SEQUENCE VARIATION IN THE *APOA1* AND *APOA4* GENES AND THEIR RELATIONSHIP WITH PLASMA HDL-CHOLESTEROL LEVELS

by

Sarah Elizabeth Hill

B.S., Arizona State University, 2006

Submitted to the Graduate Faculty of

the Department of Human Genetics, Genetic Counseling

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2009

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This thesis was presented

by

Sarah Elizabeth Hill

It was defended on

April 1, 2009

and approved by

Thesis Advisor:
M. Ilyas Kamboh, Ph.D.
Professor and Department Chair
Department of Human Genetics
Graduate School of Public Health
University of Pittsburgh

Committee Member:
F. Yesim Demirci, M.D.
Research Assistant Professor
Department of Human Genetics
Graduate School of Public Health
University of Pittsburgh

Committee Member:
David N. Finegold, M.D.
Professor
Department of Pediatrics and Department of Human Genetics
School of Medicine and Graduate School of Public Health
University of Pittsburgh

Copyright © by Sarah Elizabeth Hill

2009

SEQUENCE VARIATION IN THE *APOA1* AND *APOA4* GENES AND THEIR RELATIONSHIP WITH PLASMA HDL-CHOLESTEROL LEVELS

Sarah Elizabeth Hill, M.S.

University of Pittsburgh, 2009

Heart disease continues to be the leading cause of death in the United States, making it one of the foremost public health concerns. Many factors influence the risk to develop heart disease, including abnormal blood lipid levels. High levels of plasma high-density lipoprotein (HDL)-cholesterol have been shown to have a protective effect. Recent genome-wide association studies (GWAS) and candidate gene studies have identified genes thought to contribute to HDL-cholesterol levels. Two genes, *APOA1* and *APOA4,* have been associated with HDL-cholesterol levels in multiple studies with inconsistent results. The majority of these studies focused on the "common variant-common disease" hypothesis whereas only one study by Cohen *et al.* (2004) evaluated *APOA1* using the "rare variant-common disease" hypothesis. The aim of this study was to further investigate the role of common and rare variation in these two genes by sequencing individuals having extremely low and high HDL-cholesterol levels in two populations, U.S. Non-Hispanic Whites (NHWs), and African Blacks, and then screening the identified variants in the entire sample. In the initial sequence analysis, 54 variants were identified in *APOA1* (25 of which were new), and 43 in *APOA4* (21 of which were new). According to preliminary analysis of the sequencing data for *APOA1* and *APOA4*, no striking difference was noticed between the distribution of rare variants between high and low HDL groups in either population. To date, screening data was compiled for the entire NHWs and

Black samples for a total of seven common variants: 2 for *APOA1* (rs5070 and rs5072), and 5 in *APOA4* (rs5092, rs5100, rs5104, rs5106, and rs5109). All 7 variants were present in the Black population; five were present in NHWs (rs5070, rs5072, rs5092, rs5100, and rs5104). Modest or marginal significant p-values were observed; however, none would maintain significance after multiple testing correction in either population. Additional variants identified in sequencing remain to be screened in the entire NHWs and Black samples.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

First and foremost I would like to thank Dr. M. Ilyas Kamboh for providing me the opportunity to participate in this exciting research. I have learned so much from working on this project, and I appreciate you allowing me to be a part of it. I am also deeply in debt to Dr. F. Yesim Demirci for her tireless effort teaching and guiding me every step of the way. I have enjoyed working with you very much, and I thank you for being a mentor and role model to me these last two years. I would like to thank Dr. David Finegold who has always been an inspiration to me and showed me what it meant to truly love the science of genetics. Thank you for being such an enthusiastic teacher as well as a support system for me whenever I have needed it. I would also like to thank Amy Dressen for her assistance on the statistical analysis of this project, and for being a great ASHG roommate!

Thank you to my coworkers and fellow students in Dr. Kamboh's lab who I have enjoyed working with over these last few years. I would like to especially thank Sally Hollister who has helped to guide me through my project after doing similar work of her own with *APOA2*. I would also like to thank Ryan Minster for always answering my questions and guiding me, Fahad Waqar and Yuee Wang for their help in my training, and Jessica Figgins for her friendship and guidance.

I would also like to thank the co-directors of the Genetic Counseling Program, Ms. Elizabeth Gettig and Dr. Robin Grubs, for all that they have taught me about genetics as well as their infinite moral support throughout this project. I would also like to thank all of my

classmates in the Genetic Counseling Program. The nine of you are the most amazing and inspiring individuals I have ever met, and showed me what true friendship is all about.

Thank you to my Pittsburgh family, Cecilia Rajakaruna, Cherise Klotz, and Sully, for your love and support, especially in these last few months. Cecilia, you never fail to take care of me when I need it most. Cherise, you are my other half, I don't know what I would do without you! Thank you also to Stefanie Frace, who has been like a third roommate and sister to me.

I would also like to thank my mom and dad and sister, Rachel. I love you all very much. Thank you for always supporting my educational pursuits, and being there for me when I need it most. Mom, thank you for being my confidant and advisor. Dad, thank you for being my role model and support system. Rachel, you are my best friend, I am sorry I have missed out on so much of your growing up these last few years.

1.0 BACKGROUND AND SIGNIFICANCE

1.1 HIGH DENSITY LIPOPROTEIN (HDL)

1.1.1 The HDL Particle

High-density lipoproteins (HDL) are spherical particles that transport cholesterol esters and triglycerides in the blood. The HDL particle represents one class of lipoprotein; the classes of lipoproteins (chylomicrons, very-low-density lipoproteins, and low-density lipoproteins) are separated by density, size, and protein content. The core of the HDL particle is made up of cholesterol esters and triglycerides, which is encapsulated in an amphipathic layer containing free (unesterified) cholesterol and phospholipids (Figure 1). HDL particles also contain several proteins, specifically apolipoproteins (apos), which have many roles: they provide structural integrity, have enzymatic co-activator functions, are involved in the assembly and secretion of the HDL particle, and serve as a ligand for a variety of receptors.[1]

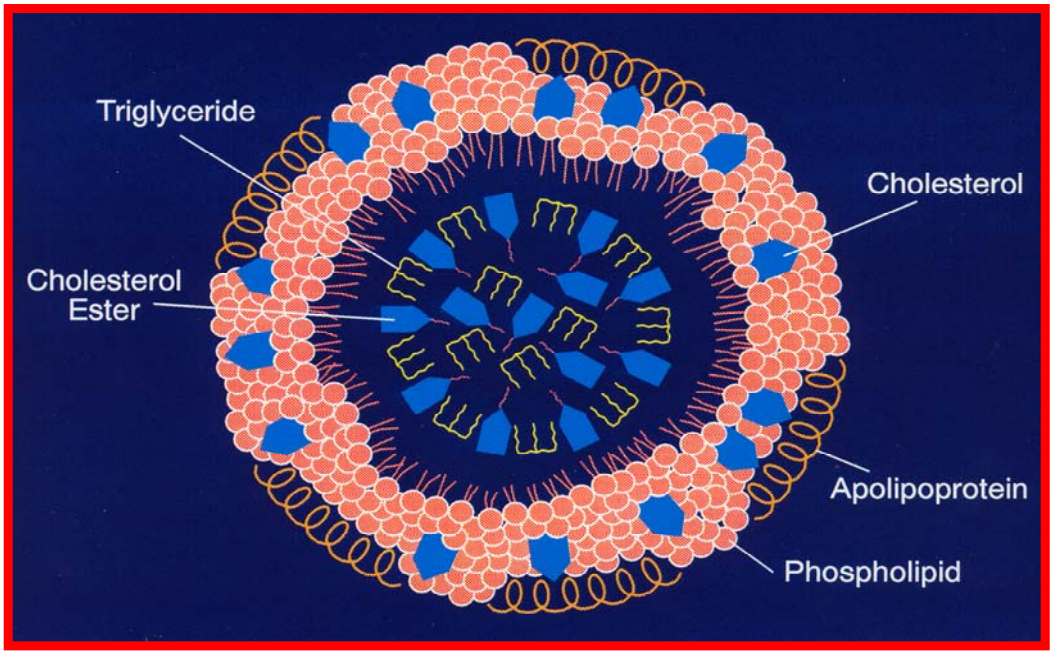


Figure 1. The HDL Particle. *Image courtesy of M. Ilyas Kamboh, Ph.D.*

1.1.2 HDL-Cholesterol Metabolism

HDL-cholesterol metabolism (Figure 2) mediates reverse cholesterol transport from the peripheral cells to the liver, where cholesterol can be excreted into the bile, or the adrenal glands, ovaries, or testes, where cholesterol can be converted into steroid hormones. HDL particles, initially containing only phospholipids and apolipoprotein A-I (apoA-I), are synthesized in the liver and intestine. They enter into the blood stream and accumulate cholesterol esters, which are converted from free cholesterols by lecithin-cholesterol acyltransferase (LCAT), through interaction with apoA-I and ATP binding cassette transporter A1 (ABCA1). The HDL particles increase in size as they move through the bloodstream accumulating cholesterol esters, and are remodeled by cholesterol ester transfer protein (CETP) and endothelial lipase (LIPG).

Eventually, the HDL particles return to the liver, where HDL removal is mediated by the scavenger receptor B1 (SRB1)[2].

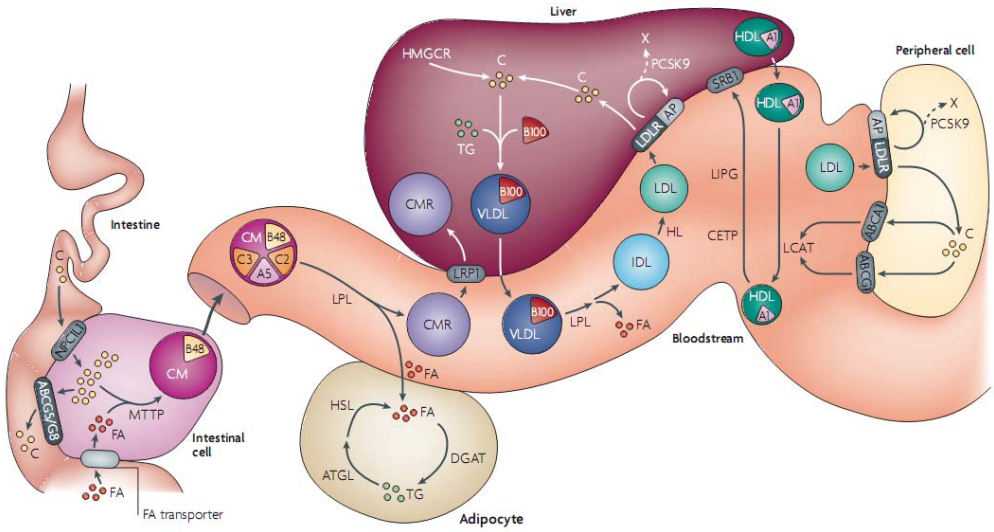


Figure 2. Overview of Lipoprotein Metabolism (Hegele, 2009)[2]

1.2 ATHEROSCLEROSIS AND HEART DISEASE

1.2.1 The Public Health Impact of Heart Disease

Coronary heart disease (CHD) is cause of one in five deaths in the United States, and diseases of the heart were the leading cause of death in 2005 according to the Center for Disease Control National Vital Statistics Reports.[3] Nearly 2,400 American die of cardiovascular disease (CVD) each day. One in three American adults, greater than eighty million individuals, have one or more types of CVD including: CHD, high blood pressure, heart failure, and stroke. Of these approximately 16 million are affected by CHD. The estimated health care cost of CVD for 2008 was $448.5 billion.[4]

1.2.2 Risk Factors for Heart Disease

Many factors influence the risk to develop heart disease. Some of the major risk factors for CHD include: abnormal blood lipid levels (or current treatment with cholesterol-lowering drugs), hypertension (or current treatment with blood pressure-lowering drugs), diabetes, abdominal obesity, cigarette smoking, a lack of physical activity, low daily fruit and vegetable consumption, and alcohol over consumption.[5,6]

1.2.3 Cholesterol Levels and Heart Disease

The American Heart Association classifies HDL-cholesterol levels of <40mg/dL for men and <50mg/dL for women as low, and considers low HDL-cholesterol to be a major risk factor for heart disease (www.americanheart.org). In addition to low HDL-cholesterol levels, high levels of low-density lipoprotein (LDL) cholesterol and total cholesterol have also been shown to increase the risk for heart disease. LDL cholesterol level of >160-189mg/dL, or a total cholesterol level of >240mg/dL is considered high.

1.2.4 Atherosclerosis and Heart Disease

The blood vessels form a system of tubes that carry blood away from the heart, through the tissues of the body, and back to the heart. The arteries are the vessels through which blood is pumped away from the heart. Atherosclerosis, commonly referred to as hardening of arteries, is an inflammatory response in the artery walls caused by the formation of plaques within the arteries. Plaques buildup in the arteries over a long period of time causing artery enlargement,

and atherosclerosis is generally asymptomatic for decades. Eventually plaques can rupture leading to stenosis of the artery or blood clot formation which blocks blood flow to the heart causing a heart attack.

