) and cystic fibrosis. Participants who were collected for high levels of cardiac risk factors (high blood pressure, high low-density lipoprotein levels, high body mass index) but who had not had any noted cardiovascular outcomes (e.g., stroke or myocardial infarction) were included as "controls" for the purposes of this study.

Genetic Ancestry:

Genetic ancestry was determined through principal component analysis (PCA). PCA was performed using the SNPRelate R statistical computing package[148]. Prior to inclusion into the correlation matrix, SNVs were selected after LD pruning at r=0.5, and a MAF > 0.03. For the sample of 4,204 ESP participants, genetically determined European ancestry was assigned to all participants with eigenvectors 1 and 2 values less than and greater than four (±2) SD from the medians of eigenvectors 1 and 2 of self-identified European ancestry participants (n=2,414). For genetically determined African ancestry, we identified all participants with values less than and greater than two (±2) SD from the medians of eigenvector 1 and 2 of self-identified African ancestry participants (n=1677). The process of calculating principal component eigenvectors was then repeated within the European and African ancestry groups, to obtain ancestry-specific eigenvectors.

Statistical Analyses:

The optimized Sequence Kernel Association Test (SKAT-O)[149] was used for testing association of SNVs in each of the *PON* genes with ischemic stroke, using an R plugin (http:// r-project.org). SKAT pools variants across loci, thereby addressing the problem of limited statistical power with

rare variants. It then applies score-based variance-component tests to assess association between SNV sets within the *PON* gene and ischemic stroke, while adjusting for potentially confounding covariates in the model. The covariates adjusted for in SKAT analyses of ischemic stroke were age, sex, current smoking status, and the first three PCA eigenvectors to adjust for population stratification. Default settings, including small sample size correction when n < 2000, were used for SKAT analyses. Single variant score test association results were calculated using skatMeta (http://cran.r-project.org/web/packages/skatMeta/index.html) to identify potential single variant associations driving the observed *PON1* association with ischemic stroke.

To determine whether one genetic ancestry group was responsible for the observed association, stratified analyses were performed in AA (n=1,677) and EA (n=2,414) subsets. For these analyses, genetic ancestry specific PCA eigenvectors were calculated considering only those of a certain genetic ancestry group to adjust for potential population substructure. These ancestry-specific PCA eigenvectors were used to adjust for population stratification, in addition to age, sex, and current smoking status.

Permutation Testing/Statistical Significance:

As the NHBLI ESP represents the largest available collection of phenotyped exome sequences, replication of our rare variant results was not possible. Moreover, dividing the existing sample set into discovery and replication groups has been shown to be less powerful than combined analysis; thus, we analyzed all 4,204 subjects together[150]. Phenotype permutation testing iterated 100,000 times was used to determine significance. In brief, ischemic stroke and control phenotypes and covariate data were randomly assigned to each of the 4,204 subjects (or 2,414

and 1,677 for EA and AA specific analyses, respectively) and analyses were repeated to obtain a p-value, using the "bootstrap" command in SKAT-O. This permutation process was repeated 100,000 times to obtain a histogram of p-values from phenotype permutation. Using the resulting permutation p-value histogram, a two-sided p-value is reported. All significant gene associations with stroke in each genetic ancestry subgroup (EA and AA) with a $p \leq 0.05$ were carried forward to permutation testing. Gene associations with a permutation p-value ≤ 0.05 in conjunction with a prior adjusted $p \leq 0.05$ were declared significant. As this was an evaluation of a specific candidate gene (*PON1*) based upon strong *a priori* data, no attempts at identifying associations across the genome or genome-wide corrections to p-values were performed.

5.4 RESULTS

Demographic information of the ESP participants in this analysis is presented in **Table 5.2**. A total of 4,204 participants had phenotype, genotype, covariate information, and passed quality control measures. The average age was 57.5 years, 32.1% of the studied population was male, and 21.1% reported being current smokers. Ischemic stroke cases were older and were comprised of proportionally more females, as WHI was a major contributor of stroke cases. Cases had an average age of 61.9 years and 19.2% were male, compared to 56.8 years and 33.8% male for controls. Rates of smoking were similar between the ischemic stroke case and non-stroke control group (21.2 and 21.1, respectively). Genetic ancestry of the cohort was 57.4% EA, 39.9% AA, and 2.7% other ancestry (including Hispanic, Asian, and Native American ancestry). Participants of EA comprised a larger proportion of stroke cases (82.7%) compared to controls (54.0%).

Using SKAT regression methods adjusting for age, sex, current smoking status, and the first three PCA eigenvectors, *PON1* ($p=1.29x10^{-3}$) was associated with ischemic stroke at nominal levels of statistical significance in pooled analyses (**Table 5.3**). Permutation testing established the significant association of *PON1* with ischemic stroke of $p=3.01x10^{-3}$.

To explore whether an individual ancestral group was responsible for the observed *PON* gene cluster associations, we stratified analyses within AA and EA subgroups (**Table 5.3**). Using ethnic-specific PCA eigenvectors in addition to age, sex, and current smoking status, *PON1* was found to be nominally significant for association with ischemic stroke in the ESP AA-subset ($p=5.73x10^{-4}$), while the EA-subset was only marginally significant (p=0.07). With permutation

Table 5.2: Description of studied subset of the NHLBI ESP (n=4204 participants).

	All Participants (n=4204)	Controls (n=3708)	Stroke cases (n=496)	EA Participants (n=2414)	AA Participants (n=1677)
Age, years	57.48 ± 13.15	56.83 ± 13.16	61.85 ± 11.79	57.65 ± 14.11	56.71 ± 11.49
Males, (%)	1348 (32.1)	1253 (33.8)	95 (19.2)	922 (38.2)	399 (23.8)
Current Smokers, (%)	887 (21.1)	782 (21.1)	105 (21.2)	528 (21.9)	342 (20.4)
EA, (%)	2414 (57.4)	2004 (54.0)	410 (82.7)	-	-
AA, (%)	1677 (39.9)	1600 (43.2)	77 (15.5)	-	-
Other Ancestry, (%)	113 (2.7)	104 (2.8)	9 (1.8)	-	-

Abbreviations: AA – African Ancestry; EA – European Ancestry; ESP = Exome Sequencing Project; NHLBI = National Heart, Lung, and Blood Institute.

Table 5.3: Results from gene-based exomic variant burden testing for association with stroke.

Gene	Total Number of SNVs [a]	EA Polymorphic SNVs [b]	EA SKAT P-Value [c]	AA Polymorphic SNVs [b]	AA SKAT P-Value [c]	Pooled Polymorphic SNVs [b]	Pooled P-Value[c]
PON1	28	19	0.0716	13	5.73×10^{-4} [d]	27	1.29×10^{-3} [e]

Abbreviations: AA – African Ancestry; EA – European Ancestry; ESP6500 – release of NHLBI ESP data containing genetic data from 6503 participants; Polymorphic – with a rare variant occurring at least once in the studied subset; PON – paraoxonase, SKAT – sequence kernel association testing; SNV – single nucleotide variation.

a Total number of coding SNVs within *PON1* in the ESP6500 dataset.

b Only polymorphic SNVs are used in the regression-based SKAT analysis.

c Analyses adjusting for age, sex, current smoking status, and first 3 principal component eigenvectors (ancestry specific eigenvectors for AA/EA analyses to adjust for potential population substructure).

d P-value after 100,000 iteration phenotype permutation for AA participants' *PON1* rare coding variation and stroke is p=0.00503.

e P-value after 100,000 iteration phenotype permutation for all participants' *PON1* rare coding variation and stroke is p=0.00301.

testing of 100,000 iterations, only the association of *PON1* and stroke in the ESP AA-subset remained significant (permutation $p=5.03x10^{-3}$).

$PON1_{Q192R}$ and $PON1_{L55M}$ are known determinants of PON1 enzyme activity and have previously been associated with CVD[28,138]. To investigate whether the associations observed between *PON1* and ischemic stroke were determined by these two functional *PON1* variants, the SKAT analyses was repeated with the two variants removed. The significance of the association with ischemic stroke for all tested groups (EA, AA, pooled) remained largely unchanged and significant (pooled $p = 0.00127$, AA $p = 5.70x10^{-4}$, EA $p = 0.07$), suggesting the two variants were not entirely responsible for our observed significant associations between *PON1* and stroke.

Individual *PON1* SNV associations with ischemic stroke were determined using a SKAT-based regression approach (**Table 5.4**). *PON1* SNV 7:94944679 (Val109Ile missense variant, rs61736513) was positively associated with ischemic stroke in the pooled ($p=7.88x10^{-3}$) and AA-subset ($p=6.52x10^{-4}$) analyses. $PON1_{V109I}$ MAF was higher in AA (1.19%) compared to EA (0.02%) participants. Two other *PON1* SNVs, 7:94937419 (rs80019660, MAF=0.23%) and 7:94953721 (rs146211440, MAF=0.01%), were associated with ischemic stroke in both the pooled and EA-subset analyses (see **Table 5.4**). The *PON1 SNV* rs80019660 is an Ala201Val missense (found more frequently in EA subjects than AA) that is protective against ischemic stroke, while rs146211440 is a Ser23Ala variant found only in EA subjects that confers increased risk.

Table 5.4: *PON1* single nucleotide variants associated with ischemic stroke.

***PON1* Variant**	**rsID**	**Missense Type**	**Pooled MAF**	**EA MAF**	**AA MAF**	**OR (95% CI)**	**SKAT Meta P-Value**
Pooled Analyses [a]							
7:94944679	rs61736513	Val109Ile	0.49%	0.02%	1.16%	**1.14 (1.08-1.20)**	7.88×10^{-3}
7:94953721	rs146211440	Ser23Ala	0.01%	[b]	[b]	2.14 (1.57-2.91)	0.0147
7:94937419	rs80019660	Ala201Val	0.23%	0.35%	0.06%	0.84 (0.79-0.91)	0.0165
7:94947661	rs141665531	Pro40Leu	0.01%	[b]	[b]	2.07 (1.52-2.83)	0.0195
AA-only Analyses [a]							
7:94944679	rs61736513	Val109Ile	1.16%	0.02%	1.16%	**1.13 (1.09-1.16)**	6.52×10^{-4}
EA-only Analyses [a]							
7:94937419	rs80019660	Ala201Val	0.35%	0.35%	0.06%	0.82 (0.75-0.90)	0.0299
7:94953721	rs146211440	Ser23Ala	0.02%	[b]	[b]	2.05 (1.42-2.96)	0.0494