1.2.5 Epidemiological Evidence for the Antiatherogenic Properties of HDL-Cholesterol

There are many different hypotheses for the biological basis of the atheroprotective role of HDL-cholesterol, including the ability of HDL to promote cholesterol efflux, as well as the antioxidant and anti-inflammatory properties of the lipid particle.[7] HDL particles have been shown to have antiatherogenic properties and HDL-cholesterol concentrations have been inversely correlated with the risk for coronary artery disease (CAD) in many studies.[8-10] The Framingham Heart Study illustrated this inverse relationship: a 1% increase in HDL-cholesterol was associated with a 2% reduction in the development of CAD.[9] A study by Gordon *et al.*[10] supported this inverse relationship as well; an increase of 1mg/dL in HDL-cholesterol levels is associated with a 2% decrease in the risk for CAD in men and a 3% decrease in the risk for CAD in women. A strong negative association has also been shown with ischemic heart disease mortality; in one meta analysis of 900,000 adults an average of approximately 13mg/dL higher HDL-cholesterol was correlated with a one third lower risk for ischemic heart disease mortality in men and women within every age group.[8] Individuals with decreased HDL-cholesterol levels have been shown to be at greater risk for heart disease; a HDL-cholesterol level <35mg/dL was associated with a 3 fold risk for CHD in one study.[11]

1.3 GENETIC STUDIES OF HDL-CHOLESTEROL LEVELS

Abnormal lipid levels are a major risk factor for heart disease. HDL-cholesterol levels have been shown to be under a considerable amount of genetic control, with heritability estimates of up to 80% and an average heritability estimate of 40-60%.[12-15] Research over the past 25 years has focused on determining the genetic basis underlying variation in HDL-cholesterol levels.

1.3.1 Candidate Gene Studies

The genes that encode the proteins responsible for HDL metabolism, including apos, cellular receptors, and enzymes are critical to HDL synthesis, processing, and catabolism. Through the elucidation of the biochemical pathway responsible for HDL metabolism candidate genes are identified for study. Numerous studies have been carried out over the last 25 years in an attempt to correlate variation in these genes with HDL-cholesterol levels.[16] Candidate gene studies of genetic polymorphisms in the genes encoding lipoprotein lipase (LPL), the major triglyceride-hydrolyzing enzyme, and apolipoprotein A-I (apoA-I), the major protein of HDL-cholesterol, have been correlated with HDL-colesterol levels with inconsistent results.[17]

1.3.2 Genome Wide Association Studies (GWAS)

GWAS utilize single nucleotide polymorphism (SNP) chip technology and a case-control study design to identify genes associated with a particular phenotype. Multiple GWAS have been carried out and shown statistically significant associations between variation in HDL-cholesterol levels and the following genes: cholesteryl ester transfer protein *(CETP),* lipoprotein lipase

(LPL), hepatic lipase *(LIPC),* endothelial lipase *(LIPG), ABCA1, LCAT,* the apolipoproteinA-I/C-3/A-IV/A-V gene cluster *(APOA1/C3/A4/A5),* apolipoprotein B *(APOB),* CCCTC-binding factor *(CTCF),* protein arginine N-methyltransferase 8 *(PRMT8),* MAP kinase-activating death domain *(MADD),* folate hydrolase 1 *(FOLH1),* acetylgalactosaminyltransferase 2 *(GALNT2),* mevalonate kinase *(MVK),* cob(I)alamin adenosyltransferase *(MMAB),* cleft lip- and palate-associated transmembrane protein 1 *(CLPTM1),* glutamate receptor, iontropic, N-methyl-D-asparate 3A *(GRIN3A)*, and nuclear receptor subfamily 1, group H, member 3 *(NR1H3)*. [18-25]

1.3.3 Genetic Models for HLD Variation

1.3.3.1 Common Variant-Common Disease Hypothesis

In the context of a Mendelian disorder a single gene can have a profound impact on a disease. This is exemplified in the case of familial hypercholesterolemia (FH), in which individuals heterozygous and homozygous for loss of function mutations in the low density lipoprotein receptor *(LDLR)* gene develop premature atherosclerosis.[26] In the context of complex disease, however, the effect of variation in most single gene candidates is small. One proposed model for genetic variation in HDL-cholesterol levels is the theory that many small effects of multiple common variants aggregate in an individual to produce disease susceptibility in common disease.[16]

1.3.3.2 Rare Variant-Common Disease Hypothesis

Another proposed model for genetic variation in HDL-cholesterol levels that is gaining increasing support is the theory that a portion of individuals in the population, those at the extremes of the Gaussian distribution, carry dysfunctional variants in genes that have more profound effects.[16] A study by Cohen *et al.*[27] established a paradigm that multiple rare alleles with major phenotypic effects underlie a substantial minority of cases of decreased HDL-cholesterol levels in the general population. This multiple rare variants model has also been used in studies looking at LDL-cholesterol levels and triglyceride levels.[28-30] The methodology used in this study, known as the 'missense-accumulation' analysis, is outlined in Figure 3.

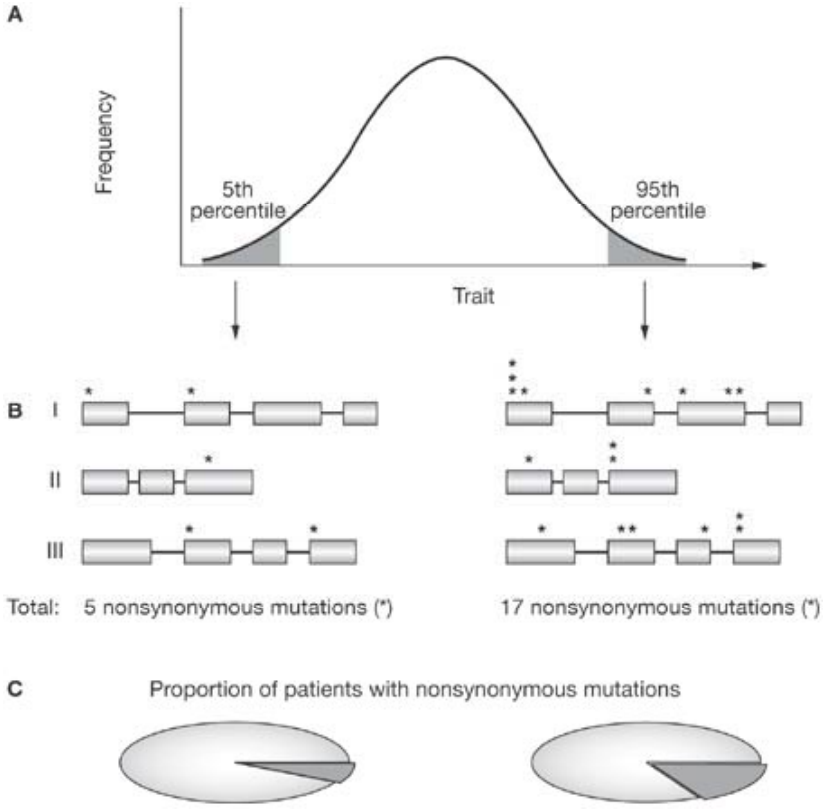


Figure 3. **'**Missense-Accumulation' Analysis. The frequency distribution of HDL-cholesterol levels is shown at the top of the figure (A). Individuals at the extremes (5th and 95th percentile) are chosen for DNA sequencing, focusing on candidate genes identified based on their biological function or in GWAS—APOA1, LCAT, ABCA1 in the case of Cohen *et al.*[27] (B). Individual samples are sequenced and the frequency of the identified variants between the two groups is compared (C). (Pollex *et al.*, 2007)[16]

1.4 APOLIPOPROTEIN A-I: THE *APOA1* GENE

The apolipoproteinA-I (apoA-I protein; *APOA1* gene) has been mapped to chromosome 11q23 in humans. The National Center for Biotechnology Information (NCBI) reference nucleotide sequence is NC_000011.8 (http://www.ncbi.nlm.gov/sites/entrez). *APOA1* gene has four exons and three introns; the lengths of the exons are: 18bp, 63bp, 157bp, 659bp, respectively (Figure 4). The mRNA nucleotide sequence is 804nt in length (NCBI mRNA locus NM_000039.1).

Figure 4. The *APOA1* Gene. (http://www.ncbi.nlm.gov/sites/entrez).

APOA1 encodes a protein, apoA-I; the coding sequence for the apoA-I protein begins in exon 2 (NCBI reference protein sequence NP_000030). ApoA-I is the major apolipoprotein of HDL.[31,32] ApoA-I is also a cofactor for LCAT, which converts free cholesterol into cholesterol ester.

1.4.1 Protein Structure

The ApoA-I protein is a single polypeptide containing 243 amino acid residues.[31] It is synthesized as a preprotein (NCBI preprotein reference sequence NP_000030.1) that undergoes proteolytic processing to form the mature protein that is present in blood.[33] Based on the amino acid sequence, the secondary structure is hypothesized to consist of repeating amphipathic

helicies of 11 or 22 amino acids in length separated by proline residues.[34] The crystal structure of apoA-I has been determined and is shown in Figure 5. The overall structure consists of two main helical domains, one in the N-terminal region containing a four-helix antiparallel bundle, and another in the C-terminal region containing a two-helix bundle.[35]

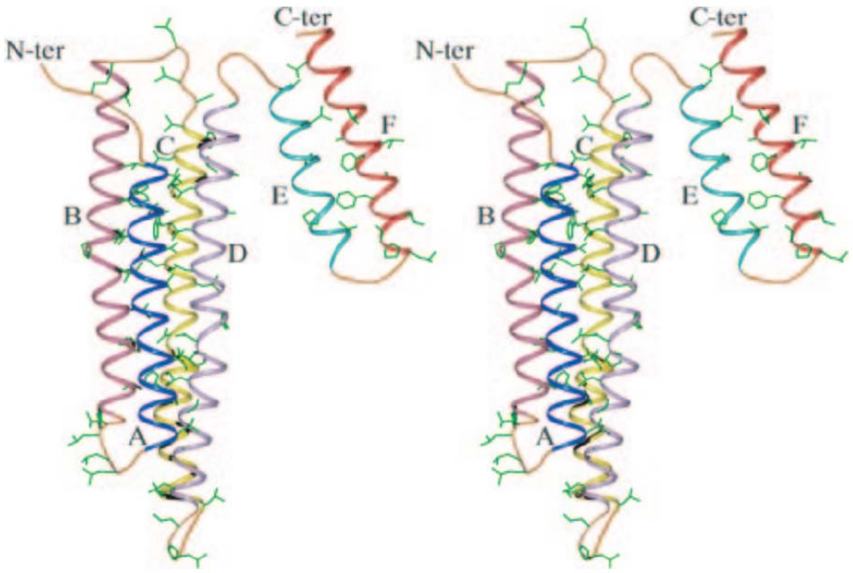


Figure 5. Crystal structure of apoA-I. The six helices in the structure are colored blue (A), pink (B), yellow (C), lavender (D), cyan (E), and red (F) and labeled. Loops are colored gold, and hydrophobic residues are shown as green sticks. (Ajees *et al.*, 2006)[35]

1.4.2 Functional Considerations

Apoa-I is the major apolipoprotein in HDL particles; it is hypothesized to have a protective effect against the development of CAD via promoting efflux of cholesterol from cells and modulating immune cell activation.[36-41] In mice, apoA-I deficiency has been correlated with atherosclerosis, and an over expression of apoA-I was shown to be atheroprotective.[42,43] ApoA-I also has proposed anti-inflammatory properties providing further evidence for its

atheroprotective role. These anti-inflammatory properties have been illustrated best in studies of D-4F, an apoA-I mimetic. Mice and monkeys given oral doses of D-4F have been shown to undergo a marked decrease in atherosclerotic lesions; clinical trials of D-4F safety and efficacy in human subjects are underway.[44]

1.4.3 *APOA1* Variants and Phenotypic Association

Studies have found a statistically significant correlation between variants in the *APOA1* gene and HDL-cholesterol levels.[45-47] While other studies have been less successful in correlating variation in this gene with a clinical phenotype.[48] Decreased levels of apoA-I protein have also shown to be an independent risk factor for CAD, leading to the conclusion that apoA-I plays an important role in the atherogenic process, even in patients with no other risk factors for heart disease.[48] The heretibility of apoA-I levels have been estimated as high as 90% in multiple studies.[49]

A variety of specific sequence variants have been identified in the *APOA1* gene and correlated with Mendelian disorders. NCBI Online Mendelian Inheritance in Man (OMIM) lists 26 rare allelic variants that are associated with different Mendelian disorders (Table 1) (http://www.ncbi.nlm.nih.gov/sites/entrez?db=omim).

Table 1. Allelic Variants in APOA1

Variant	Amino Acid Change	dbSNP	Clinical Phenotype
ApoA-I (Milano)[50]	ARG173CYS	rs28931573	Hypertriglyceridemia and decrease in HDL levels without clinical signs of atherosclerosis[51]
ApoA-I (Marburg)	LYS107TER		Hypertriglyceridemia and decrease in HDL levels[52]
ApoA-I (Munster4)[53]	GLU198LYS		No relationship to premature atherosclerosis[54]
ApoA-I (Norway)[53]	GLU136LYS		
ApoA-I (Giessen)[52]	PRO143ARG		
ApoA-I (Munster3C)[53]	PRO3ARG		
ApoA-I (Munster3B)[53]	PRO4ARG		
ApoA-I	PRO165ARG		Decrease in HDL and ApoA-I levels[55]
Iowa or Van Allen type Amyloid Polyneuropathy-Nephropathy[56,57]	GLY26ARG	rs28931574	Autosomal dominant early onset amyloidosis and neuropathy and variable onset nephropathy with peptic ulcer, cataracts, and hearing loss[56]
Combined Deficiency of ApoA-I and ApoC-III; Detroit type HDL Deficiency	APOA1/APOC 3 FUSION		Autosomal recessive very low HDL and heart failure from CAD with arcus cornealis and xanthoma[58]
Absence of ApoA-I due to Deletion of APOA1/APOC3/APOA4 Gene Complex	APOA1 DELETION		Heterozygotes demonstrate decrease in HDL and apoA-I with CAD; no detectable apoA-I, very low HDL, reduced ApoB/C, CAD, corneal clouding, and diffuse lipid deposits in the epithelium seen in homozygous individual[59]
ApoA-I (Baltimore)	ARG10LEU	rs28929476	Decrease in apoA-I levels (linkage not demonstrated)[60]
Corneal Clouding due to ApoA-I Deficiency	1-BP DELETION CODON 202		Corneal clouding in a homozygous individual[61]
ApoA-I Deficiency	GLN84TER		ApoA-I deficiency and premature atherosclerosis in a homozygous individual [62]
Systemic Nonneuropathic Amyloidosis	LEU60ARG		Autosomal dominant nonneuropathic systemic amyloidosis[63]
Analphalipoproteinemia	GLN2TER		Very low HDL-cholesterol,

Table 1 (Continued)

			undetectable apoA-I, xanthelasmata, cataracts, and cerebellar ataxia in a homozygous individual[64]
Primary Hypoalphalipoproteinemia	1-BP INSERTION IN CODON 325		Autosomal dominant decreased HDL-cholesterol and apoA-I[65]
Periorbital Xanthlasmas	GLN32TER		Periorbital xanthlasmas without CAD or atherosclerosis in a homozygous individual[66]
Hepatic and Systemic Amyloidosis	12-BP DELETION AND 2-BP INSERTION IN EXON 4		Autosomal dominant nonneuropathic amyloidosis with a unique hepatic presentation and death from liver failure[67]
Systemic Nonneuropathic Amyloidosis	TRP50ARG		Hereditary amyloidosis[67]
ApoA-I (Oita)	VAL156GLU		Less than 10% normal HDL and apoA-I, CAD, and corneal opacities in a homozygous individual[68]
Primary Hypoalphalipoproteinemia	DONOR SPLICE SITE MUTATION IN INTRON 2 G-C, +1		Primary Hypoalphalipoproteinemia[69]
Cardiac and Cutaneous Amyloidosis	LEU90PRO		Autosomal dominant hereditary amyloidosis with unique cutaneous and cardiac presentation and death from heart failure[70]
Cardiac and Cutaneous Amyloidosis	ARG173PRO		Hereditary amyloidosis that showed expression mainly in the skin and heart[71]
Systemic Nonneuropathic Amyloidosis	LEU174SER		Amyloid deposits mainly in the heart[72]
Systemic Nonneuropathic Amyloidosis	ALA175PRO		Renal amyloidosis with renal failure, sterility, and hoarseness due to laryngeal amyloid deposits[73]

Additionally, Pisciotta *et al.*[74] reported two siblings with HDL deficiency, no plasma apoA-I, corneal opacities, and planar xanthomas who were homozygous for a deletion in exon 3 (c.85 del C) leading to a premature termination codon; one sibling also had premature CAD. This mutation was also reported in unrelated individuals, some of which were heterozygous, while others were compound heterozygous for other mutations in *APOA1*. A novel mutation in *APOA1* was also reported by Hovingh *et al.*[75], a C>T point mutation at nucleotide 643 in exon 4, predicting the exchange of a leucine for a proline at codon 178. This change was correlated with low levels of apoA-I and HDL in Caucasian Dutch heterozygotes. The heterozygous individuals also had endothelial dysfunction, and statistically significant increased arterial wall thickness and increased rates of premature artery disease as compared to their unaffected siblings. Another new mutation in *APOA1,* leading to severe HDL-cholesterol deficiency in a group of 54 unrelated French Canadian subjects, was reported by Dastani *et al.*[76] The novel mutation in this population was a G>T point mutation at nucleotide 478 in exon 4, leading to a substitution of glutamic acid for a stop codon. In the study, five out of nine carriers over the age of 35 had developed CAD. Esperon *et al.*[77] recently reported a 2006G>C point mutation in exon 4 leading to an arginine to proline substitution in codon 153 in a family with a strong history of premature CAD. They named this variant ApoA-$I_{Montevideo}$.

Studies have been done looking at common and rare variants that contribute to complex disease. Thirty SNPs in the *APOA1* gene, plus the insertion/deletion polymorphism have been reported by Fullerton *et al.*[78] and summarized in the SeattleSNPs database (http://pga.mbt.washington.edu). Tables 2, 3, and 4 below summarize the variants found in each of the three populations: Jackson, MS, North Karelia Finland, and Rochester, MN, respectively.

Table 2. Allelic Variants in the Jackson, MS Population (25 total)

Site	rs Number	nt Change	Minor Allele Frequency (MAF)
206	rs7123454	A>C	0.50
631	rs7948159	A>G	0.35
1049	rs1263162	T>A	0.24
1128	rs11216153	G>T	0.17
1308	rs12721030	C>T	0.06
1407	rs127211027	ins	0.03
1541	rs127211029	C>T	0.03
1546	rs525028	G>A	0.23
1620	rs12721028	A>G	0.45
1749	rs12718462	T>C	0.05
2077	rs12721025	G>A	0.04
2198	rs12721026	T>G	0.04
2373	rs12718463	T>C	0.42
2376	rs5081	A>T	0.23
3220	rs5076	G>A	0.27
3368	rs7116797	G>A	0.31
3431	rs12718464	G>A	0.04
3543	rs5073	C>T	0.12
3766	rs12718465	C>T	0.10
4050	rs5070	A>G	0.41
4245	rs12721032	G>A	0.02
4284	rs5069	G>A	0.27
4443	rs670	C>T	0.17
4732	rs12718467	C>A	0.04
4807	rs12691374	C>T	0.11

Table 3. Allelic Variants in the North Karelia, Finland Population (20 total)

Site	rs Number	nt Change	Minor Allele Frequency (MAF)
206	rs7123454	A>C	0.25
533	rs12721031	C>T	0.08
1128	rs11216153	G>T	0.10
1308	rs12721030	C>T	0.17
1546	rs525028	A>G	0.35
1598	rs10750098	T>G	0.18
1620	rs12721028	A>G	0.10
1749	rs12718462	T>C	0.08
2077	rs12721025	G>A	0.08
2198	rs12721026	T>C	0.09
2373	rs12718463	T>C	0.04
3220	rs5076	G>A	0.06
3368	rs7116797	G>A	0.23
3431	rs12718464	G>A	0.08
3613	rs5072	G>A	0.17
3714	rs2070665	G>A	0.17
4050	rs5070	G>A	0.35
4284	rs5069	G>A	0.06
4443	rs670	C>T	0.10
4693	rs12718466	T>G	0.06

Table 4. Allelic Variants in the Rochester, MN Population (25 total)

Site	rs Number	nt Change	Minor Allele Frequency (MAF)
206	rs7123454	A>C	0.08
533	rs12721031	C>T	0.02
1049	rs1263162	T>A	0.02
1128	rs11216153	G>T	0.31
1308	rs12721030	C>T	0.33
1407	rs127211027	ins	0.02
1546	rs525028	A>G	0.40
1598	rs10750098	T>G	0.06
1620	rs12721028	A>G	0.28
1717	rs12718461	G>C	0.02
1749	rs12718462	T>C	0.02
2077	rs12721025	G>A	0.02
2198	rs12721026	T>G	0.02
2373	rs12718463	T>C	0.02
2376	rs5081	A>T	0.02
3220	rs5076	G>A	0.02
3368	rs7116797	G>A	0.08
3431	rs12718464	G>A	0.04
3613	rs5072	G>A	0.06
3714	rs2070665	G>A	0.06
3766	rs12718465	C>T	0.09
4050	rs5070	G>A	0.35
4284	rs5069	G>A	0.02
4443	rs670	C>T	0.31
4693	rs12718466	T>G	0.02

Both Brown *et al.*[45] and Shioji *et al.*[47] previously identified the T>C change (rs5070); Shioji *et al.*[47] correlated this change with a statistically significant increase in HDL-cholesterol levels in an Japanese population, but Brown *et al.*[45] did not see this same association in a multi-ethnic population. Resequencing of the *APOA1* gene in Brown *et al.*[45] identified one variant in *APOA1* (rs5076) that was statistically significant in European-American males and had a consistent genotype-phenotype relationship across all populations and gender subgroups. However, Brown

et al.[45] did not see any statistically significant correlation with *APOA1* SNPs (rs5069, rs2070665, and rs2073) and HDL levels in a multi-ethnic population of Caucasians and African-Americans. Knoblauch *et al.*[79] did not see an association between *APOA1* variants (rs525028, rs5081, rs5070, rs1799837, and rs5069) and HDL levels.

1.5 APOLIPOPROTEIN A-IV AND THE *APOA4* GENE

The apolipoproteinA-IV (apoA-IV protein; *APOA4* gene) has been mapped to chromosome 11q23 in humans (NCBI reference nucleotide sequence NC_000011.8). *APOA4* gene has three exons and two introns; the length of the exons are: 153bp, 127bp, and 1180bp, respectively (Figure 6). The mRNA nucleotide sequence is 1191nt in length (NCBI mRNA locus NM_000482.3).

Figure 6. The *APOA4* Gene (http://www.ncbi.nlm.gov/sites/entrez).

APOA4 encodes a protein, apoA-IV, that is 396 amino acids in length and has two major isoforms (The NCBI reference protein sequences are AAI13597 and AAI13595). It is synthesized as a preprotein (NCBI preprotein reference sequence NP_000473.2) and undergoes post-translations modifications. While the exact function of apoA-IV is not known, it has a number of proposed functions including involvement in the assembly and secretion of chylomicrons and the reverse cholesterol transport system.

1.5.1 Structure

The crystalline structure of ApoA-IV has yet to be determined. A three-dimension homology model of the protein has been proposed, and studies have looked at the structure of the protein and the possible functional implications.[80,81]

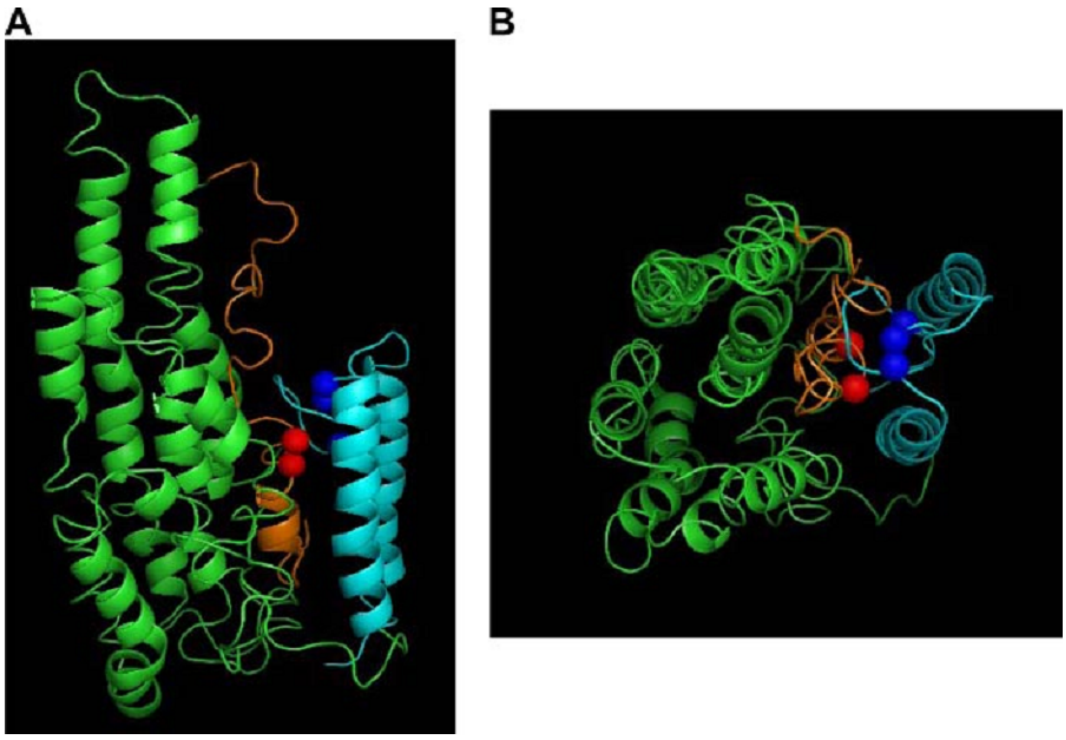


Figure 7. Homology Model of apoA-IV. The majority of the protein is colored green. The N-terminal 39 amino acids, encoded by a separate exon, are colored orange with the amino acids Trp^{12} and Phe^{15} shown as red spheres. The C-terminal 66 amino acids are colored light blue, with residues Phe^{334}, $Phe^{335}$and Phe^{338} shown as blue spheres. (Tubb, 2008) [81]

1.5.2 *APOA4* Variants and Phenotypic Association

Many allelic variants in *APOA4* have been reported and correlated with a clinical phenotype. In 1990, Lohse *et al.*[82] reported the molecular basis of a common protein polyprophism (APOA4*1 and APOA4*2), and demonstrated that a G>T nucleotide substitution leads to the conversion of a glutamine to a histidine at codon 360. This change has been categorized further in many studies. In the Framingham Offspring Study of 2322 Caucasian men and women, Cendoroglo *et al*.[83] examined the effect of the APOA4 (G>T) polymorphism on plasma lipid and lipoprotein levels. No significant allele effect of the was observed on HDL-cholesterol levels, other lipid levels, or Lp(a) levels.