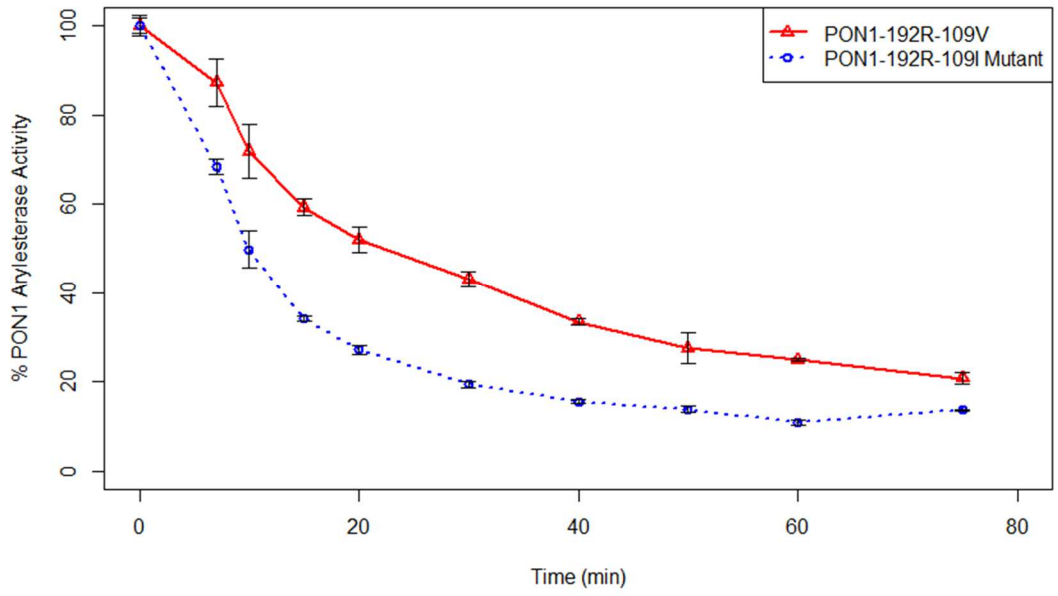
Abbreviations: AA – African Ancestry; CI – confidence interval; EA – European Ancestry; ESP – Exome Sequencing Project; MAF – minor allele frequency; OR – odds ratio; PON – paraoxonase, SKAT – sequence kernel association testing.

a Analyses adjusting for age, sex, current smoking status, and first 3 principal component eigenvectors (ancestry specific for AA/EA analyses).

b *PON1* SNV observed only for EA subjects in the ESP6500 data.

To determine the effects of *PON1* SNV rs61736513 (*PON1*$_{V109I}$) on PON1 enzyme function, several functional tests were performed. First the *PON1*$_{V109I}$ missense variant was inserted into a *PON1* plasmid containing the *PON1*$_{192R}$ variant, forming the *PON1*$_{192R-109I}$ construct. Both *PON1*$_{192R-109I}$ and the *PON1*$_{192R}$ variants were then expressed and purified via liquid chromatography separately. Post-hoc mass spectrometry analysis of the PON1 proteins used in each experiment structurally confirmed both *PON1*$_{192R}$ and *PON1*$_{192R-109I}$ as the PON1 protein variants tested. While, assays of PON1 enzyme function at baseline showed that the two had approximately normal hydrolysis rates of arylester (AREase activity, data not shown), heat resistance assays for *PON1*$_{192R-109I}$ demonstrated significantly lower ($p < 0.05$) residual AREase activity at each successive time point when compared to the nonmutated *PON1*$_{192R}$ protein (see **Figure 5.1**). These data suggest that, under some circumstances, the function of PON1 is affected by this substitution.

Figure 5.1. PON1$_{192R-109I}$ protein variant retains less PON1 arylesterase activity with heat stress compared to PON1$_{192R}$ protein. The x-axis represents time in minutes after the initiation of heat stress. The y-axis represents the % of original PON1 arylesterase (measured at t=0 min) activity retained. % PON1 arylesterase activity retained was significantly different ($p < 0.05$) for all time points comparing PON1$_{192R-109I}$ and PON1$_{192R}$ protein.



5.5 DISCUSSION

In light of the recent evidence that challenges the assertion that HDL-C levels mark the cardioprotective properties of HDL[42,43], a more thorough understanding of *PON1* and, specifically, how deleterious genetic SNVs might alter PON1 enzyme function, may provide new insights as to how HDL and its associated components act in concert to prevent atherosclerotic disease.

Within this context, we have completed the first large-scale study of the effects of rare coding variation in the *PON* gene cluster on the cardiovascular outcome of non-cardioembolic ischemic stroke. Rare coding variation in *PON1*, likely to alter function and be deleterious, is associated with ischemic stroke risk (permutation p-value=3.01×10^{-3}). Moreover, the association between this coding variation in *PON1* and stroke is independent of the common functional *PON1* variants, $PON1_{Q192R}$ and $PON1_{L55M}$. These effects of *PON1* are more pronounced in participants of AA (permutation p=5.03×10^{-3}) compared to participants of EA, which may be attributed to the $PON1_{V109I}$ mutation that is found more frequently in AA subjects. Finally, we have demonstrated that the $PON1_{V109I}$ mutation results in a protein that is functionally compromised.

The finding that *PON1* is more significantly associated with ischemic stroke in participants of AA than EA is interesting, although the finding requires replication. Previous investigations into *PON1* SNVs and cardiovascular and cerebrovascular disease have largely focused on European and Asian cohorts[28,30,138,139]. However, relative to EA patients, those of AA have a higher rate of ischemic stroke in the United States[51], receive fewer evidence-based treatments when in hospital, and thus have a longer length-of-stay relative to white patients[151]. Given these considerations, an

association of *PON1* SNVs with ischemic stroke in patients of AA may have consequences for genetic risk prediction in this high-risk population and could potentially help reduce the high morbidity and mortality of stroke. Moreover, the finding that the $PON1_{V109I}$ protein is less stable under heat stress testing warrants further functional testing within human cells. Although the $PON1_{V109I}$ protein has normal baseline PON1 enzyme activity, it is possible that it more rapidly degraded *in vivo*, thus leading to lower levels of the cardioprotective PON1 protein and an increased risk of ischemic stroke.

Although rare variation could account for a large portion of complex trait inheritance, such as for ischemic stroke, alternative and potentially complementary hypotheses have been proposed. One of these hypotheses is that gene-by-environment interactions among common SNVs comprise a large portion of heritability[55,152]. Given the wide variety of pharmacologic and dietary determinants on *PON1* expression and enzyme activity[53], the potential interaction of these environmental factors with *PON1* variants could represent another important source of trait heritability.

Some limitations of this study should be considered. First, although the ESP data contained two coding *PON1* functional SNVs ($PON1_{Q192}$ and $PON1_{L55M}$), $PON1_{-108C/T}$ was not captured by the exome sequencing methods. $PON1_{-108C/T}$ is a major determinant of PON1 activity, accounting for approximately 14% of PON1 activity variance[30,53]. However, as $PON1_{-108C/T}$ has not been associated with heart disease in meta-analyses[138,139] or carotid artery disease in smaller cohorts[28-30], and neither of the other *PON1* functional SNVs affected results, it may not have accounted for increased risk of ischemic stroke in this study. Second, participants of African Ancestry

represented only a small portion of total ischemic stroke cases in this data (77 of 496 total cases). As replication data were not available, we permuted the phenotype 100,000 times and obtained a permutation p-value that remained significant (permutation $p=5.03x10^{-3}$) and suggestive of a true positive result. Separate replication using exome or whole genome sequence data that captures rare coding variation is needed to verify our result. Third, the cohort was comprised primarily of females for both the ischemic stroke cases and controls; this limits generalizability of our findings. Fourth, our definition of controls for ischemic stroke in this study included subjects with high cardiac risk factors, but no cardiovascular events. However, when we performed a smaller and more restrictive analysis using only subjects collected as "controls" or "deeply phenotyped resources" we found that the association between *PON1* and ischemic stroke remained significant. We therefore believe that our definition of controls for ischemic stroke were valid, and may have more accurately represented the broader population. Finally, studies of paraoxonase and CVD would optimally include measures of paraoxonase activity. Unfortunately, PON1 activity assays could not be completed for the purposes of this study. Most sites used specimens derived from their stored plasma in tubes containing ethylenediaminetetraacetic acid (EDTA); however, EDTA irreversibly inactivates PON1 by chelation of calcium. This also limited a potential source of functional validation of our findings through testing of participant plasma for PON1 enzyme activity.

In conclusion, we present the first known application of exome sequence data to the *PON* gene cluster and describe the strong association between rare coding variation in *PON1* and non-cardioembolic ischemic stroke in 4,404 participants. We also present evidence that participants of AA have a stronger association between *PON1* variation and stroke risk than those of EA, and

that the activity of the PON1$_{V109I}$ protein variant found almost exclusively in participants of AA is less stable compared to the common allele. These results strengthen the link between PON1 and CVD by demonstrating that rare coding variation, which is likely to change PON1 protein function, is associated with non-cardioembolic ischemic stroke where common variant studies in the past have failed to find an association.

Disclosures: The authors declare that they have no conflicts of interest.

Acknowledgements: The authors wish to acknowledge the support of the National Heart, Lung, and Blood Institute (NHLBI) and the contributions of the research institutions, study investigators, field staff and study participants in creating this resource for biomedical research. This work was funded in part by WA State Life Sciences Discovery Fund (265508, to the Northwest Institute of Genetic Medicine). The NHLBI funded the multicenter effort with the following grants: the Lung GO Sequencing Project (HL-102923), the WHI Sequencing Project (HL-102924), the Broad GO Sequencing Project (HL-102925), the Seattle GO Sequencing Project (HL-102926) and the Heart GO Sequencing Project (HL-103010). SSM was supported in part by NIH Grants ES09883 and T32ES007032. MAN was supported in part by the Intramural Research Program of the National Institute of Aging, NIH project Z01 AG-000954-06. DSK was supported in part by the Benjamin and Margaret Hall Endowed Fellowship in Genome Sciences, a Markey Foundation award, and National Institutes of Health 5T31HG000035-18 and 1F31MH101905-01.

CHAPTER 6: CONCLUSIONS AND FUTURE DIRECTIONS

6.1 Summary of dissertation work

Since recent studies have disputed the role of HDL-C in the causal pathway of atherosclerotic disease, focus has shifted to other aspects of the HDL particle that may instead be responsible for its long-held association with decreased CVD incidence[88]. In this dissertation, I have made small contributions in this field by providing epidemiologic evidence that small and dense HDL-P is a stronger predictor of cerebrovascular disease; first, by identifying that HDL3 accounts for greater CAAD status variance in the CLEAR study[48], and then by following up in the MESA study to report that small/medium HDL-P is robustly associated with decreased common and internal cIMT (Kim *et al*, in review). I have also focused on two specific proteins that are physically associated with the smaller and denser HDL particles[14,23]: PLTP and PON1. In this work I have elucidated the cardioprotective role of PLTP activity, as measured by the gold-standard radiometric assay, and CAAD status in the CLEAR study[49]. I have also demonstrated Mendelian Randomization for *PON1* rare variants by demonstrating that the $PON1_{V109I}$ variant is associated with decreased PON1 activity and increased ischemic stroke risk[50]. For the remainder of this chapter, I will attempt to highlight research needs in the field of HDL research, and predict where the field is headed in the future.

6.2 HDL-P as a clinical correlate

HDL-P is currently being investigated as a potential biomarker for use in the clinic[153], after the robust findings in MESA[154] and separately in the CLEAR study[78]. Within this context, I have reported that HDL3 (as measured by standard cholesterol affinity assays in the CLEAR study) and small/medium HDL-P (as measured by NMR in the MESA data) are superior predictors of

CAAD[48] and cIMT (unpublished data), respectively, as compared to HDL-C and large HDL-P measures. Separately, our work in the CLEAR study that found that small/medium HDL-P as measured by ion mobility is similarly protective against CAAD[78]. These data do narrow the focus to the small and dense HDL particles as potential sources of cardioprotection, in contrast with the long-held belief that cholesterol carried on HDL is itself protective.

Despite these results, there are currently disputes over the methodology used in determining HDL-P: nuclear magnetic resonance (NMR) vs. calibrated ion mobility (which is based on mass spectrometry) that must first be resolved before clinical implementation can begin. As noted in separate work, the NMR method identifies a minority of large HDL-P, which is in conflict with standard methodologies, such as gel electrophoresis and ultracentrifugation.[78]. Moreover, the NMR and ion mobility methods both give divergent estimates of mean HDL-P (30 μmol/L vs 5 μmol/L, respectively) and neither method yields estimates of apoA1 that are similar to previously reported literature[78]. These data suggest that although HDL-P as measured by either NMR or ion mobility is a superior predictor of CVD as compared to HDL-C, further work is required to refine these technologies before clinical implementation is prudent.

6.3 PON1: arylesterase activity and rare variation

PON1 activity has long been reported to be the superior predictor of both CVD and cerebrovascular disease as compared to the common SNPs (*PON1*$_{Q192R}$, *PON1*$_{L55M}$, *PON1*$_{-108C/T}$, *PON1*$_{-162G/A}$) that account for a majority of its organophosphate hydrolysis activity (paraoxonase and diazoxonase). In part due to the dependence of these organophosphate activities on common genotypes, PON1 arylesterase (AREase) has long been held as the ideal measure of PON1

activity[30,70]. PON1 AREase, in addition to being robust to the common *PON1* SNPs, is also a better correlate of PON1 protein levels[61]. PON1 AREase also better predicts CAAD as compared to both paraoxonase and diazoxonase (unpublished data).

Despite this knowledge of the superiority of PON1 AREase activity for evaluating the association between PON1 and CVD outcomes, the vast majority of epidemiologic studies of PON1 have used the common *PON1* genotypes as proxies for association studies. As a result, numerous studies have found only a slight increase in CVD risk for the $PON1_{Q192R}$ SNP, but not association for the other three common *PON1* SNPs[138]. Thus, at the time of my work in *PON1* rare genetics, many researchers had disregarded PON1 as a potential biomarker in CVD.

The demonstration of Mendelian Randomization for rare variants in *PON1* (specifically, $PON1_{V109I}$) for association with ischemic stroke will, hopefully, revive interest in PON1 as a potential intermediate phenotype in the causal pathway of cerebrovascular disease. In such studies, however, I advocate for use of both PON1 AREase activity and *PON1* sequence data for optimal epidemiologic determination of the association of PON1 and CVD outcomes (a revised "PON status", as discussed in[70]). Moreover, I note that replication and validation of my results must be sought in independent data; at the time of the study, the ESP represented the largest collection of phenotyped exomic data, and thus replication of the $PON1_{V109I}$ association with CVD outcomes could not be completed (the $PON1_{V109I}$ variant was not included on the Illumina HumanBead ExomeChip)[50].

6.4 PLTP: gold-standard activity measures and conflicting genetics

The majority of PLTP activity association studies preceding my own used the high-throughput commercial assay from Roar Biomedical that uses fluorescence as the measure of PLTP activity. In my work, I demonstrate that this is not a strong reflection of PLTP activity, as measured by the low-throughput, but gold-standard, radiometric assay[49]. I then demonstrated that radiometrically-measured PLTP activity was strongly associated with decreased CAAD risk[49], in contrast with numerous studies using the commercial PLTP assay[107-109,123].

Despite these findings, conflicting data also presented itself in the CLEAR study when considering the genetics of PLTP activity. In this work, I also investigated *PLTP* region SNPs and their association with both PLTP activity and CAAD status. Interestingly, I found that *PLTP* SNPs that were associated with increased PLTP activity were associated with increased CAAD risk, which would indicate that PLTP is pro-atherogenic[49]. These results portray a complex relationship between PLTP and atherosclerotic disease. This complexity is exemplified in its mouse studies, whereby *PLTP* KO mice have decreased atherosclerotic disease, but macrophage-specific *PLTP* activity (through bone marrow transplant) is atheroprotective. We have hypothesized that the *PLTP* SNPs we reported on were more strongly associated with pro-atherogenic systemic or hepatic PLTP activity, which may comprise the majority of the total plasma PLTP activity that we measured.

As a result of this intricacy, I believe that, while challenging, the development of tissue-specific assays of PLTP activity are imperative to determine which subtypes of PLTP are pro- vs. anti-atherogenic. Similarly, tissue-specific *PLTP* expression experiments can be performed to

correlate with *PLTP* region SNPs, to determine if there are any differences in association between the tissues and cell types studied. Until the time of such assay development, however, I also would argue that additional studies using the gold-standard radiometric PLTP assay should be conducted to replicate the CAAD-protective finding that I have reported[49].

6.5 Longitudinal studies to establish clinical evidence

All of the work that I have performed in this dissertation represents cross-sectional or case-control studies. While the results may be compelling, epidemiologic rigor requires that longitudinal studies be performed to provide strong scientific evidence of possible causality (which would then ideally be followed up in a randomized clinical trial with an intervention that affects the intermediate phenotypes studied: HDL-P, PLTP, or PON1). While such studies are underway or can be easily completed with existing plasma for HDL-P[155,156], no such studies are currently investigating PLTP activity (measured either by the fluorescent commercial assay or the gold-standard radiometric assay) for association with incident CVD. Moreover, the studies of HDL-P and incident disease continue to use divergent methods: some with ion mobility[156] and others with NMR[154,155], which makes interpretation of the effects of HDL-P with CVD outcomes more difficult. Finally, it is notable that such studies are not possible for PON1, as the vast majority of longitudinal studies only used EDTA plasma, which irreversibly inactivates PON1 enzyme activity through calcium chelation[61,146]. Thus, determination of the longitudinal association of PON1 activity with CVD incidence will likely not be possible without the collection of an entirely new cohort of study.

6.6 Conclusions

In summary, through this dissertation I have presented evidence that small and dense HDL-P particles (and possibly their associated proteins, PLTP and PON1) are associated with decreased cerebrovascular disease. However, none of the data presented in my dissertation offer strong longitudinal evidence on causality (though the *PON1* rare variation work does show a form of Mendelian Randomization); thus, more work needs to be done to further elucidate the role of these HDL-associated intermediate phenotypes on incident CVD. Through further research and validation, it is my hope that researchers can strongly propose a role for specific functions of HDL in the causal pathway for atherosclerosis, thereby reviving interest in this multi-faceted and ubiquitous particle.

APPENDIX A: SUPPLEMENTAL TABLES FOR CHAPTER 3

Supplemental Table A1. Full multivariable regression model coefficients for association of small/medium HDL-P, large HDL-P, and common/internal cIMT.

	Common cIMT *N*=6512		Internal cIMT *N*=6418	
Covariates	**Beta ± SE**	***P***	**Beta ± SE**	***P***
Age	7.13 ± 0.23	<0.0001	15.83 ± 0.76	<0.0001
Black	20.62 ± 5.44	0.00015	-60.43 ± 18.11	0.00086
Hispanic	-13.07 ± 5.72	0.022	-73.33 ± 19.02	0.00012
Asian	-24.59 ± 7.13	0.00059	-204.86 ± 23.80	<0.0001
Male gender	32.19 ± 4.81	<0.0001	78.24 ± 16.01	<0.0001
BMI	3.03 ± 0.44	<0.0001	1.58 ± 1.47	0.25
Former smoker	11.28 ± 5.24	0.032	47.54 ± 17.45	0.0064
Current smoker	5.09 ± 7.37	0.49	96.52 ± 24.51	<0.0001
Pack years	0.46 ± 0.12	0.00014	2.49 ± 0.40	<0.0001
SBP	1.53 ± 0.11	<0.0001	2.72 ± 0.36	<0.0001
HTN Rx	4.03 ± 4.77	0.39	70.19 ± 15.86	<0.0001
Fasting glucose	0.16 ± 0.094	0.081	0.50 ± 0.31	0.11
Diabetes status	20.31 ± 8.70	0.020	83.89 ± 28.96	0.0038
LDL-C	0.41 ± 0.069	<0.0001	1.05 ± 0.23	<0.0001
HDL-C	-0.19 ± 0.41	0.63	-0.45 ± 1.37	0.74
LLM Rx	20.19 ± 5.87	0.00058	135.23 ± 19.52	<0.0001
Large HDL-P	-1.61 ± 1.74	0.36	-3.13 ± 1.59	0.59
Small/medium HDL-P	-1.68 ± 0.48	0.00042	-4.26 ± 1.59	0.0073

Abbreviations: BMI – body mass index; HDL-C – high-density lipoprotein cholesterol; HDL-P – high-density lipoprotein particle concentration; HTN Rx – hypertension medication treatment; LDL-C – low-density lipoprotein cholesterol; LLM Rx – lipid lowering medication treatment; SBP – systolic blood pressure.

Supplemental Table A2. Sensitivity analysis by race of association of small/medium HDL-P with common cIMT.

	White *N*=2622		Black *N*=1892		Hispanic *N*=1496		Asian *N*=804	
Covariates	**Beta ± SE**	***P***	**Beta ± SE**	***P***	**Beta ± SE**	***P***	**Beta ± SE**	***P***
Age	8.05 ± 0.36	<0.0001	6.76 ± 0.46	<0.0001	7.05 ± 0.50	<0.0001	4.87 ± 0.63	<0.0001
Male gender	24.57 ± 7.87	0.0018	49.57 ± 9.45	<0.0001	33.29 ± 10.48	0.00153	12.10 ± 13.85	0.38
BMI	3.43 ± 0.73	<0.0001	2.49 ± 0.78	0.0015	3.19 ± 0.95	0.00082	4.98 ± 1.85	0.0073
Former smoker	15.34 ± 7.98	0.055	9.21 ± 10.48	0.38	9.62 ± 11.36	0.40	9.22 ± 18.73	0.63
Current smoker	14.54 ± 12.48	0.24	-14.29 ± 13.38	0.29	23.12 ± 14.93	0.12	20.96 ± 27.53	0.45
Pack years	0.38 ± 0.16	0.020	0.54 ± 0.26	0.036	0.28 ± 0.32	0.39	0.91 ± 0.49	0.065
SBP	1.71 ± 0.18	<0.0001	1.23 ± 0.20	<0.0001	1.54 ± 0.23	<0.0001	1.75 ± 0.29	<0.0001
HTN Rx	7.57 ± 7.58	0.32	-3.52 ± 9.06	0.69	17.14 ± 10.57	0.11	-10.02 ± 13.94	0.47
Fasting glucose	0.22 ± 0.21	0.31	0.18 ± 0.17	0.30	0.044 ± 0.16	0.78	0.37 ± 0.27	0.17
Diabetes status	28.13 ± 18.48	0.13	7.38 ± 14.79	0.62	32.69 ± 16.82	0.052	25.57 ± 23.34	0.27
LDL-C	0.32 ± 0.12	0.0066	0.43 ± 0.13	0.00091	0.55 ± 0.14	0.00013	0.36 ± 0.20	0.077
HDL-C	-0.29 ± 0.63	0.65	-1.51 ± 0.82	0.067	1.22 ± 0.92	0.18	0.49 ± 1.17	0.67
LLM Rx	11.45 ± 8.96	0.20	46.86 ± 11.43	<0.0001	-8.77 ± 13.72	0.52	33.40 ± 17.21	0.053
Large HDL-P	-1.40 ± 2.77	0.61	3.94 ± 3.46	0.25	-6.85 ± 3.87	0.077	-4.07 ± 4.66	0.38
Small/Med HDL-P	-1.33 ± 0.75	0.075	-1.58 ± 0.93	0.089	-2.77 ± 1.06	0.0089	-0.88 ± 1.41	0.53

Linear regression analyses of small/medium HDL-P and all other covariates in **Supplemental Table A1** association with common cIMT in the MESA study, stratified by race.