Another common variant, with a reported allele frequency of 0.20-0.25, is a A>T nucleotide substitution which leads to a conversion of a threonine to serine at codon 347.[84] One study of 2808 healthy individuals correlated this variant with a decreased plasma apoA-IV levels and an increased risk for CHD.[85] Another study correlated this variant with an increased risk for cardiovascular disease in individuals with diabetes.[86] However, a third study reported no association between this variant and hyperlipidemia in otherwise healthy individuals.[87]

Other more rare variants have also been reported in the literature. Lohse *et al.*[88] reported a four amino acid insertion (Glu-Gln-Gln-Gln) between codons 361 and 362 which was termed APOA4*0. Lohse *et al.*[88] also reported a G>A nucleotide substitution that converted glutamic acid to lysine at codon 230, termed APOA4*3. In another study, Lohse *et al.*[89] reported three novel variants in APOA4: an A>T point mutation at nucleotide 2346 in exon 3 causing a Thr347Ser amino acid substitution, an A>G point mutation at nucleotide 1806 causing a Lys167Glu amino acid substitution (this was in cis with an APOA4*2 variant), and a G>A point mutation at nucleotide 1800 in exon 3 causing a Glu165Lys amino acid substitution. Deeb *et*

al.[90] also reported three novel variants in *APOA4* in individuals with familial combined hyperlipidemia: a C>T causing a S158L amino acid substitution (termed Seattle-1), a G>A change causing a R244Q amino acid subsitution (termed Seattle-2), and a G>T change causing a A141S amino acid substitution (termed Seattle-3). Knoblauch *et al.*[79] did not see an association between *APOA4* variants (rs675, rs5104, rs5092, and rs2542051) and HDL levels.

Thirty SNPs in the APOA4 gene, plus the deletion polymorphism have been reported by Fullerton *et al.*[78] and summarized in the SeattleSNPs database (http://pga.mbt.washington.edu). Tables 5, 6, and 7 below summarize the variants found in each of the three populations: Jackson, MS, North Karelia Finland, and Rochester, MN, respectively.

Table 5. Allelic Variants in the Jackson, MS Population (18 total)

Site	rs Number	nt Change	Minor Allele Frequency (MAF)
165	rs9282602	DEL	0.02
315	rs675	T>A	0.19
406	rs5109	C>A	0.11
568	rs5106	G>A	0.05
974	rs5104	T>C	0.06
1183	rs12721042	C>A	0.02
1198	rs5101	G>A	0.23
1334	rs5100	A>G	0.35
1453	rs5098	G>C	0.14
1735	rs5096	A>G	0.35
1803	rs5095	A>G	0.20
1853	rs5094	G>A	0.11
1993	rs2239013	C>T	0.13
1994	rs5093	G>A	0.09
2104	rs5092	T>C	0.02
2511	rs12721041	C>T	0.05
2645	rs5091	C>T	0.15
2981	rs5089	C>T	0.05

Table 6. Allelic Variants in the North Karelia, Finland Population (14 total)

Site	rs Number	nt Change	Minor Allele Frequency (MAF)
120	rs12721040	G>A	0.02
165	rs9282602	DEL	0.41
274	rs5110	C>A	0.02
315	rs675	T>A	0.15
933	rs12721043	C>A	0.09
964	rs2234668	G>A	0.04
974	rs5104	T>C	0.20
1192	rs5103	A>G	0.12
1334	rs5100	A>G	0.50
1735	rs5096	A>G	0.50
1803	rs5095	A>G	0.15
1993	rs2239013	C>T	0.08
2104	rs5092	T>C	0.35
2695	rs5090	C>G	0.08

Table 7. Allelic Variants in the Rochester, MN Population (15 total)

Site	rs Number	nt Change	Minor Allele Frequency (MAF)
165	rs9282602	DEL	0.31
274	rs5110	C>A	0.08
315	rs675	T>A	0.11
933	rs12721043	C>A	0.02
964	rs2234668	G>A	0.04
974	rs5104	T>C	0.15
1192	rs5103	A>G	0.02
1334	rs5100	A>G	0.31
1735	rs5096	A>G	0.31
1803	rs5095	A>G	0.15
1853	rs5094	G>A	0.02
1993	rs2239013	C>T	0.04
1994	rs5093	G>A	0.02
2104	rs5092	T>C	0.17
2695	rs5090	C>G	0.06

1.6 SPECIFIC AIMS

This study aims to further evaluate the genetic variation in the *APOA1* and *APOA4* genes, and correlate this variation with HDL-cholesterol levels in two well-characterized populations: African Blacks from Benin, Nigeria and Non-Hispanic Whites (NHWs) from Colorado, U.S.

Aim 1: Sequence the *APOA1* and *APOA4* genes in a subset of individuals with HDL-cholesterol in the upper 5th percentile (47 NHWs and 48 Blacks) and lower 5th percentile (48 NHWs and 47 Blacks).

Aim 2: Identify rare and common variants within the data generated from sequencing the *APOA1* and *APOA4* genes in this population subset.

Aim 3: Determine the associations of both rare and common *APOA1* and *APOA4* variants on HDL-cholesterol levels in the general population of NHWs and African blacks.

2.0 SUBJECTS AND METHODS

2.1 SUBJECTS

2.1.1 Study Populations

The subjects used in this study are summarized in Table 8. Samples from the African Black population were obtained as part of a study on coronary heart disease risk factors in blacks. Subjects were recruited from junior and senior staff, at a variety of salary grades, in government ministries in Sokoto and Benin City, Nigeria. During the initial study demographic and health information was gathered from participants; detailed information about the study population is available elsewhere.[91,92]

Samples from the Non-Hispanic Whites (NHWs) were obtained from the San Luis Valley Southern Colorado Diabetes Study. The subjects involved in this study were normoglycemic. A detailed description of the sample population is available elsewhere.[93,94]

The esterase-oxidase method was used to measure fasting total serum cholesterol.[95,96] Following dextran sulfate magnesium precipitation, total HDL-cholesterol was determined enzymatically.[95,97] The DNA used for sequencing and TaqMan genotyping was extracted from clot sample (Blacks) and from buffy coat (NHWs) using standard DNA extraction procedures.

Table 8. Sample populations

Population	Men (%)	Women (%)	Total (%)
African Blacks	495 (62.8)	293 (37.2)	788 (55.8)
U.S. Whites	295 (47.4)	328 (52.6)	623 (44.2)
Total	790 (56.0)	621 (44.0)	1,411 (100)

2.1.2 Subset of the Study Population Used for Sequencing

Ninety-five samples (47 NHW and 48 Black) from individuals with HDL-cholesterol levels in the upper 5th percentile and 95 samples (48 NWH and 47 Black) from individuals with HDL-cholesterol levels in the lower 5th percentile were selected for sequencing and screening for common and rare variants. The sample characteristics of selected individuals in the high and low HDL groups, are summarized in Table 9, including: gender, age, BMI, LDL, HDL, total cholesterol, and triglyceride levels.

Table 9. Biometric and Quantitative Data (mean±SD) of NHWs and Blacks Used for DNA Sequencing

Variable	*NHWs (n=95)*			*Blacks (n=95)*		
	High HDL (*n* =47)	Low HDL (*n* =48)	*p* -value	High HDL (*n* =48)	Low HDL (*n* =47)	*p* -value
Sex (M/F)	24/23	24/24	0.92	24/24	23/24	0.92
Age (years)	55.45 ± 9.80	53.03 ± 10.54	0.25	41.29 ± 8.72	40.87 ± 7.16	0.8
BMI (kg/m2)	23.17 ± 3.17	27.35±3.90	<0.001	22.06 ± 4.71	23.91 ± 5.51	0.08
Total cholesterol (mg/dl)	227.34±51.76	208.81±44.65	0.06	201 ± 39.68	141.68 ± 31.03	<0.001
LDL cholesterol (mg/dl)	126.84±46.95	136.95±41.28	0.28	112.55 ± 39.75	95.04 ± 28.28	0.02
HDL cholesterol (mg/dl)	**77.68±13.32**	**31.81±4.37**	**<0.001**	**76.05 ± 7.53**	**25.51 ± 5.66**	**<0.001**
Triglycerides (mg/dl)	114.09±60.88	240.21±153.22	<0.001	61.98 ± 19.85	95.79 ± 73.21	0.003

2.2 DNA SEQUENCING

APOA1 primers were designed using Primer3 software version 0.4.0 (http://frodo.wi.mit.edu) to create nine overlapping PCR amplicons. The area covered by the primers includes the four exons and three introns in the gene, plus approximately 940bp of the 5' flanking region (putative promoter region) and approximately 2.5Kb of the 3' flanking region. Primers for amplicon 5 were redesigned because the PCR fragment did not amplify using the original primer set. Additional primers were also designed to amplify a PCR fragment spanning the amplicon 4 and 5 junction because of a sequencing gap in this region. Sequencing was performed in both the forward and reversed direction of all of the samples except for the PCR fragment spanning the amplicon 4 and 5 junction; for this amplicon only forward sequencing was performed. Table 10 is a comprehensive list of the primers used to sequence *APOA1* in this study.

Table 10. *APOA1* Polymerase Chain Reaction (PCR) Primers

PCR Amplicons	Forward Primer	Reverse Primer
1	5'-GCCTTCCTTGACAGCTTTGT-3'	5'-CTGCACCAACTGAGCAGAAT-3'
2	5'-AGAGGCTGCTTCCTTTGTGT-3'	5'-CCTGGCACTCAAGTTCACAT-3'
3	5'-TTCAGACATGAGTGCAAGGAG-3'	5'-AGAAGCTGGCCTGAGTAAGAA-3'
4	5'-CAGTGTCCTCATCCATGCTC-3'	5'-GTCTTAGGGCCAAGATCGAC-3'
4-5 junction	5'-CCAGCTAAAGCAACCCTGTT-3'	5'-GTTTCCAAAGTGGGAAGCAG-3'
5	5'-TTGGATTGTCTGTGGCTTTG-3'	5'-AGAAGAAGTGGCAGGAGGAG-3'
5-new	5'-TCCGCTGTGACTTCCTTTCT-3'	5'-ATGAGCAAGGATCTGGAGGA-3'
6	5'-AGTGGGCTCAGCTTCTCTTG-3'	5'-AAGCCCCTTTCCCTTCTTC-3'
7	5'-AGTGGCCTAGCATTTCCAGT-3'	5'-CTAACCTAGGGAGCCAACCA-3'
8	5'-GGGAGGGGAGACCCAGAT-3'	5'-CCCACTGAACCCTTGACC-3'
9	5'-GTCCTGGCAATGTGGAACTT-3'	5'-TAACTTGCCCACGATCTTCC-3'

Publicly available information from the Seattle SNP database (http://pga.mbt.washington.edu) was used to order M13-tagged primers for sequencing of *APOA4*; four overlapping PCR amplicons were created using these primers. The area covered by the primers includes the three exons and two introns in the gene, plus approximately 780bp of the 5' flanking region (putative promoter region) and approximately 10bp of the 3' flanking region. Sequencing was performed in both the forward and reversed direction of all of the samples. Table 11 is a comprehensive list of the primers used to sequence *APOA4* in this study; M13-tag sequence is underlined.