Abbreviations: BMI – body mass index; HDL-C – high-density lipoprotein cholesterol; HDL-P – high-density lipoprotein particle concentration; HTN Rx – hypertension medication treatment; LDL-C – low-density lipoprotein cholesterol; LLM Rx – lipid lowering medication treatment; SBP – systolic blood pressure.

Supplemental Table A3. Sensitivity analysis by race of association of small/medium HDL-P with internal cIMT.

	White *N*=2622		Black *N*=1892		Hispanic *N*=1496		Asian *N*=804	
Covariates	**Beta ± SE**	***P***	**Beta ± SE**	***P***	**Beta ± SE**	***P***	**Beta ± SE**	***P***
Age	18.48 ± 1.24	<0.0001	15.43 ± 1.53	<0.0001	17.31 ± 1.66	<0.0001	6.69 ± 1.77	0.00017
Male gender	66.82 ± 26.97	0.013	54.06 ± 31.71	0.088	142.07 ± 34.49	<0.0001	56.58 ± 39.62	0.15
BMI	2.74 ± 2.49	0.27	1.15 ± 2.62	0.66	2.02 ± 3.13	0.52	-4.97 ± 5.29	0.35
Former smoker	21.49 ± 27.38	0.43	99.98 ± 35.10	0.0043	56.41 ± 37.39	0.13	3.72 ± 53.59	0.94
Current smoker	91.38 ± 42.79	0.033	107.60 ± 44.88	0.017	167.10 ± 49.14	0.00069	-21.92 ± 78.77	0.78
Pack years	3.29 ± 0.56	<0.0001	0.96 ± 0.86	0.26	1.46 ± 1.04	0.16	3.23 ± 1.41	0.023
SBP	2.89 ± 0.62	<0.0001	3.79 ± 0.69	<0.0001	2.02 ± 0.75	0.0070	1.58 ± 0.84	0.059
HTN Rx	147.83 ± 25.99	<0.0001	4.91 ± 30.38	0.87	54.91 ± 34.79	0.11	16.96 ± 39.88	0.67
Fasting glucose	0.41 ± 0.73	0.57	0.045 ± 0.58	0.94	0.065 ± 0.52	0.90	1.79 ± 0.77	0.021
Diabetes status	21.53 ± 63.36	0.73	69.26 ± 49.62	0.16	202.07 ± 55.37	0.00027	96.73 ± 66.78	0.15
LDL-C	1.39 ± 0.39	0.00049	0.40 ± 0.44	0.36	1.57 ± 0.47	0.00087	0.63 ± 0.58	0.28
HDL-C	-1.31 ± 2.18	0.55	-2.31 ± 2.76	0.40	2.79 ± 3.02	0.36	-0.11 ± 3.34	0.97
LLM Rx	118.85 ± 30.72	0.00011	204.33 ± 38.34	<0.0001	49.07 ± 45.17	0.28	199.71 ± 49.26	<0.0001
Large HDL-P	0.64 ± 9.48	0.95	3.92 ± 11.62	0.74	-13.57 ± 12.74	0.29	-8.39 ± 13.33	0.53
Small/Med HDL-P	-5.51 ± 2.56	0.032	-3.54 ± 3.11	0.26	-2.66 ± 3.48	0.44	-3.59 ± 4.04	0.37

Linear regression analyses of small/medium HDL-P and all other covariates in **Supplemental Table A1** association with internal cIMT in the MESA study, stratified by race.

Abbreviations: BMI – body mass index; HDL-C – high-density lipoprotein cholesterol; HDL-P – high-density lipoprotein particle concentration; HTN Rx – hypertension medication treatment; LDL-C – low-density lipoprotein cholesterol; LLM Rx – lipid lowering medication treatment; SBP – systolic blood pressure.

Supplemental Table A4. Sensitivity analysis by sex of association of small/medium HDL-P with common cIMT and internal cIMT.

	Common cIMT				Internal cIMT			
	Male *N*=3173		Female *N*=3553		Male *N*=3137		Female *N*=3492	
Covariates	**Beta ± SE**	***P***	**Beta ± SE**	***P***	**Beta ± SE**	***P***	**Beta ± SE**	***P***
Age	7.53 ± 0.35	<0.0001	6.80 ± 0.30	<0.0001	16.65 ± 1.16	<0.0001	15.26 ± 1.02	<0.0001
Black	29.36 ± 8.24	0.00037	17.03 ± 7.24	0.019	-76.33 ± 27.38	0.0053	-42.67 ± 24.26	0.079
Hispanic	-11.52 ± 8.53	0.18	-13.67 ± 7.70	0.076	-41.08 ± 28.35	0.15	-100.30 ± 25.79	0.00010
Asian	-18.48 ± 10.67	0.085	-24.95 ± 9.73	0.010	-207.95 ± 35.62	<0.0001	-195.39 ± 32.59	<0.0001
BMI	5.78 ± 0.81	<0.0001	1.90 ± 0.52	0.00025	3.63 ± 2.70	0.18	0.68 ± 1.74	0.70
Former smoker	12.67 ± 7.68	0.099	10.37 ± 7.34	0.16	11.59 ± 25.52	0.65	86.83 ± 24.57	0.00041
Current smoker	14.08 ± 10.86	0.19	-2.41 ± 10.17	0.81	82.98 ± 36.11	0.022	113.48 ± 34.07	0.00088
Pack years	0.42 ± 0.16	0.0076	0.45 ± 0.20	0.022	2.97 ± 0.52	<0.0001	1.52 ± 0.66	0.021
SBP	1.55 ± 0.18	<0.0001	1.52 ± 0.13	<0.0001	3.23 ± 0.59	<0.0001	2.47 ± 0.45	<0.0001
HTN Rx	7.46 ± 7.23	0.30	0.85 ± 6.28	0.89	95.35 ± 24.05	<0.0001	49.22 ± 21.04	0.019
Fasting glucose	0.015 ± 0.13	0.91	0.39 ± 0.14	0.0056	0.46 ± 0.43	0.28	0.55 ± 0.47	0.24
Diabetes status	40.59 ± 12.49	0.0012	-7.07 ± 12.17	0.56	85.90 ± 41.55	0.039	76.34 ± 40.77	0.061
LDL-C	0.57 ± 0.11	<0.0001	0.30 ± 0.090	0.00084	1.23 ± 0.36	0.00062	0.93 ± 0.30	0.0021
HDL-C	0.53 ± 0.68	0.43	-0.61 ± 0.51	0.23	1.42 ± 2.25	0.53	-1.87 ± 1.70	0.27
LLM Rx	24.76 ± 8.88	0.0053	18.19 ± 7.78	0.020	106.13 ± 29.51	0.00033	163.73 ± 26.07	<0.0001
Large HDL-P	-4.40 ± 2.85	0.12	0.039 ± 2.17	0.99	-10.08 ± 9.49	0.29	1.52 ± 7.26	0.83
Small/Med HDL-P	-3.54 ± 0.83	<0.0001	-0.70 ± 0.57	0.22	-4.91 ± 2.76	0.076	-3.69 ± 1.92	0.055

Linear regression analyses of small/medium HDL-P and all other covariates in **Supplemental Table A1** association with common and internal cIMT in the MESA study, stratified by sex.

Abbreviations: BMI – body mass index; HDL-C – high-density lipoprotein cholesterol; HDL-P – high-density lipoprotein particle concentration; HTN Rx – hypertension medication treatment; LDL-C – low-density lipoprotein cholesterol; LLM Rx – lipid lowering medication treatment; SBP – systolic blood pressure.

APPENDIX B: SUPPLEMENTAL TABLE FOR CHAPTER 4

Supplemental Table B1. Demographic and clinical characteristics of the studied CLEAR subset stratified by CAAD case status.

Variable	N	CAAD Controls (n=622)	CAAD Cases (n=493)	Total (n=1115)	P-Value
Censored age, y	1115	66.7 ± 8.0	66.5 ± 9.1	66.6 ± 8.5	0.98[a]
Female, %	1115	106 (17%)	60 (12%)	166 (15%)	0.031[b]
Diabetic, %	1115	61 (10%)	115 (23%)	176 (16%)	< 0.001[b]
BMI, kg/m^2	1115	28.4 ± 5.0	27.9 ± 4.7	28.2 ± 4.9	0.14[a]
Statin use, %	1115	109 (18%)	353 (72%)	462 (41%)	< 0.001[b]
Current smoking status, %	1115	35 (6%)	94 (19%)	129 (11%)	< 0.001[b]
Total cholesterol, mg/dl	1100	197 ± 40	176 ± 40	187 ± 41	< 0.001[a]
apoA1, mg/dl	1110	148 ± 28	138 ± 27	144 ± 28	< 0.001[a]
HDL-C, mg/dl	1100	53 ± 17	47 ± 15	50 ± 17	< 0.001[a]
HDL2, mg/dl	1098	10 ± 6.5	8.6 ± 5.8	9.6 ± 6.3	< 0.001[a]
HDL3, mg/dl	1099	43 ± 12	38 ± 11	41 ± 11	< 0.001[a]
Ln(Triglycerides)	1100	4.7 ± 0.56	4.8 ± 0.55	4.8 ± 0.56	< 0.001[a]
Ln(VLDL)	1100	3.1 ± 0.55	3.3 ± 0.54	3.2 ± 0.55	< 0.001[a]
PON1 AREase activity, IU	1077	148 ± 50	128 ± 50	139 ± 50	< 0.001[a]
PLTPa, μM/hr	1115	13.8 ± 4.2	13.2 ± 4.7	13.5 ± 4.4	0.008[a]

Abbreviations: apoA1 = apolipoprotein A1; BMI = body mass index; CAAD = carotid artery disease; censored age = age at CAAD diagnosis for cases, age at enrollment for controls; CLEAR = Carotid Lesion Epidemiology and Risk cohort; HDL = high-density lipoprotein; PON1 AREase activity = Paraoxonase 1 arylester hydrolysis activity; PLTPa = phospholipid transfer protein activity; VLDL = very-low-density lipoprotein.

Tests used for p-value calculations for differences between subsets: [a] = Wilcoxon rank sum test; [b] = Pearson chi-square test.

Supplemental Table B2. Sensitivity analysis of PLTPa predictors from stepwise linear regression in controls (n=601).*

Variable	Coefficient ± SE	% PLTPa	P-Value
Intercept	11.79 ± 2.16	-	-
Current age	0.00014 ± 0.025	0.12%	0.99
Female gender	-0.13 ± 0.52	0.066%	0.79
Ln(Triglycerides)	0.92 ± 0.35	1.31%	0.0097
Diabetes status	1.73 ± 0.61	1.69%	0.0045
PON1 AREase activity	0.0056 ± 0.0042	0.53%	0.19
Current smoking status	1.07 ± 0.77	0.37%	0.17
apoA1, mg/dl	0.0078 ± 0.0079	0.19%	0.32
Statin use	-1.02 ± 0.48	0.54%	0.037

Abbreviations: apoA1 = apolipoprotein A1; PON1 AREase activity = paraoxonase 1 arylester hydrolysis enzymatic activity; PLTPa = phospholipid transfer protein activity.

These analyses relate to **Table 11 in Chapter 4. They show the association of each individual covariate identified as predictive of PLTP activity stratified by CAAD status (associations for CAAD controls reported in this table).*

Supplemental Table B3. Sensitivity analysis of PLTPa predictors from stepwise linear regression in CAAD cases (n=455).*

Variable	Coefficient ± SE	% PLTPa	P-Value
Intercept	11.28 ± 2.67	-	-
Current age	-0.031 ± 0.028	1.12%	0.48
Female gender	-1.04 ± 0.72	0.19%	0.15
Ln(Triglycerides)	0.98 ± 0.46	1.59%	0.034
Diabetes status	0.93± 0.51	0.52%	0.074
PON1 AREase activity	0.011 ± 0.0048	1.31%	0.023
Current smoking status	0.18 ± 0.59	0.016%	0.75
apoA1, mg/dl	0.0043 ± 0.0091	0.048%	0.64
Statin use	0.13 ± 0.53	0.13%	0.80

Abbreviations: apoA1 = apolipoprotein A1; PON1 AREase activity = paraoxonase 1 arylester hydrolysis enzymatic activity; PLTPa = phospholipid transfer protein activity.

These analyses relate to **Table 11 in Chapter 4. They show the association of each individual covariate identified as predictive of PLTP activity stratified by CAAD status (associations for CAAD cases reported in this table).*

Supplemental Table B4. PLTPa predicts CAAD status adjusting for confounders through multivariate logistic regression (n=1056)*.

Variable	OR (95% CI)	% CAAD Prediction	P-Value
Censored Age	1.00 (0.99-1.02)	0.054%	0.58
Female Gender	1.07 (0.69-1.64)	0.65%	0.79
Ln(Triglycerides) [a]	1.27 (1.06-1.53)	2.34%	0.005
Diabetes Status	3.25 (2.24-4.74)	3.59%	6.31×10^{-9}
PON1 AREase activity [b]	0.94 (0.92-0.98)	2.69%	0.001
Current Smoking Status	4.55 (2.92-7.21)	4.34%	1.38×10^{-12}
apoA1 [b]	0.96 (0.90-1.02)	0.30%	0.12
PLTPa [c]	0.91 (0.88-0.94)	2.92%	7.90×10^{-9}

Abbreviations: apoA1 = apolipoprotein A1; CAAD = carotid artery disease; PON1 AREase activity = paraoxonase 1 arylester hydrolysis enzymatic activity; PLTPa = phospholipid transfer protein activity.

[a] Reported odds ratio for ln(triglyceride) association with CAAD odds is for a 2-fold increase in triglycerides.
[b] Reported odds ratios for PON1 AREase activity and apoA1 associations with CAAD odds are for a 10 unit increase in each respective covariate.
[c] PLTPa adjusted for statin use by the average decrease seen in controls using statins (-1.02).

*1056 participants had complete data on all parameters considered in this stepwise linear regression model. All included covariates were previously associated with PLTPa (see **Table 11**).

Supplemental Table B5. *PLTP* and local region SNP associations (P ≤ 0.05) with covariate-adjusted odds of CAAD (logistic regression) and PLTPa (linear regression).

SNP	Position[a]	Gene[b]	Function	MAF	N	CAAD OR (95% CI)[c]	CAAD P[c]	PLTP Beta[d]	PLTP P[d]
rs4810479	20:44545048	*(PLTP)*	5'	0.253	1025	0.84 (0.67-1.04)	0.097	-1.183	2.63E-07
rs4812975	20:44545460	*(PLTP)*	5'	0.253	1022	0.84 (0.68-1.05)	0.12	-1.043	2.28E-06
rs2868346	20:44547970	*(PLTP)*	5'	0.2486	1015	0.84 (0.67-1.04)	0.11	-1.012	4.60E-06
rs6065904	20:44534651	*PLTP*	Intron 8[e]	0.2212	1025	0.74 (0.59-0.94)	0.012	-1.054	1.35E-05
chr20:44546198:D	20:44546198	*(PLTP)*	TA>T Del	0.1867	1001	0.77 (0.60-0.99)	0.040	-1.056	1.91E-05
rs73307905	20:44545773	*(PLTP)*	5'	0.1863	1021	0.77 (0.59-0.98)	0.034	-1.058	1.94E-05
rs7679	20:44576502	*PCIF1*	3'UTR	0.1772	987	0.76 (0.59-0.98)	0.035	-1.102	2.84E-05
rs58847685	20:44544947	*(PLTP)*	5'	0.1863	1018	0.78 (0.61-0.99)	0.044	-1.03	3.28E-05
rs111602331	20:44557474	*(PLTP)*	5'	0.1753	986	0.73 (0.57-0.95)	0.017	-1.048	4.35E-05
rs139953093	20:44559748	*(PLTP)*	5'	0.1767	982	0.76 (0.59-0.99)	0.038	-1.03	4.73E-05
rs6073957	20:44549476	*(PLTP)*	5'	0.1841	1019	0.77 (0.60-0.98)	0.034	-0.9936	5.98E-05
rs118024629	20:44546197	*(PLTP)*	5'	0.1852	1019	0.76 (0.60-0.98)	0.031	-0.9902	6.33E-05
rs12185764	20:44550020	*(PLTP)*	5'	0.1847	1019	0.76 (0.60-0.98)	0.031	-0.9902	6.33E-05
rs59329875	20:44547672	*(PLTP)*	5'	0.1852	1019	0.76 (0.60-0.98)	0.031	-0.9902	6.33E-05
rs73307913	20:44548301	*(PLTP)*	5'	0.1852	1019	0.76 (0.60-0.98)	0.031	-0.9902	6.33E-05
rs6065906	20:44554015	*(PLTP)*	5'	0.1849	1011	0.75 (0.59-0.96)	0.025	-0.986	6.80E-05
rs12185776	20:44551600	*(PLTP)*	5'	0.1841	1017	0.77 (0.60-0.98)	0.035	-0.9682	9.18E-05
rs6073958	20:44551855	*(PLTP)*	5'	0.1937	998	0.82 (0.64-1.04)	0.099	-0.9515	0.0001
rs6073952	20:44536932	*PLTP*	Intron 4[e]	0.1974	1025	0.80 (0.63-1.01)	0.069	-0.8945	0.0003
rs6073966	20:44570192	*PCIF1*	Intron 7[f]	0.1644	986	0.75 (0.58-0.98)	0.031	-0.9076	0.0004
rs6065905	20:44537837	*PLTP*	Intron 4[e]	0.198	1018	0.81 (0.64-1.03)	0.080	-0.8009	0.0009
rs378114	20:44538427	*PLTP*	Intron 3[e]	0.2656	1025	0.95 (0.77-1.17)	0.61	0.5145	0.024
rs435306	20:44538484	*PLTP*	Intron 3[e]	0.2656	1025	0.95 (0.77-1.17)	0.61	0.5145	0.024
rs4810478	20:44544732	*(PLTP)*	5'	0.05424	1004	1.13 (0.75-1.71)	0.55	-0.8958	0.036
rs6073955	20:44544571	*(PLTP)*	5'	0.05424	1004	1.13 (0.75-1.71)	0.55	-0.8958	0.036
rs2903809	20:44552895	*(PLTP)*	5'	0.2705	998	0.93 (0.75-1.15)	0.49	0.447	0.037
rs3843763	20:44548193	*(PLTP)*	5'	0.2721	1004	0.93 (0.76-1.15)	0.53	0.4247	0.048

rs41305805	20:44587926	*ZNF335*	Phe>Val	0.01749	995	1.18 (0.61-2.28)	0.63	1.335	0.049
rs118013520	20:44563133	*(PCIF1)*	5’	0.0175	995	1.18 (0.61-2.27)	0.62	1.329	0.04933
rs181914932	20:44561279	*(PCIF1)*	5’	0.0175	995	1.18 (0.61-2.27)	0.62	1.329	0.04933
rs79629788	20:44562148	*(PCIF1)*	5’	0.0175	995	1.18 (0.61-2.27)	0.62	1.329	0.04933

Abbreviations: Beta = beta coefficient from linear regression; CI = confidence interval; MAF = minor allele frequency; OR = odds ratio.

[a] = Position based on hg19/GRCh37.

[b] = Noncoding SNPs are represented in parentheses naming the nearest gene, e.g. *(PLTP)*.

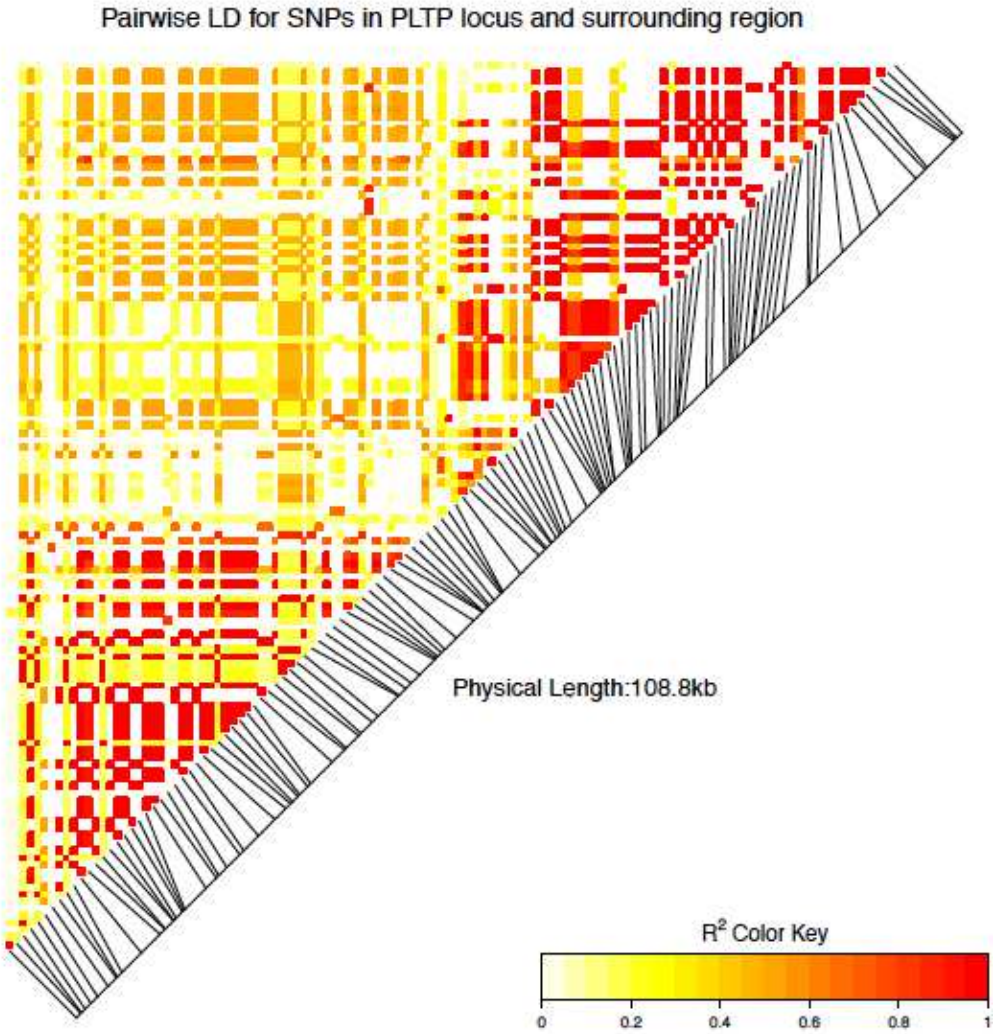
[c] = Logistic regression analyses adjusted by censored age, sex, diabetes status, ln(triglyceride levels), current smoking status, and PON1 AREase activity.