Table 11. *APOA4* Polymerase Chain Reaction (PCR) Primers

PCR Amplicons	Forward Primer
1	5'-TGTAAAACGACGGCCAGTCAACCAGTTGAGGCTAGATTCTC-3'
2	5'-TGTAAAACGACGGCCAGTTTCTTCTTCATCTGGAAGGTCAG-3'
3	5'-TGTAAAACGACGGCCAGTCTCAGGATCTCCCACATAGTTTG-3'
4	5'-TGTAAAACGACGGCCAGTTTTCCCTGTCTGAGCTTAGCTTT-3'
PCR Amplicons	Reverse Primer
1	5'-CAGGAAACAGCTATGACCTCAAAGTCAAGATTGACCAGACC-3'
2	5'-CAGGAAACAGCTATGACCGCAGAGGTCAAGAAGACAACATT-3'
3	5'-CAGGAAACAGCTATGACCGGACACTTCTGAGTGCCCAT-3'
4	5'-CAGGAAACAGCTATGACCATGGAGACTGAGAGATGACCGTA-3'

The polymerase chain reaction (PCR) was performed using the GeneAMP® PCR System 9700 thermal cycler with a heated lid (Applied Biosystems, Foster City, CA). The PCR reaction reagents and cycling conditions used in this study are given in Table 12. PCR conditions were optimized through adjusting the $MgCl_2$ or annealing temperature within the range indicated in the table.

Table 12. PCR Reaction and Cycling Conditions

PCR Reaction (total volume 25 µL)		PCR conditions
DNA	3.0 µL	1. 95°C for 5 minutes
dH_2O	12.25-13.75 µL	
10x BufferGold	2.5 µL	2. 95°C for 45 seconds
$MgCl_2$ (25 mM)	1-2.5 µL	3. 58-60°C for 45 seconds
dNTPs (1.25mM)	3.8 µL	4. 72°C for 1 minute
Forward Primer (20mM)	0.4 µL	-repeat 2-4 39x
Reverse Primer (20mM)	0.4 µL	5. 72°C for 10 minutes
AmpliTaqGold (5U/µL)	0.15 µL	6. Cool to 4°C

Gel electrophoresis was used to check for successful amplification of each of the PCR fragments. For each sample, 7µL of PCR product was combined with 5µL of loading dye and 8µL distilled water, and loaded into a 96-well pre-cast agarose gel (Invitrogen™ E-Gel® 96 2% with SYBR® Safe). The EG program on the electrophoresis base (Invitrogen™ E-Base™) was used to run the gel for 8 minutes. Reamplification was done for a subset of samples that failed the initial amplification. For this subset, confirmation was performed using agarose gels with ethidium bromide. The 7µL of PCR product was combined with 5µL of loading dye and loaded into a 2% agarose gel in TBE buffer (tris, boric acid, and disodium EDTA dihydrate) and stained with ethidium bromide. Electrophoresis was run for 15 minutes at 250V. Both the 96-well and hand-poured gels were visualized using a UV transilluminator.

A commercial sequencing facility (Genomic Services of Agencourt Bioscience Corporation, Beverly, MD) performed automated sequencing and capillary electrophoresis. The programs used to analyze the sequence results were: Sequencher version 4.8 (Gene Codes

Corporation, Ann Arbor, MI), and Variant Reporter version 1.0 (Applied Biosystems, Foster City, CA).

2.3 GENOTYPING WITH TAQMAN

For common SNPs (MAF≥5%) available pre-made TaqMan SNP genotyping assays were ordered. Seven assays were available, two for *APOA1* and five for *APOA4*. Table 13 lists the seven TaqMan assays that were used for genotyping in the NHW, Black population, or both.

Table 13. TaqMan® SNP Genotyping Assays

Reference SNP ID	Gene	Assay ID	Population
rs5070	*APOA1*	C_2679584_10	NHWs, Blacks
rs5072	*APOA1*	C_11482715_1_	NHWs, Blacks
rs5092	*APOA4*	C_2679569_10	NHWs, Blacks
rs5100	*APOA4*	C_2679565_10	NHWs, Blacks
rs5104	*APOA4*	C_11482766_10	NHWs, Blacks
rs5106	*APOA4*	C_11482768_10	Blacks
rs5109	*APOA4*	C_11482772_10	Blacks

The TaqMan procedure involved DNA amplification and endpoint fluorescent reading using the ABI Prism 9700HT Sequence Detection System (Applied Biosystems, Foster City, CA). The reagents listed in Table 14 were added to 384-well plates containing dried whole genome amplified DNA. The TaqMan genotyping Assay Mix contains: sequence specific forward and reverse primers, a TaqMan minor groove binder (MGB) probe labeled with VIC dye at the 5' end and a nonfluorescent quencher (NFQ) at the 3' end, and a TaqMan MGD labeled with FAM dye at the 5' end and a NFQ at the 3' end. The GeneAMP® PCR System 9700 thermal cycler with a heated lid (Applied Biosystems, Foster City, CA) was used for PCR amplification; cycling conditions are displayed in Table 14.

Table 14. TaqMan Reaction and Cycling Conditions

TaqMan Reaction	(total volume 5 μL)	PCR conditions
dH_2O	2.435 μL	1. 95°C for 10 minutes
TaqMan Genotyping Master Mix (2x)	2.5 μL	2. 95°C for 15 seconds
TaqMan Genotyping Assay Mix (40x)	0.065 μL	3. 60°C for 1 minute -repeat 2-3 49x

Discrimination of alleles is possible because of the selective annealing of the TaqMan probes; each MGB probe binds to the target sequence harboring the SNP of interest during the annealing step (step 3 in Table 14) AmpliTaq Gold polymerase, which is part of the TaqMan Genotyping Master Mix, cleaves the probes that hybridize to the target sequence. The reporter dye is separated from the NFQ, releasing a fluorescent signal. Fluorescence is suppressed if the probes do not hybridize to the target sequence; the reporter dye does not separate from the NFQ.

The genotyping call rates for all seven assays are shown in Table 15. The genotyping discrepancy rate was <2.2% for each variant based on a 20-30% repeat of the samples.

Table 15. Genotyping Call Rates for TaqMan

Reference SNP ID	**NHWs (%)**	**Blacks (%)**
rs5070	98.56	95.94
rs5072	99.52	98.22
rs5092	99.52	96.70
rs5100	99.68	96.32
rs5104	99.04	94.92
rs5106	-	96.57
rs5109	-	98.22

2.4 STATISTICAL METHODS

Direct allele counting was used to determine allele frequencies in this study. Concordance of the genotype distribution to Hardy-Weinberg equilibrium (HWE) was tested for each variant using a χ^2 goodness-of-fit test. A standard Z-test of two binomial proportions was used to compare the allele frequencies. Linkage disequilibrium (LD) pattern and tagSNPs were determined using Haploview version 4.3 (http://www.broad.mit.edu/mpg/haploview). All dependant quantitative variables were transformed (using a log or square root transformation) when necessary to reduce the effects of non-normality. The significant covariates for each dependant variable were identified using stepwise regression in both directions. The most parsimonious set of covariats was determined separately for males and females within the NWH and Black populations. One-way analysis of variance (ANOVA) was performed separately for males and females within the NHW and Black populations to test for the effects of genotypes on the means of quantitative traits (which were transformed and adjusted when necessary). The R statistical software package version 2.3.1 (http://www.r-project.org) and Statistical Analysis Software (SAS) was used to perform all computations. Two genetic models were used for data analysis, the additive and codominant models. A p-value of <0.05 under one of these models was considered as suggestive evidence of association.

3.0 RESULTS

3.1 DNA SEQUENCING

3.1.1 *APOA1*

A total of 53 single nucleotide substitutions plus one insertion/deletion (indel) polymorphism were identified in *APOA1*. For the indel variant in 3' flanking region three different alleles were observed; insertion of T (9 T's) in NHWs, deletion of T (7 T's) in Blacks, and the wild type of 8 T's in both populations (in the reverse strand sequence). The insertion allele was previously reported in public databases, whereas the deletion allele was novel. The location of variants were as follows: 8 in putative promoter region, 6 in exons, 12 in introns, and the remaining in 3' flanking region.

Twenty-nine of the identified variant locations had already been reported in publicly available databases (SeattleSNPs database, CHIP Bioinformatics database, dbSNP), while 25 were novel (not previously reported). Seventeen single nucleotide substitutions were observed only in NHWs; 20 single nucleotide substitutions were observed only in Blacks. Of a total of 25 identified new variants, 10 were in NHWs and 15 were in Blacks; thus, none of the novel variants were observed in both populations. Of 34 variants identified in NHWs 23 were relatively rare, with MAF <5%. Of 37 variants identified in Blacks, 17 were relatively rare with

MAF <5%. All newly identified variants in each population had <5% MAF. Table 16 lists all of the variants identified in this study. The chromatograms illustrating the 25 novel variants in *APOA1* are shown in figure 8. The annotated FASTA file and related information is given is section 3.1.3.

Table 16. *APOA1* Sequence Variants.

APOA1 Variant*/**	rs# (CHIP&GB)	Location	Amino Acid Change	Population	MAF (NHWs)	MAF (Blacks)
206A>C	rs7123454	3'-flanking	---	Both	0.174	0.349
338A>G	Novel Variant	3'-flanking	---	Blacks	---	0.032
386G>A	Novel Variant	3'-flanking	---	Blacks	---	0.005
477C>T	Novel Variant	3'-flanking	---	Blacks	---	0.016
533C>T	rs12721031	3'-flanking	---	NHWs	0.016	---
631A>G	rs7948159	3'-flanking	---	Blacks	---	0.484
656C>T	Novel Variant	3'-flanking	---	Blacks	---	0.005
689C>T	Novel Variant	3'-flanking	---	NHWs	0.005	---
894G>A	Novel Variant	3'-flanking	---	Blacks	---	0.016
959G>C	Novel Variant	3'-flanking	---	NHWs	0.005	---
1049T>A	rs1263162	3'-flanking	---	Both	0.011	0.128
1128G>T	rs11216153	3'-flanking	---	Both	0.191	0.095
1143G>T	Novel Variant	3'-flanking	---	Blacks	---	0.005
1308C>T	rs12721030	3'-flanking	---	Both	0.234	0.011
1407del/insT***	rs12721027****	3'-flanking	---	ins(NHWs); del(Blacks)	0.005	0.032
1507T>C	Novel Variant	3'-flanking	---	NHWs	0.005	---
1546A>G	rs525028	3'-flanking	---	Both	0.372	0.079
1549C>T	Novel Variant	3'-flanking	---	NHWs	0.005	---
1598T>G	rs10750098	3'-flanking	---	Both	0.132	0.084
1620A>G	rs12721028	3'-flanking	---	Both	0.184	0.300
1749T>C	rs12718462	3'-flanking	---	NHWs	0.037	---
1965T>C	Novel Variant	3'-flanking	---	Blacks	---	0.032
2077G>A	rs12721025	3'-flanking	---	NHWs	0.037	---
2120C>A	Novel Variant	3'-flanking	---	Blacks	---	0.011
2198T>G	rs12721026	3'-flanking	---	NHWs	0.037	---
2215C>A	Novel Variant	3'-flanking	---	Blacks	---	0.005
2373T>C	rs12718436	3'-flanking	---	Both	0.042	0.389
2376A>T	rs5081	3'-flanking	---	Both	0.011	0.126
2626G>C	rs5080	exon 4	syn	Blacks	---	0.016
2652C>A*****	Novel Variant	exon 4	Glu>Ter	NHWs	0.005	---
2880C>G	Novel Variant	exon 4	Glu>Gln	Blacks	---	0.005
3220G>A	rs5076	intron 3	---	Both	0.042	0.437
3307C>A	Novel Variant	intron 3	---	NHWs	0.011	---
3368G>A	rs7116797	intron 3	---	Both	0.147	0.358
3431G>A	rs12718464	intron 3	---	NHWs	0.026	---
3543C>T	rs5073	intron 3	---	Blacks	---	0.058
3613G>A	rs5072	intron 3	---	Both	0.105	0.100
3714G>A	rs2070665	intron 3	---	Both	0.105	0.096
3769A>C	Novel Variant	exon 3	Ser>Ala	NHWs	0.005	---
3867G>T	Novel Variant	exon 3	Pro>His	Blacks	---	0.005
3959G>T	Novel Variant	intron 2	---	NHWs	0.005	---
4050G>A	rs5070	intron 2	---	Both	0.332	0.436
4151G>C	Novel Variant	exon 2 / 5'-UTR	---	NHWs	0.005	---
4208C>T	Novel Variant	intron 1	---	Blacks	---	0.005
4283C>T	rs1799837	intron 1	---	NHWs	0.005	---
4284G>A	rs5069	intron 1	---	Both	0.032	0.437
4443C>T	rs670	5'-flanking / promoter	---	Both	0.189	0.121
4693T>G	rs12718466	5'-flanking / promoter	---	NHWs	0.042	---
4732C>A	rs12718467	5'-flanking / promoter	---	Blacks	---	0.106
4807C>T	rs12691374	5'-flanking / promoter	---	Blacks	---	0.068
4987T>C	Novel Variant	5'-flanking / promoter	---	Blacks	---	0.026
5055A>T	Novel Variant	5'-flanking / promoter	---	Blacks	---	0.068
5066G>T	Novel Variant	5'-flanking / promoter	---	Blacks	---	0.011
5131C>T	Novel Variant	5'-flanking / promoter	---	NHWs	0.011	---

* The nucleotide change represented in the table is for the minor allele in the NHW population.
** The locations and nucleotide changes are based on the reverse strand sequence used in the SeattleSNPs database.
*** Triallelic insertion/deletion polymorphism.
**** rs number is for the insertion and wild type alleles.
***** Suspicious variants with low sequence quality.