[d] = Linear regression analyses adjusted by age, sex, diabetes status, statin use, ln(triglyceride levels), current smoking status and PON1 AREase activity.

[e] = *PLTP* gene has 14 total introns.

[f] = *PCIF1* gene has 15 total introns.

Supplemental Figure B1. Linkage disequilibrium (LD) of *PTLP* and surrounding genomic region SNPs with minor allele frequency greater than 5%. This plot demonstrates the local correlation (R^2), or LD, of *PLTP* region SNPs that were studied. Red denotes strong correlation, indicating that these SNPs' genotypes (which create "haplotypes") are inherited together frequently, while white indicates no correlation. Note the lack of solid "blocks" of SNPs in LD (continuous red blocks) that would indicate large regions or haplotypes are inherited together.



REFERENCES

1. Go, A. S. *et al.* Heart Disease and Stroke Statistics--2013 Update: A Report From the American Heart Association. *Circulation* **127,** e6–e245 (2013).
2. ECST. MRC European Carotid Surgery Trial: interim results for symptomatic patients with severe (70-99%) or with mild (0-29%) carotid stenosis. European Carotid Surgery Trialists' Collaborative Group. *Lancet* **337,** 1235–1243 (1991).
3. ECST. Endarterectomy for moderate symptomatic carotid stenosis: interim results from the MRC European Carotid Surgery Trial. *The Lancet* **347,** 1591–1593 (1996).
4. North American Symptomatic Carotid Endarterectomy Trial Collaborators. Beneficial effect of carotid endarterectomy in symptomatic patients with high-grade carotid stenosis. *N Engl J Med* **325,** 445–453 (1991).
5. Willeit, J. & Kiechl, S. Prevalence and risk factors of asymptomatic extracranial carotid artery atherosclerosis. A population-based study. *Arteriosclerosis, Thrombosis, and Vascular Biology* **13,** 661–668 (1993).
6. Steg, P. G. *et al.* One-year cardiovascular event rates in outpatients with atherothrombosis. *JAMA* **297,** 1197–1206 (2007).
7. Castelli, W. P. Cardiovascular disease and multifactorial risk: challenge of the 1980s. *Am. Heart J.* **106,** 1191–1200 (1983).
8. Barter, P. J. *et al.* Effects of torcetrapib in patients at high risk for coronary events. *N Engl J Med* **357,** 2109–2122 (2007).
9. Rohatgi, A. *et al.* HDL Cholesterol Efflux Capacity and Incident Cardiovascular Events. *N Engl J Med* 141118051511004 (2014). doi:10.1056/NEJMoa1409065
10. Watson, A. D. *et al.* Protective effect of high density lipoprotein associated paraoxonase.

Inhibition of the biological activity of minimally oxidized low density lipoprotein. *J. Clin. Invest.* **96,** 2882–2891 (1995).

11. Ashby, D. T. *et al.* Factors Influencing the Ability of HDL to Inhibit Expression of Vascular Cell Adhesion Molecule-1 in Endothelial Cells. *Arteriosclerosis, Thrombosis, and Vascular Biology* **18,** 1450–1455 (1998).

12. Cockerill, G. W., Rye, K. A., Gamble, J. R., Vadas, M. A. & Barter, P. J. High-density lipoproteins inhibit cytokine-induced expression of endothelial cell adhesion molecules. *Arteriosclerosis, Thrombosis, and Vascular Biology* **15,** 1987–1994 (1995).

13. Navab, M. Oral Administration of an Apo A-I Mimetic Peptide Synthesized From D-Amino Acids Dramatically Reduces Atherosclerosis in Mice Independent of Plasma Cholesterol. *Circulation* **105,** 290–292 (2002).

14. Vaisar, T. *et al.* Shotgun proteomics implicates protease inhibition and complement activation in the antiinflammatory properties of HDL. *J. Clin. Invest.* **117,** 746–756 (2007).

15. Henning, M. F., Herlax, V. & Bakas, L. Contribution of the C-terminal end of apolipoprotein AI to neutralization of lipopolysaccharide endotoxic effect. *Innate Immunity* **17,** 327–337 (2011).

16. Suzuki, M. *et al.* High-Density Lipoprotein Suppresses the Type I Interferon Response, a Family of Potent Antiviral Immunoregulators, in Macrophages Challenged With Lipopolysaccharide. *Circulation* **122,** 1919–1927 (2010).

17. Yuhanna, I. S. *et al.* High-density lipoprotein binding to scavenger receptor-BI activates endothelial nitric oxide synthase. *Nat. Med.* **7,** 853–857 (2001).

18. Blanche, P. J., Gong, E. L., Forte, T. M. & Nichols, A. V. Characterization of human

high-density lipoproteins by gradient gel electrophoresis. *Biochim. Biophys. Acta* **665,** 408–419 (1981).

19. Asztalos, B. F., Tani, M. & Schaefer, E. J. Metabolic and functional relevance of HDL subspecies. *Current Opinion in Lipidology* **22,** 176–185 (2011).

20. Tian, L. & Fu, M. The relationship between high densitylipoprotein subclass profile and plasma. *Lipids Health Dis* **9,** 118 (2010).

21. Asztalos, B. F. *et al.* High-density lipoprotein subpopulation profile and coronary heart disease prevalence in male participants of the Framingham Offspring Study. *Arteriosclerosis, Thrombosis, and Vascular Biology* **24,** 2181–2187 (2004).

22. Asztalos, B. F. *et al.* Value of high-density lipoprotein (HDL) subpopulations in predicting recurrent cardiovascular events in the Veterans Affairs HDL Intervention Trial. *Arteriosclerosis, Thrombosis, and Vascular Biology* **25,** 2185–2191 (2005).

23. Kontush, A., Chantepie, S. & Chapman, M. J. Small, dense HDL particles exert potent protection of atherogenic LDL against oxidative stress. *Arteriosclerosis, Thrombosis, and Vascular Biology* **23,** 1881–1888 (2003).

24. Shuhei, N., Söderlund, S., Jauhiainen, M. & Taskinen, M.-R. Effect of HDL composition and particle size on the resistance of HDL to the oxidation. *Lipids Health Dis* **9,** 104 (2010).

25. Mackness, M. I., Arrol, S. & Durrington, P. N. Paraoxonase prevents accumulation of lipoperoxides in low-density lipoprotein. *FEBS Lett.* **286,** 152–154 (1991).

26. Mackness, M. I., Arrol, S., Abbott, C. & Durrington, P. N. Protection of low-density lipoprotein against oxidative modification by high-density lipoprotein associated paraoxonase. *Atherosclerosis* **104,** 129–135 (1993).

27. Aviram, M., Rosenblat, M., Bisgaier, C. L. & Newton, R. S. Atorvastatin and gemfibrozil metabolites, but not the parent drugs, are potent antioxidants against lipoprotein oxidation. *Atherosclerosis* **138,** 271–280 (1998).

28. Jarvik, G. P. *et al.* Paraoxonase (PON1) Phenotype Is a Better Predictor of Vascular Disease Than Is PON1192 or PON155 Genotype. *Arteriosclerosis, Thrombosis, and Vascular Biology* **20,** 2441–2447 (2000).

29. Jarvik, G. P. *et al.* Paraoxonase activity, but not haplotype utilizing the linkage disequilibrium structure, predicts vascular disease. *Arteriosclerosis, Thrombosis, and Vascular Biology* **23,** 1465–1471 (2003).

30. Kim, D. S. *et al.* Additional Common Polymorphisms in the PON Gene Cluster Predict PON1 Activity but Not Vascular Disease. *J Lipids* **2012,** 476316–11 (2012).

31. Bhattacharyya, T. *et al.* Relationship of paraoxonase 1 (PON1) gene polymorphisms and functional activity with systemic oxidative stress and cardiovascular risk. *JAMA* **299,** 1265–1276 (2008).

32. Cheung, M. C. & Albers, J. J. Active plasma phospholipid transfer protein is associated with apoA-I-but not apoE-containing lipoproteins. *The Journal of Lipid Research* **47,** 1315–1321 (2006).

33. O'Brien, K. D. *et al.* Cell-associated and extracellular phospholipid transfer protein in human coronary atherosclerosis. *Circulation* **108,** 270–274 (2003).

34. Desrumaux, C. M. *et al.* Phospholipid transfer protein is present in human atherosclerotic lesions and is expressed by macrophages and foam cells. *The Journal of Lipid Research* **44,** 1453–1461 (2003).

35. Wolfbauer, G., Albers, J. J. & Oram, J. F. Phospholipid transfer protein enhances

removal of cellular cholesterol and phospholipids by high-density lipoprotein apolipoproteins. *Biochim. Biophys. Acta* **1439,** 65–76 (1999).

36. Oka, T. *et al.* Distribution of phospholipid transfer protein in human plasma: presence of two forms of phospholipid transfer protein, one catalytically active and the other inactive. *The Journal of Lipid Research* **41,** 1651–1657 (2000).

37. Tall, A. R., Krumholz, S., Olivecrona, T. & Deckelbaum, R. J. Plasma phospholipid transfer protein enhances transfer and exchange of phospholipids between very low density lipoproteins and high density lipoproteins during lipolysis. *The Journal of Lipid Research* **26,** 842–851 (1985).

38. Bailey, D. *et al.* Analysis of lipid transfer activity between model nascent HDL particles and plasma lipoproteins: implications for current concepts of nascent HDL maturation and genesis. *The Journal of Lipid Research* **51,** 785–797 (2010).

39. Chirackal Manavalan, A. P. *et al.* Phospholipid transfer protein is expressed in cerebrovascular endothelial cells and involved in high density lipoprotein biogenesis and remodeling at the blood-brain barrier. *Journal of Biological Chemistry* **289,** 4683–4698 (2014).

40. Sacks, F. M. *et al.* Coronary Heart Disease in Patients With Low LDL-Cholesterol Benefit of Pravastatin in Diabetics and Enhanced Role for HDL-Cholesterol and Triglycerides as Risk Factors. *Circulation* **105,** 1424–1428 (2002).

41. Taylor, A. J. Arterial Biology for the Investigation of the Treatment Effects of Reducing Cholesterol (ARBITER) 2: A Double-Blind, Placebo-Controlled Study of Extended-Release Niacin on Atherosclerosis Progression in Secondary Prevention Patients Treated With Statins. *Circulation* **110,** 3512–3517 (2004).