Figure 8. Chromatograms for New Variants in the *APOA1* Gene.

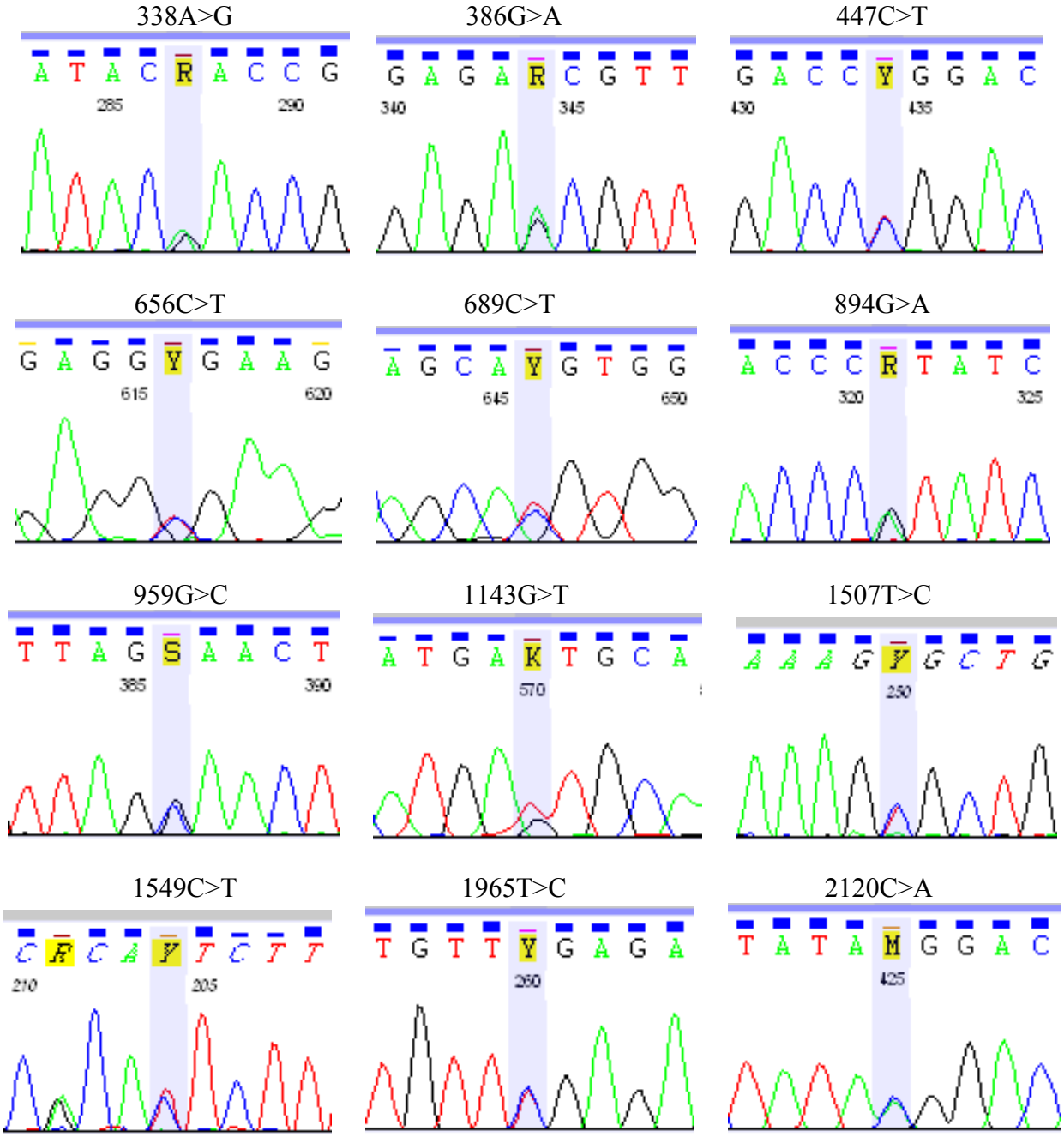
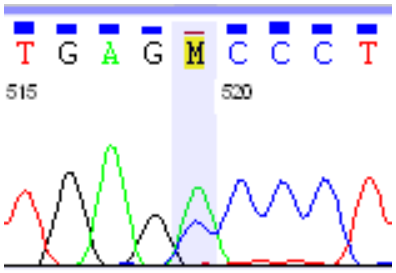
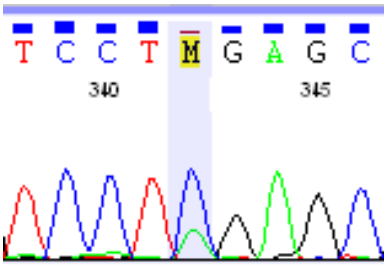


Figure 8 Continued

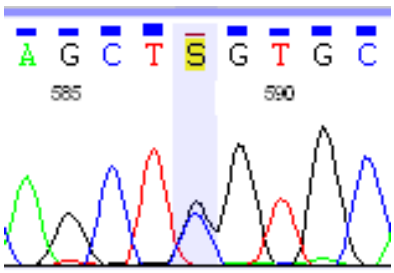
2215C>A



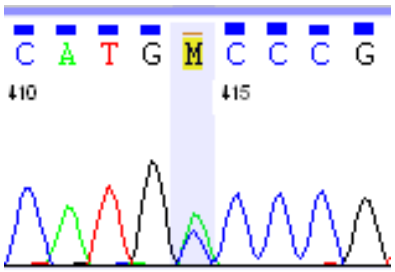
2652C>A*



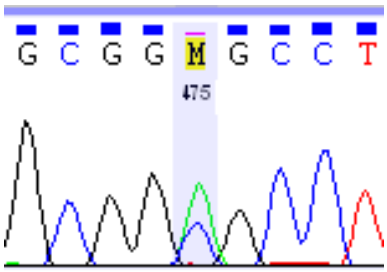
2880C>G



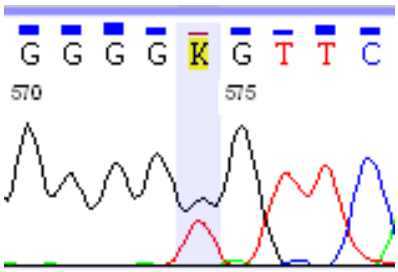
3307C>A



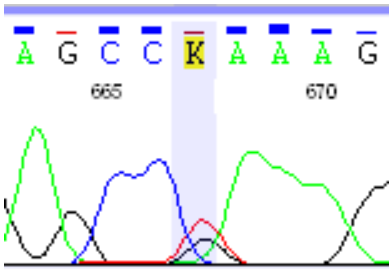
3769A>C



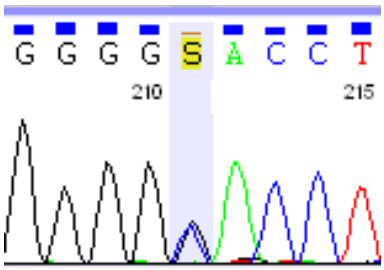
3867G>T



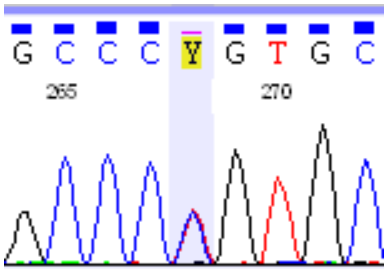
3959G>T



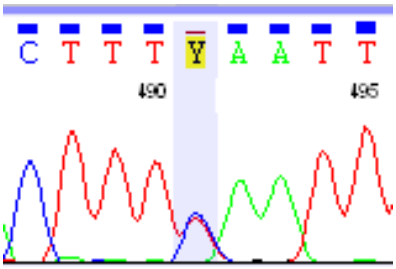
4151G>C



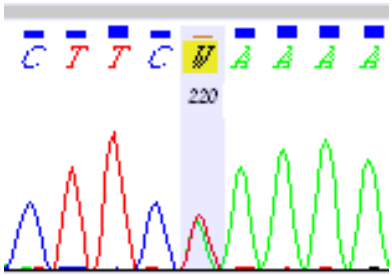
4208C>T



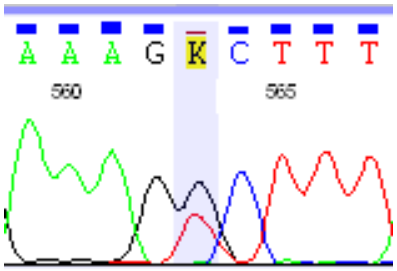
4987T>C



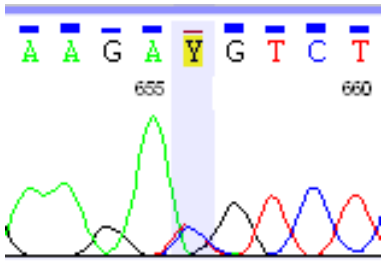
5055A>T



5066G>T



5131C>T



3.1.2 *APOA4*

A total of 41 single nucleotide substitutions were observed in *APOA4*, plus two indel polymorphisms. Both indels were located in exon 3 and included a four nucleotide deletion (ACAG) at the 3' untranslated region (UTR), and a 12 nucleotide insertion (CTGTTCCTGCTG) affecting the coding region. Although the latter variant was reported in the literature, it is shown as a novel variant in Table 17 because it is not present in the public databases.[88]

Twenty-three of the identified variants had already been reported in publicly available databases, while 20 were novel (not previously reported). Thirteen variants were observed only in NHWs; 20 were observed only in Blacks. Of a total of 20 identified new variants, 7 were in NHWs and 13 were in Blacks; thus, none of the novel variants were observed in both populations. Of 23 variants identified in NHWs 12 were relatively rare, with MAF <5%. Of 30 variants identified in Blacks, 17 were relatively rare with MAF <5%. Of the 20 novel variants identified in this study, 18 were relatively rare with MAF <5%. Two of the novel variants in Blacks had a MAF of 0.053. Table 17 lists all of the variants identified in this study. The chromatograms illustrating the 20 novel variants in *APOA4* are shown in figure 9. The annotated FASTA file and related information is given is section 3.1.3.

Table 17. *APOA4* Sequence Variants.

APOA4 Variant*/**	rs# (CHIP&GB)	Location	Amino Acid Change	Population	MAF (NHWs)	MAF (Blacks)
120G>A	rs12721040	exon 3 / 3'-UTR	---	NHWs	0.021	---
165delACAG	rs9282602	exon 3 / 3'-UTR	---	Both	0.495	0.058
274C>A	rs5110	exon 3	Gln>His	NHWs	0.089	---
288ins12	Novel Variant	exon 3	---	Blacks	---	0.016
315T>A	rs675	exon 3	Thr>Ser	Both	0.195	0.074
357A>C	Novel Variant	exon 3	Ser>Ala	Blacks	---	0.053
406C>A	rs5109	exon 3	syn	Blacks	---	0.126
422G>T***	Novel Variant	exon 3	Pro>His	NHWs	0.005	---
520C>T	Novel Variant	exon 3	syn	NHWs	0.005	---
568G>A	rs5106	exon 3	syn	Blacks	---	0.042
634G>A	rs5105	exon 3	syn	Blacks	---	0.021
755C>T	Novel Variant	exon 3	Arg>His	Blacks	---	0.005
945G>A	Novel Variant	exon 3	syn	NHWs	0.005	---
952C>T	Novel Variant	exon 3	syn	NHWs	0.005	---
964G>A	rs2234668	exon 3	syn	NHWs	0.058	---
974T>C	rs5104	exon 3	Asn>Ser	Both	0.163	0.106
1033G>T	Novel Variant	exon 3	Asn>Lys	NHWs	0.005	---
1192A>G	rs5103	exon 3	syn	NHWs	0.053	---
1198G>A	rs5101	exon 3	syn	Blacks	---	0.425
1274G>A	Novel Variant	intron 2	---	Blacks	---	0.011
1326A>G	Novel Variant	intron 2	---	Blacks	---	0.053
1334A>G	rs5100	intron 2	---	Both	0.411	0.404
1371C>T	Novel Variant	intron 2	---	Blacks	---	0.043
1453G>C	rs5098	intron 2	---	Blacks	---	0.012
1735A>G	rs5096	intron 2	---	Both	0.411	0.402
1743T>G	Novel Variant	intron 2	---	Blacks	---	0.005
1803A>G	rs5095	intron 2	---	Both	0.195	0.065
1853G>A	rs5094	intron 2	---	Both	0.011	0.081
1948C>A	Novel Variant	intron 2	---	Blacks	---	0.021
1993C>T	rs2239013	intron 2	---	Both	0.042	0.043
1994G>A	rs5093	intron 2	---	Both	0.032	0.011
2104T>C	rs5092	exon 2	syn	Both	0.216	0.160
2287G>A	Novel Variant	intron 1	---	NHWs	0.005	---
2327C>A***	Novel Variant	intron 1	---	Blacks	---	0.005
2406C>G***	Novel Variant	intron 1	---	Blacks	---	0.005
2645C>T	rs5091	exon 1 / 5'-UTR	---	Blacks	---	0.050
2685C>T***	Novel Variant	5'-flanking / promoter	---	Blacks	---	0.005
2695C>G	rs5090	5'-flanking / promoter	---	NHWs	0.068	---
2705C>T	Novel Variant	5'-flanking / promoter	---	Blacks	---	0.005
2978C>A	rs7929134	5'-flanking / promoter	---	NHWs	0.021	---
2981C>T	rs5089	5'-flanking / promoter	---	Blacks	---	0.037
2984G>A	Novel Variant	5'-flanking / promoter	---	NHWs	0.005	---
3146G>A	Novel Variant	5'-flanking / promoter	---	Blacks	---	0.005

* The nucleotide change represented in the table is for the minor allele in the NHW population.
** The locations and nucleotide changes are based on the reverse strand sequence used in the SeattleSNPs database.
*** Suspicious variants with low sequence quality.