42. AIM-HIGH Investigators *et al.* Niacin in patients with low HDL cholesterol levels receiving intensive statin therapy. *N Engl J Med* **365,** 2255–2267 (2011).

43. Voight, B. F. *et al.* Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. *The Lancet* **380,** 572–580 (2012).

44. Schwartz, G. G. *et al.* Effects of Dalcetrapib in Patients with a Recent Acute Coronary Syndrome. *N Engl J Med* **367,** 2089–2099 (2012).

45. Cannon, C. P. *et al.* Safety of Anacetrapib in Patients with or at High Risk for Coronary Heart Disease. *N Engl J Med* **363,** 2406–2415 (2010).

46. deGoma, E. M., deGoma, R. L. & Rader, D. J. Beyond High-Density Lipoprotein Cholesterol Levels. *Journal of the American College of Cardiology* **51,** 2199–2211 (2008).

47. Mackey, R. H. *et al.* High-density lipoprotein cholesterol and particle concentrations, carotid atherosclerosis, and coronary events: MESA (multi-ethnic study of atherosclerosis). *Journal of the American College of Cardiology* **60,** 508–516 (2012).

48. Kim, D. S. *et al.* HDL-3 is a Superior Predictor of Carotid Artery Disease in a Case-Control Cohort of 1725 Participants. *Journal of the American Heart Association* **3,** e000902–e000902 (2014).

49. Kim, D. S. *et al.* PLTP activity inversely correlates with CAAD: effects of PON1 enzyme activity and genetic variants on PLTP activity. *The Journal of Lipid Research* **56,** 1351–1362 (2015).

50. Kim, D. S. *et al.* Rare coding variation in paraoxonase-1 is associated with ischemic stroke in the NHLBI Exome Sequencing Project. *The Journal of Lipid Research* **55,** 1173–1178 (2014).

51. Go, A. S. *et al.* Heart disease and stroke statistics--2014 update: a report from the American Heart Association. *Circulation* **129,** e28–e292 (2014).

52. Lozano, R. *et al.* Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **380,** 2095–2128 (2012).

53. Kim, D. S. *et al.* Dietary cholesterol increases paraoxonase 1 enzyme activity. *J. Lipid Res.* **53,** 2450–2458 (2012).

54. Kim, D. S. *et al.* Novel common and rare genetic determinants of paraoxonase activity: FTO, SERPINA12, and ITGAL. *The Journal of Lipid Research* **54,** 552–560 (2013).

55. Kim, D. S. *et al.* Novel gene-by-environment interactions: APOB and NPC1L1 variants affect the relationship between dietary and total plasma cholesterol. *The Journal of Lipid Research* **54,** 1512–1520 (2013).

56. Kim, D. S. *et al.* Dietary fatty acid intake is associated with paraoxonase 1 activity in a cohort-based analysis of 1,548 subjects. *Lipids Health Dis* **12,** 183 (2013).

57. Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155,** 945–959 (2000).

58. Gunderson, K. L., Steemers, F. J., Lee, G., Mendoza, L. G. & Chee, M. S. A genome-wide scalable SNP genotyping assay using microarray technology. *Nat. Genet.* **37,** 549–554 (2005).

59. Steemers, F. J. *et al.* Whole-genome genotyping with the single-base extension assay. *Nat Meth* **3,** 31–33 (2006).

60. Marcovina, S. M., Albers, J. J., Henderson, L. O. & Hannon, W. H. International Federation of Clinical Chemistry standardization project for measurements of

apolipoproteins A-I and B. III. Comparability of apolipoprotein A-I values by use of international reference material. *Clin. Chem.* **39,** 773–781 (1993).

61. Richter, R. J., Jarvik, G. P. & Furlong, C. E. Determination of paraoxonase 1 status without the use of toxic organophosphate substrates. *Circulation: Cardiovascular Genetics* **1,** 147–152 (2008).

62. Richter, R. J., Jarvik, G. P. & Furlong, C. E. Paraoxonase 1 (PON1) status and substrate hydrolysis. *Toxicology and Applied Pharmacology* **235,** 1–9 (2009).

63. Superko, H. R. *et al.* High-density lipoprotein subclasses and their relationship to cardiovascular disease. *J Clin Lipidol* **6,** 496–523 (2012).

64. Lamarche, B. *et al.* Associations of HDL2 and HDL3 subfractions with ischemic heart disease in men. Prospective results from the Québec Cardiovascular Study. *Arteriosclerosis, Thrombosis, and Vascular Biology* **17,** 1098–1105 (1997).

65. Asztalos, B. F. *et al.* Change in alpha1 HDL concentration predicts progression in coronary artery stenosis. *Arteriosclerosis, Thrombosis, and Vascular Biology* **23,** 847–852 (2003).

66. Bergmeier, C., Siekmeier, R. & Gross, W. Distribution spectrum of paraoxonase activity in HDL fractions. *Clin. Chem.* **50,** 2309–2315 (2004).

67. Mackness, M. I. *et al.* Serum paraoxonase activity in familial hypercholesterolaemia and insulin-dependent diabetes mellitus. *Atherosclerosis* **86,** 193–199 (1991).

68. Shih, D. M. *et al.* Mice lacking serum paraoxonase are susceptible to organophosphate toxicity and atherosclerosis. *Nature* **394,** 284–287 (1998).

69. Aviram, M. *et al.* Paraoxonase inhibits high-density lipoprotein oxidation and preserves its functions. A possible peroxidative role for paraoxonase. *J. Clin. Invest.* **101,** 1581–

1590 (1998).

70. Kim, D. S., Marsillach, J., Furlong, C. E. & Jarvik, G. P. Pharmacogenetics of paraoxonase activity: elucidating the role of high-density lipoprotein in disease. *Pharmacogenomics* **14,** 1495–1515 (2013).

71. Lövkvist, H. *et al.* Are 25 SNPs from the CARDIoGRAM study associated with ischaemic stroke? *Eur J Neurol* **20,** 1284–1291 (2013).

72. O'Donnell, C. J. *et al.* Genome-wide association study for coronary artery calcification with follow-up in myocardial infarction. *Circulation* **124,** 2855–2864 (2011).

73. Schunkert, H. *et al.* Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat. Genet.* **43,** 333–338 (2011).

74. Traylor, M. *et al.* Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE Collaboration): a meta-analysis of genome-wide association studies. *The Lancet Neurology* **11,** 951–962 (2012).

75. Ikram, M. A. *et al.* Genomewide association studies of stroke. *N Engl J Med* **360,** 1718–1728 (2009).

76. Bis, J. C. *et al.* Meta-analysis of genome-wide association studies from the CHARGE consortium identifies common variants associated with carotid intima media thickness and plaque. *Nat. Genet.* **43,** 940–947 (2011).

77. HPS2-THRIVE Collaborative Group *et al.* Effects of extended-release niacin with laropiprant in high-risk patients. *N Engl J Med* **371,** 203–212 (2014).

78. Hutchins, P. M. *et al.* Quantification of HDL particle concentration by calibrated ion mobility analysis. *Clin. Chem.* **60,** 1393–1401 (2014).

79. Tiozzo, E. *et al.* High-density lipoprotein subfractions and carotid plaque: The Northern

Manhattan Study. *Atherosclerosis* **237,** 163–168 (2014).

80. Jeyarajah, E. J., Cromwell, W. C. & Otvos, J. D. Lipoprotein particle analysis by nuclear magnetic resonance spectroscopy. *Clin Lab Med* **26,** 847–870 (2006).

81. Otvos, J. D. Measurement of lipoprotein subclass profiles by nuclear magnetic resonance spectroscopy. *Clinical laboratory* **48,** 171–180 (2002).

82. Friedewald, W. T., Levy, R. I. & Fredrickson, D. S. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin. Chem.* **18,** 499–502 (1972).

83. O'Leary, D. H. *et al.* Carotid-artery intima and media thickness as a risk factor for myocardial infarction and stroke in older adults. Cardiovascular Health Study Collaborative Research Group. *N Engl J Med* **340,** 14–22 (1999).

84. Mora, S. *et al.* LDL particle subclasses, LDL particle size, and carotid atherosclerosis in the Multi-Ethnic Study of Atherosclerosis (MESA). *Atherosclerosis* **192,** 211–217 (2007).

85. Jarvik, G. P. *et al.* Genetic and nongenetic sources of variation in phospholipid transfer protein activity. *J. Lipid Res.* **51,** 983–990 (2010).

86. Javaheri, A. & Rader, D. J. Apolipoprotein A-I and Cholesterol Efflux: The Good, the Bad, and the Modified. *Circulation Research* **114,** 1681–1683 (2014).

87. Rader, D. J. & Tall, A. R. The not-so-simple HDL story: Is it time to revise the HDL cholesterol hypothesis? *Nat. Med.* **18,** 1344–1346 (2012).

88. Kim, D. S., Hutchins, P. M. & Jarvik, G. P. The HDL particle: frontiers for new discovery in cardioprotection. *Clinical Laboratory International* 16–20 (2014).

89. Asztalos, B. F. *et al.* Distribution of ApoA-I-Containing HDL Subpopulations in

Patients With Coronary Heart Disease. *Arteriosclerosis, Thrombosis, and Vascular Biology* **20,** 2670–2676 (2000).

90. Li, W. F., Costa, L. G. & Furlong, C. E. Serum paraoxonase status: a major factor in determining resistance to organophosphates. *J Toxicol Environ Health* **40,** 337–346 (1993).

91. Riedmaier, S. *et al.* Paraoxonase (PON1 and PON3) polymorphisms: impact on liver expression and atorvastatin-lactone hydrolysis. *Frontiers in Pharmacology* **2,** (2011).

92. Tougou, K. *et al.* Paraoxonase has a major role in the hydrolysis of prulifloxacin (NM441), a prodrug of a new antibacterial agent. *Drug Metabolism and Disposition* **26,** 355–359 (1998).

93. Stevens, R. C. *et al.* Engineered recombinant human paraoxonase 1 (rHuPON1) purified from Escherichia coli protects against organophosphate poisoning. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 12780–12784 (2008).

94. Cheung, M. C., Vaisar, T., Han, X., Heinecke, J. W. & Albers, J. J. Phospholipid Transfer Protein in Human Plasma Associates with Proteins Linked to Immunity and Inflammation. *Biochemistry* **49,** 7314–7322 (2010).

95. Cheung, M. C., Wolfbauer, G. & Albers, J. J. Different phospholipid transfer protein complexes contribute to the variation in plasma PLTP specific activity. *Biochim. Biophys. Acta* **1811,** 343–347 (2011).

96. Albers, J. J., Vuletic, S. & Cheung, M. C. Role of plasma phospholipid transfer protein in lipid and lipoprotein metabolism. *BBA - Molecular and Cell Biology of Lipids* **1821,** 345–357 (2012).

97. Jiang, X. C. *et al.* Apolipoprotein B secretion and atherosclerosis are decreased in mice

with phospholipid-transfer protein deficiency. *Nat. Med.* **7,** 847–852 (2001).

98. Lie, J. *et al.* Elevation of plasma phospholipid transfer protein increases the risk of atherosclerosis despite lower apolipoprotein B-containing lipoproteins. *The Journal of Lipid Research* **45,** 805–811 (2004).