Figure 9. Chromatograms for Novel Variants in the *APOA4* Gene.

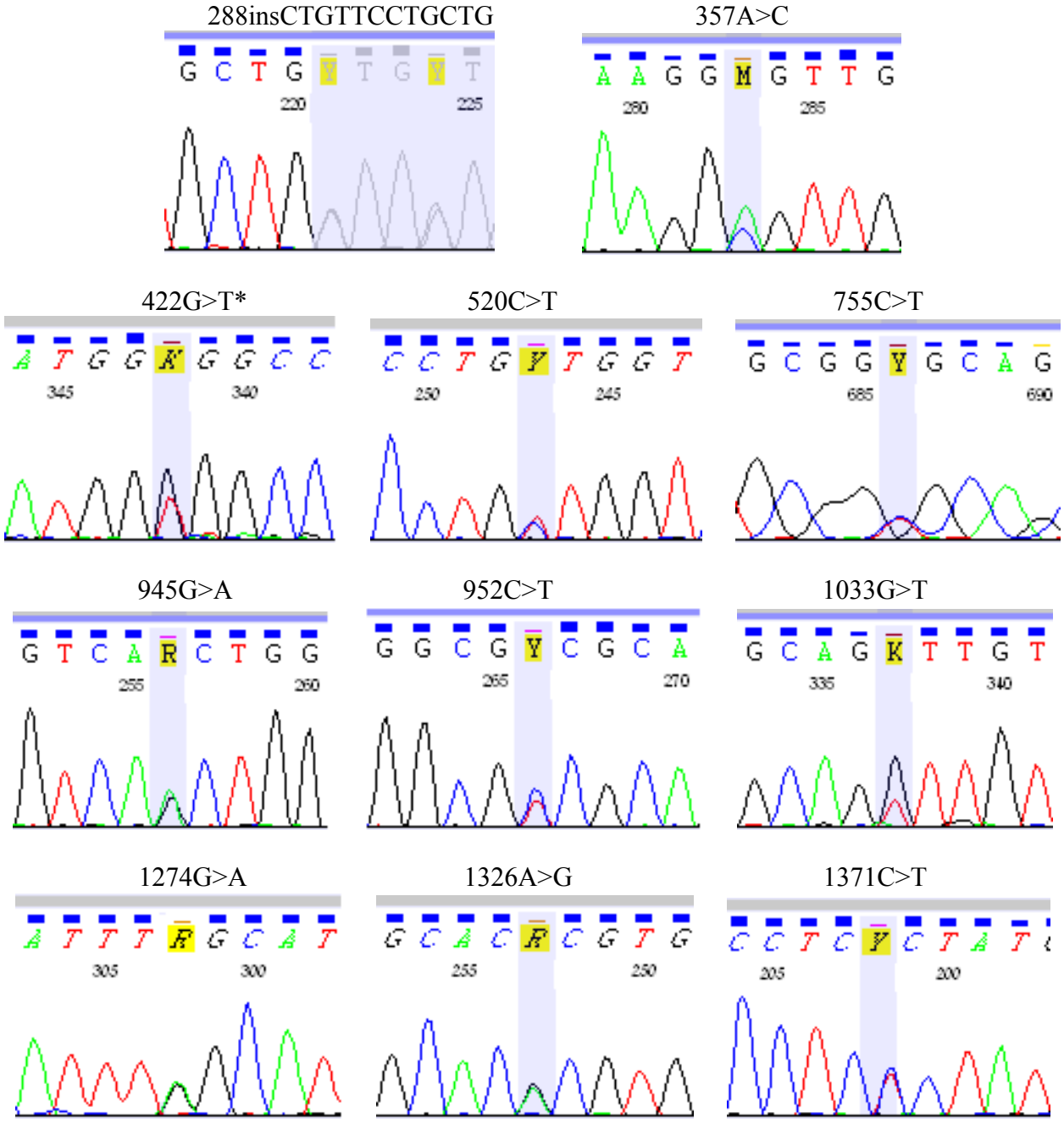
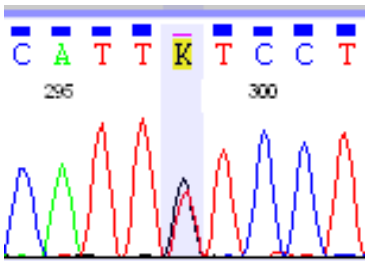
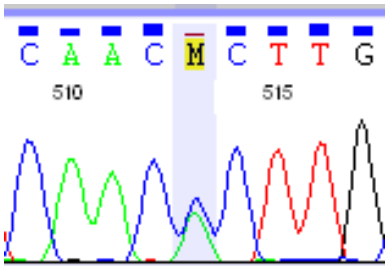


Figure 9 Continued

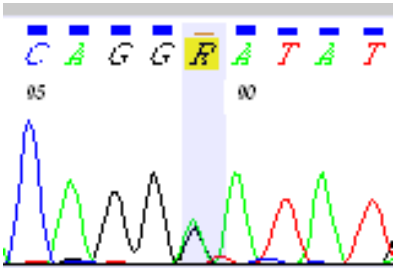
1743T>G



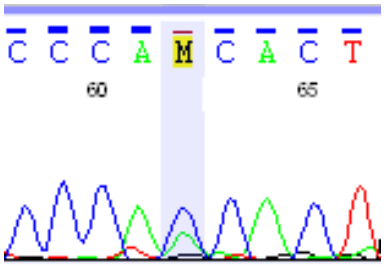
1948C>A



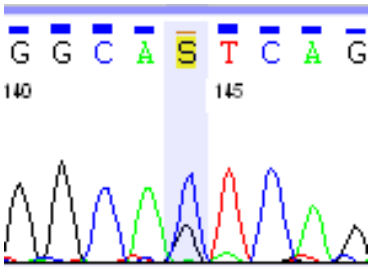
2287G>A



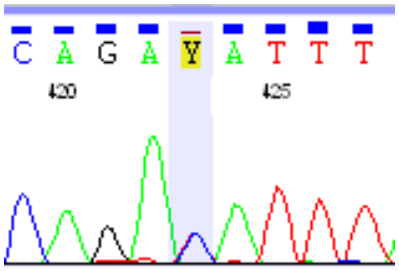
2327C>A*



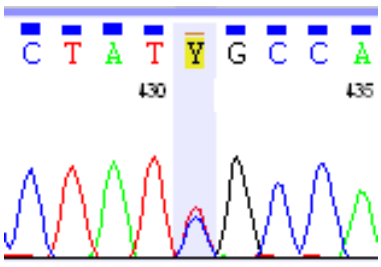
2406C>G*



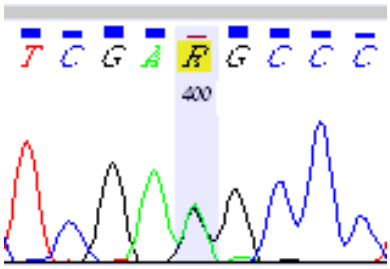
2685C>T*



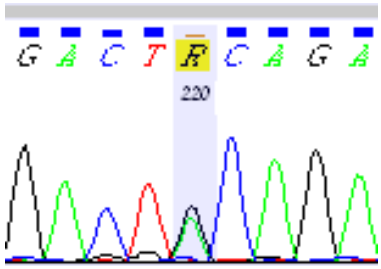
2705C>T



2984G>A



3146G>A



3.1.3 *APOA1* and *APOA4* Annotated Sequence

Figures 10 and 11 depict the variants identified in *APOA1* and *APOA4* within a color FASTA representation of the annotated reference sequence from the CHIP Bioinformatics database (http://snpper.chip.org). The sequences from the CHP Bioinformatics database were used as a reference instead of the sequences from the SeattleSNPs database because the SeattleSNPs database sequences were not given in the forward direction (forward strand). However, the SeattleSNPs database was used as a reference to design and order PCR primers, and for comparison with sequencing results in this. Therefore, the SeattleSNPs location nomenclature has been used throughout the text, tables, and figures. The variants identified in this study also reported in public databases and with rs numbers in GenBank are shown in **blue font**; rs number followed by SeattleSNPs database-based locations using the reverse strand (variants reported in SeattleSNPs database are in paranthesis, variants not reported in SeattleSNPs database are in brackets). The new variants identified in this study are shown in **red font** in brackets. The suspicious variants with low sequence quality are marked with a *. Variants reported in public databases that were not identified in this study are show in **purple font**.