99. Yang, X. P. *et al.* Increased atherosclerotic lesions in apoE mice with plasma phospholipid transfer protein overexpression. *Arteriosclerosis, Thrombosis, and Vascular Biology* **23,** 1601–1607 (2003).

100. Samyn, H. *et al.* Elevation of systemic PLTP, but not macrophage-PLTP, impairs macrophage reverse cholesterol transport in transgenic mice. *Atherosclerosis* **204,** 429–434 (2009).

101. Teslovich, T. M. *et al.* Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466,** 707–713 (2010).

102. Vergeer, M. *et al.* Genetic Variation at the Phospholipid Transfer Protein Locus Affects Its Activity and High-Density Lipoprotein Size and Is a Novel Marker of Cardiovascular Disease Susceptibility. *Circulation* **122,** 470–477 (2010).

103. Liu, R., Hojjati, M. R., Devlin, C. M., Hansen, I. H. & Jiang, X. C. Macrophage Phospholipid Transfer Protein Deficiency and ApoE Secretion: Impact on Mouse Plasma Cholesterol Levels and Atherosclerosis. *Arteriosclerosis, Thrombosis, and Vascular Biology* **27,** 190–196 (2007).

104. Valenta, D. T., Bulgrien, J. J., Bonnet, D. J. & Curtiss, L. K. Macrophage PLTP is atheroprotective in LDLr-deficient mice with systemic PLTP deficiency. *The Journal of Lipid Research* **49,** 24–32 (2008).

105. Valenta, D. T. *et al.* Atheroprotective potential of macrophage-derived phospholipid

transfer protein in low-density lipoprotein receptor-deficient mice is overcome by apolipoprotein AI overexpression. *Arteriosclerosis, Thrombosis, and Vascular Biology* **26,** 1572–1578 (2006).

106. Curtiss, L. K. What Is So Special About Apolipoprotein AI in Reverse Cholesterol Transport? *Arteriosclerosis, Thrombosis, and Vascular Biology* **26,** 12–19 (2006).

107. Robins, S. J., Lyass, A., Brocia, R. W., Massaro, J. M. & Vasan, R. S. Plasma lipid transfer proteins and cardiovascular disease. The Framingham Heart Study. *Atherosclerosis* **228,** 230–236 (2013).

108. Schlitt, A. *et al.* High plasma phospholipid transfer protein levels as a risk factor for coronary artery disease. *Arteriosclerosis, Thrombosis, and Vascular Biology* **23,** 1857–1862 (2003).

109. Schlitt, A. *et al.* PLTP activity is a risk factor for subsequent cardiovascular events in CAD patients under statin therapy: the AtheroGene Study. *The Journal of Lipid Research* **50,** 723–729 (2009).

110. Schgoer, W. *et al.* Low phospholipid transfer protein (PLTP) is a risk factor for peripheral atherosclerosis. *Atherosclerosis* **196,** 219–226 (2008).

111. Jarvik, G. P. *et al.* Vitamin C and E intake is associated with increased paraoxonase activity. *Arteriosclerosis, Thrombosis, and Vascular Biology* **22,** 1329–1333 (2002).

112. Jarvik, G. P. *et al.* Novel paraoxonase (PON1) nonsense and missense mutations predicted by functional genomic assay of PON1 status. *Pharmacogenetics* **13,** 291–295 (2003).

113. Kim, D. S. *et al.* Effects of dietary components on high-density lipoprotein measures in a cohort of 1,566 participants. *Nutrition and Metabolism* **11,** 44 (2014).

114. Howie, B. N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* **5,** e1000529 (2009).

115. 1000 Genomes Project Consortium *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491,** 56–65 (2012).

116. Delaneau, O., Zagury, J.-F. & Marchini, J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat Meth* **10,** 5–6 (2013).

117. Cheung, M. C. *et al.* Phospholipid transfer protein activity is associated with inflammatory markers in patients with cardiovascular disease. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **1762,** 131–137 (2006).

118. Cheung, M. C., Wolfbauer, G. & Albers, J. J. Plasma phospholipid mass transfer rate: relationship to plasma phospholipid and cholesteryl ester transfer activities and lipid parameters. *Biochim. Biophys. Acta* **1303,** 103–110 (1996).

119. Carlson, C. S. *et al.* Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *The American Journal of Human Genetics* **74,** 106–120 (2004).

120. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics* **81,** 559–575 (2007).

121. Ronald, J. *et al.* Analysis of recently identified dyslipidemia alleles reveals two loci that contribute to risk for carotid artery disease. *Lipids Health Dis* **8,** 52 (2009).

122. Cholesterol Treatment Trialists' (CTT) Collaborators *et al.* The effects of lowering LDL cholesterol with statin therapy in people at low risk of vascular disease: meta-analysis of individual data from 27 randomised trials. *Lancet* **380,** 581–590 (2012).

123. de Vries, R. *et al.* Elevated plasma phospholipid transfer protein activity is a determinant of carotid intima-media thickness in type 2 diabetes mellitus. *Diabetologia* **49,** 398–404 (2005).

124. Shih, D. M. *et al.* Combined serum paraoxonase knockout/apolipoprotein E knockout mice exhibit increased lipoprotein oxidation and atherosclerosis. *J. Biol. Chem.* **275,** 17527–17535 (2000).

125. Vuletic, S., Dong, W., Wolfbauer, G., Tang, C. & Albers, J. J. PLTP regulates STAT3 and NFκB in differentiated THP1 cells and human monocyte-derived macrophages. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1813,** 1917–1924 (2011).

126. Vuletic, S. *et al.* SAA and PLTP activity in plasma of periodontal patients before and after full-mouth tooth extraction. *Oral Diseases* **14,** 514–519 (2008).

127. Jänis, M. T. *et al.* Active and low-active forms of serum phospholipid transfer protein in a normal Finnish population sample. *The Journal of Lipid Research* **45,** 2303–2309 (2004).

128. Settasatian, N. *et al.* The Mechanism of the Remodeling of High Density Lipoproteins by Phospholipid Transfer Protein. *Journal of Biological Chemistry* **276,** 26898–26905 (2001).

129. Oram, J. F., Wolfbauer, G., Vaughan, A. M., Tang, C. & Albers, J. J. Phospholipid transfer protein interacts with and stabilizes ATP-binding cassette transporter A1 and enhances cholesterol efflux from cells. *J. Biol. Chem.* **278,** 52379–52385 (2003).

130. Sliwkowski, M. B. & Windmueller, H. G. Rat liver and small intestine produce proapolipoprotein A-I which is slowly processed to apolipoprotein A-I in the circulation.

J. Biol. Chem. **259,** 6459–6465 (1984).

131. Lee-Rueckert, M. *et al.* Absence of endogenous phospholipid transfer protein impairs ABCA1-dependent efflux of cholesterol from macrophage foam cells. *The Journal of Lipid Research* **47,** 1725–1732 (2006).

132. Kaess, B. M. *et al.* Large-Scale Candidate Gene Analysis of HDL Particle Features. *PLoS ONE* **6,** e14529 (2011).

133. Chasman, D. I. *et al.* Forty-Three Loci Associated with Plasma Lipoprotein Size, Concentration, and Cholesterol Content in Genome-Wide Analysis. *PLoS Genet* **5,** e1000730 (2009).

134. Kathiresan, S. *et al.* Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat. Genet.* **41,** 334–341 (2009).

135. Fan, H. *et al.* PCIF1, a novel human WW domain-containing protein, interacts with the phosphorylated RNA polymerase II. *Biochemical and Biophysical Research Communications* **301,** 378–385 (2003).

136. Liu, A., Desai, B. M. & Stoffers, D. A. Identification of PCIF1, a POZ domain protein that inhibits PDX-1 (MODY4) transcriptional activity. *Molecular and Cellular Biology* **24,** 4372–4383 (2004).

137. Brophy, V. H. *et al.* Effects of 5′ regulatory-region polymorphisms on paraoxonase-gene (PON1) expression. *Am. J. Hum. Genet.* **68,** 1428 (2001).

138. Wheeler, J. G., Keavney, B. D., Watkins, H., Collins, R. & Danesh, J. Four paraoxonase gene polymorphisms in 11 212 cases of coronary heart disease and 12 786 controls: meta-analysis of 43 studies. *Lancet* **363,** 689–695 (2004).

139. Wang, M., Lang, X., Zou, L., Huang, S. & Xu, Z. Four genetic polymorphisms of paraoxonase gene and risk of coronary heart disease: a meta-analysis based on 88 case-control studies. *Atherosclerosis* **214,** 377–385 (2011).

140. Eichler, E. E. *et al.* Missing heritability and strategies for finding the underlying causes of complex disease. *Nat Rev Genet* **11,** 446–450 (2010).

141. Tennessen, J. A. *et al.* Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science* **337,** 64–69 (2012).

142. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25,** 1754–1760 (2009).

143. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25,** 2078–2079 (2009).

144. Manichaikul, A. *et al.* Robust relationship inference in genome-wide association studies. *Bioinformatics* **26,** 2867–2873 (2010).

145. Cooper, G. M. *et al.* Distribution and intensity of constraint in mammalian genomic sequence. *Genome Research* **15,** 901–913 (2005).

146. Richter, R. J. & Furlong, C. E. Determination of paraoxonase (PON1) status requires more than genotyping. *Pharmacogenetics* **9,** 745–753 (1999).

147. Eng, J. K., McCormack, A. L. & Yates, J. R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom* **5,** 976–989 (1994).

148. Zheng, X. *et al.* A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **28,** 3326–3328 (2012).

149. Wu, M. C. *et al.* Rare-Variant Association Testing for Sequencing Data with the

Sequence Kernel Association Test. *Am. J. Hum. Genet.* **89,** 82–93 (2011).

150. Skol, A. D., Scott, L. J., Abecasis, G. R. & Boehnke, M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat. Genet.* **38,** 209–213 (2006).

151. Schwamm, L. H. *et al.* Race/ethnicity, quality of care, and outcomes in ischemic stroke. *Circulation* **121,** 1492–1501 (2010).

152. Zuk, O., Hechter, E., Sunyaev, S. R. & Lander, E. S. The mystery of missing heritability: genetic interactions create phantom heritability. *Proceedings of the National Academy of Sciences* **109,** 1193–1198 (2012).

153. Matyus, S. P. *et al.* HDL particle number measured on the Vantera®, the first clinical NMR analyzer. *Clinical Biochemistry* **48,** 148–155 (2015).

154. Mackey, R. H. *et al.* High-Density Lipoprotein Cholesterol and Particle Concentrations, Carotid Atherosclerosis, and Coronary Events. *Journal of the American College of Cardiology* **60,** 508–516 (2012).

155. Mackey, R. H. *et al.* Lipoprotein particles and size, total and high molecular weight adiponectin, and leptin in relation to incident coronary heart disease among severely obese postmenopausal women: The Women's Health Initiative Observational Study. *BBA Clin* **3,** 243–250 (2015).

156. Mora, S. *et al.* Atherogenic Lipoprotein Subfractions Determined by Ion Mobility and First Cardiovascular Events After Random Allocation to High-Intensity Statin or Placebo: The JUPITER Trial. *Circulation* (2015). doi:10.1161/CIRCULATIONAHA.115.016857