Figure 10. *APOA1* Annotated Sequence

```
116,214,548  aacaaccctg accattcttg ccccattttg cagatagaaa accgaggctc
116,214,498  agagagatta tataacttgc ccacgatctt cctccagcaa gatggaggcc
116,214,448  aagtgaaatg agaaagcagg tctcctgcca cttcctttgc ccagaggtct
116,214,398  tctccccaca ccagggcttc ccaagggctg agatccagtc acacctgtgc
116,214,348  gtgatcaaat ataagtgtga acaatgcaaa gggagacgtc ttcaatctaa [5131]
116,214,298  ggggcttcaa ttctgtaatg taattctgag attatgccct tttttgttaa
116,214,248  agcctttcct tttgaagtg atggtcactg tagatggtga gggttttttg [5066];[5055]
116,214,198  gaggcggaca atatctttac atgacaaaat taaaagttgg cagctccgaa [4987]
116,214,148  ttgatctctg gagtgttttg aaatgcaaga ggtctccgaa acctcagtct
116,214,098  gggagccacg gagggctctc ccctctcccc aggtttacca gtttgggagg
116,214,048  cttggagaga ggcctggagg acctgctggg gactaaagaa gagcactggt
116,213,998  gggaggacag ggcgggggaa gggggagggg agtgaagtag tctccctgga rs12691374(4807)
116,213,948  atgctggtgg tgggggaggc agtctccttg gtggaggagt cccagcgtcc rs12718467(4732)
116,213,898  ctcccctccc ctcctctgcc aacacaatgg acaatggcaa ctgcccacac rs12718466(4693)
116,213,848  actcccatgg aggggaaggg gatgagtgca gggaaccccg accccacccg
116,213,798  ggagacctgc aagcctgcag acactcccct cccgccccca ctgaaccctt
116,213,748  gacccctgcc ctgcagccccc cgcagcttgc tgtttgccca ctctatttgc rs2727786;rs2542054
116,213,698  ccagccccag ggacagagct gatccttgaa ctcttaagtt ccacattgcc rs2542053
116,213,648  aggaccagtg agcagcaaca gggccggggc tgggcttatc agcctcccag rs670(4443)
116,213,598  cccagaccct ggctgcagac ataaataggc cctgcaagag ctggctgctt

116,213,548  agagactgcg agaaggaggt gcgtcctgct gcctgccccg gtcactctgg Exon 1 Intron 1
116,213,498  ctccccagct caaggttcag gccttgcccc aggccggggcc tctgggtacc rs5069(4284);rs1799837[4283]
116,213,448  tgaggtcttc tcccgctctg tgcccttctc ctcacctggc tgcaatgagt rs12721032(4245)
116,213,398  gggggagcac ggggcttctg catgctgaag gcaccccact cagccaggcc [4208]
116,213,348  cttcttctcc tccaggtccc ccacggccct tcaggATGAA AGCTGCGGTG Exon 2 [4151]
116,213,298  CTGACCTTGG CCGTGCTCTT CCTGACGGGT AGGTGTCCCC TAACCTAGGG Intron 2
116,213,248  AGCCAACCAT CGGGGGGCTT TCTCCCTAAA TCCCCGTGGC CCACCCTCCT rs5070(4050)
116,213,198  GGGCAGAGGC AGCAGGTTTC TCACTGGCCC CCTCTCCCCC ACCTCCAAGC
116,213,148  TTGGCCTTTC GGCTCAGATC TCAGCCCACA GCTGGCCTGA TCTGGGTCTC [3959]
116,213,098  CCCTCCCACC CTCAGGGAGC CAGGCTCGGC ATTTCTGGCA GCAAGATGAA Exon 3
116,213,048  CCCCCCCAGA GCCCCTGGGA TCGAGTGAAG GACCTGGCCA CTGTGTACGT [3867];rs28929476
116,212,998  GGATGTGCTC AAAGACAGCG GCAGAGACTA TGTGTCCCAG TTTGAAGGCT rs28931574;[3769]
116,212,948  CCGCCTTGGG AAAACAGCTA AAGTAAGGAC CCAGCCTGGG GTTGAGGGCA rs12718465(3766);rs5071;rs17407917 Intron 3
116,212,898  GGGGTAGGGG GCAGAGGCCT GTGGGATGAT GTTGAAGCCA GACTGGCCGA rs2070665(3714)
116,212,848  GTCCTCACCT AATATCTGAT GAGCTGGGCC CCACAGATGG TCTGGATGGA
116,212,798  GAAACTGGAA TGGGATCTCC AGGCAGGGTC ACAGCCCATG TCCCCTGCAA rs5072(3613)
116,212,748  AGGACAGACC AGGGCTGCCC GATGCGTGAT CACAGAGCCA CATTGTGCCT rs5073(3543)
116,212,698  GCAAGTGTAG CAAGCCCCTT TCCCTTCTTC ACCACCTCCT CTGCTCCTGC rs13306170
116,212,648  CCAGCAAGAC TGTGGGCTGT CTTCGGAGAG GAGAATGCGC TGGAGGCATA rs12718464(3431)
116,212,598  GAAGCGAGGT CCTTCAAGGG CCCACTTTGG AGACCAACGT AACTGGGCAC
116,212,548  TAGTCCCAGC TCTGTCTCCT TTTTAGCTCC TCTCTGTGCC TCGGTCCAGC rs7116797(3368)
116,212,498  TGCACAACGG GGCATGGCCT GGCGGGGCAG GGGTGTTGGT TGAGAGTGTA [3370]
116,212,448  CTGGAAATGC TAGGCCACTG CACCTCCGCG GACAGGTGTC ACCCAGGGCT rs5075;rs5076(3220)
116,212,398  CACCCCTGAT AGGCTGGGGC GCTGGGAGGC CAGCCCTCAA CCCTTCTGTC
116,212,348  TCACCCTCCA GCCTAAAGCT CCTTGACAAC TGGGACAGCG TGACCTCCAC Exon 4
116,212,298  CTTCAGCAAG CTGCGCGAAC AGCTCGGCCC TGTGACCCAG GAGTTCTGGG
116,212,248  ATAACCTGGA AAAGGAGACA GAGGGCCTGA GGCAGGAGAT GAGCAAGGAT rs17145083;rs2727787
116,212,198  CTGGAGGAGG TGAAGGCCAA GGTGCAGCCC TACCTGGACG ACTTCCAGAA rs28931575;rs5077
116,212,148  GAAGTGGCAG GAGGAGATGG AGCTCTACCG CCAGAAGGTG GAGCCGCTGC rs4882
116,212,098  GCGCAGAGCT CCAAGAGGGC GCGCGCCAGA AGCTGCACGA GCTGCAAGAG [2880]
116,212,048  AAGCTGAGCC CACTGGGCGA GGAGATGCGC GACCGCGCGC GCGCCCATGT
116,211,998  GGACGCGCTG CGCACGCATC TGGCCCCCTA CAGCGACGAG CTGCGCCAGC rs5078;rs1052925;rs28931573
116,211,948  GCTTGGCCGC GCGCCTTGAG GCTCTCAAGG AGAACGGCGG CGCCAGACTG rs5079
116,211,898  GCCGAGTACC ACGCCAAGGC CACCGAGCAT CTGAGCACGC TCAGCGAGAA rs1053223;rs14081
116,211,848  GGCCAAGCCC GCGCTCGAGG ACCTCCGCCA AGGCCTGCTG CCCGTGCTGG [2652]*;rs5080[2626]
116,211,798  AGAGCTTCAA GGTCAGCTTC CTGAGCGCTC TCGAGGAGTA CACTAAGAAG
116,211,748  CTCAACACCC AGTGAggcgc ccgccgccgc ccccttccc ggtgctcaga
116,211,698  ataaacgttt ccaaagtggg
```

Figure 10 Continued

```
116,211,678  aagcagcttc tttcttttgg gagaatagag gggggtgcgg ggacatccgg
116,211,628  gggagcccgg gtggggcctt tggccctgga gcagggactt cctgccggat rs2849171
116,211,578  ctcaacaact ccgtgcccag actggacgtc ttagggccaa gatcgacgtt rs5081(2376);rs12718463(2373)
116,211,528  ggaggacctg ctggacgcct ggctgcttac gagtgaggga gtagagtctg
116,211,478  ccttagcaag gctcaagtag aaaggaagtc acagcggacc aggcaaagcc
116,211,428  acagacaatc caaggccagg tgccctgaaa ggggctcaaa caaggcctgc [2215]
116,211,378  agccctgtct gaggcgggcc aggaaacagg gttgctttag ctgggagcag rs12721026(2198)
116,211,328  tgggttcccc gtccccagag gtgtgtccgt atagagcctt ctccagccca [2120]
116,211,278  gccgctgtca gcggggcggg acggagcggg gcggcctcag ggagccagcc rs12721025(2077)
116,211,228  actgggattg gggtttggtc ccgggtgcaa gtgaagcgct tggagtttgc
116,211,178  gcctgtcctc ctttactaat tcaaaaacct ctcaaacaga cacttccctt [1965]
116,211,128  ttcttctcac aaggccagta tccccctccc actactccca tcccgcccag
116,211,078  aaacagccgc ggcttcctca ggcacagcag tggaagccag tcctccaccc
116,211,028  cctgcggctc catgccatgc caccccctct ttctgccagc cctggcagaa
116,210,978  gctggcctga gtaagaaaat tcaccaccac ctcttgcagg tacatttttt rs12718462(1749)
116,210,928  tttccaagat gctctcatat ctgtgctctc actgcatcct cccttcccca rs12718461(1717)
116,210,878  catcctggct agattgccat cagacgcaga gcatggatga ggacactgaa
116,210,828  gcctggacct gtgacgtcgc ttgcccagtg aacagcagga tgggctaggc (1620)
116,210,778  cgcgcttttt agaccctgca cccctggcca tccatgatta ttgaaaagag rs10750098(1598);[1549]
116,210,728  tgtgcgggtc gggtgcggtg gctcaagcct gtaatcccag cactttggga rs525028(1546);rs12721029(1541);[1507]
116,210,678  ggctgaggtg ggcgtatcac ttcaggccag gagtttgaga ccagcctggc
116,210,628  caatatggtg aaaccctgtc tctactaaaa atacaaaaaa aatcagctgg rs12721027(1407)
116,210,578  gcatggtggc ttgcacccgt aatcccagct actaggaagg ctgaggcagg
116,210,528  agaatcgctt gaacctggga ggcagaggtc acagtgagcc ggaaatcatgc rs12721030(1308)
116,210,478  cactgcactc cagcctgggc gacggagcaa gactccagct aaaaaaaaaa
116,210,428  aaaaaaaaaa aaaagagtgt gtggcctggc actcaagttc acatgggtgt
116,210,378  gcaggcatgc ctgtgtattc tcacatgacc tccctgctca cggtccctcc
116,210,328  ttgcactcat gtctgaatgt ccccgcgtgc acgcacatgg cttcacagat [1143]; rs11216153(1128)
116,210,278  ctgggcagtg ccttccctac cctctctctg cagggccttt tgccccctca rs1263162(1049)
116,210,228  tgcaggcccc tggataatcg gccccatccc catgtcccca tctccagtgt
116,210,178  atcttagcta ccctaggtaa aggagtgggc tttttagttc ctaaccttcc [959]
116,210,128  agagctacaa cagcagtcat ccagccaggt ctgggtggga acatttttcta
116,210,078  gatacgggtg ctgagatctc tcagcccaga gagaagccct ggggaatttt [894]
116,210,028  cagagagaaa gcagtctcca ggtggggctg gatgtactga tgccactgag
116,209,978  atctgtaaag gagtccctaa cacctgacat aggagtgaca aaactgtttt
116,209,928  ctgcaccaac tgagcagaat acacgcagct gacctgggct caaggtctgg
116,209,878  ccctgccacg tgctggctct gtgatgctgg ccaagtgcct tcgcctctcc [689]; [656]
116,209,828  gggccacagt tttttgatct gaagagtgga gccctactca agccatctgc rs7948159(631)
116,209,778  agctctcggg ctctctgacc tgacatcttt cgggtggtgg ggacacaaag
116,209,728  gaagcagcct ctattggagg accttgtgct tctttttggt cccaggacac rs12721031(533)
116,209,678  tgccccccac cactccagtc cggtcccaa gggcccagtc agctcaactg [477]
116,209,628  taatcatgac aacattgatc aagcatcttt acgtgcaggt gctgtgccaa
116,209,578  acggttcgaa cgctctctca tttcaatctc acggcaaacc tacggtggag [386]
116,209,528  ggggtacggt tgtatccact ttacatgtaa gaaactgagg ctgatatcaa [338]
116,209,478  gtggtggagc caagaatagt gcctcgttgc atcttactcc aacctctagc
116,209,428  ccatccggcc tcctcccttc acgtgcgcct aagagggcta ggtggcctgg rs7123454(206)
116,209,378  ataggggagg tcagctccac agttttgagt aaacacacac agtctcaact
116,209,328  ctgatgacaa cttaagtgcc aggcatagtg gctggcatgg ggcacacact
116,209,278  caagtcatgt tgtgcagcac ctaacagttt atcaaagtat cagcaaactt
116,209,228  attgtcctgt ttgaccttcc gcacaaagct gtcaaggaag gcagggta
```

Figure 11. *APOA4* Annotated Sequence

```
116,200,221  tgtactaaac atatacagac ttttttggtg attacctcct aaacaataca
116,200,171  gtataacaac cattcacaca gcattttccgt cgtattaggt attataagta rs10892035
116,200,121  atctagagat ttaaagtatg tgggaggctg tgcttaggtt atatgcaaat rs35781859
116,200,071  actatgccat tttatatcaa ggacttgaac atccatggat tttggtatct
116,200,021  gcagagggtc ctggaatcaa tttcccatgg agactgagag atgaccgtac
116,199,971  tacccacttc gcaagcaatg tcttctttaa tgtactgaac catcccattg
116,199,921  ttcagaggag aaactgaagc tcagggcttt gaataactag accaaggagg
116,199,871  cacagcatgg gagtgggaca tgaagcactc tacaattaac cctttcagga
116,199,821  caaggccctg tctcccacac ccatctgccc aaaggctctc cagggccccc
116,199,771  tcctcttggg tgtaccttga caagagacct agattttagc tcactatgct
116,199,721  gtctgcagtc ctggatggtc ccactccagt gtctggtgct ctgagatgga [3146]
116,199,671  gtcagcatta gtggcggatg tggagactgg ggggacctgt cttcactggg
116,199,621  gtagacagag gagatgtgga ctttgccccc catgagcccg gcacaaaccc
116,199,571  agagccgcca gcagggcctc gaggcatcag tcccgggctc atgggctccc [2984];rs5089(2981);rs7929134[2978]
116,199,521  tgaggtgttt ctcctactgt tttccgttcc cctcctccct tccatgctga rs1263179
116,199,471  ggttggtggg gtgggggtgg gggtgcccac gcacggaaca gccaccactt rs7926125;rs13306180
116,199,421  ctaactatcg cctgagccct gatctgctgt cagcttccac gtagtctcag rs1263178
116,199,371  ggtcacaaaa gtccaagagg cctcttggga atgtgtcacc ttccagcgtg
116,199,321  gagtcacact gaggaaggag gaggggaggg cagccagggg ggtggcgata [2705]
116,199,271  gggagagagt ttaaatgtct ggctggctct gagcttcagt cagttcccac rs5090(2695);[2685]*

116,199,221  tgcagcgcag gtgagctctc ctgaggacct ctctgtcagc tcccctgatt Exon 1 rs5091(2645)
116,199,171  gtagggagga tccagtgtgg caagaaactc ctccagccca gcaagcagct
116,199,121  caggATGTTC CTGAAGGCCG TGGTCCTGAC CCTGGCCCTG GTGGCTGTCG rs12721041(2511)
116,199,071  CCGGTGAGTA GAAGCTGTCT TTGGATGGCA CTCCTGGGCT GCTGCTCTGA Intron 1
116,199,021  GTAGTGCAGG ATGGAGGCTG AGCCAAAGCA AAAGGACACT TCTGAGTGCC  [2406]*
116,198,971  CATCAGCCCC CAGCTGGACA TGAGGTCTGC CTGGCTGCCA AGTGGCTCAC
116,198,921  AGGAGAGCTG GCCCAGTCCC AGTGGTGGGC CCATTGGCAT TGGTGCTATA [2327]*
116,198,871  CCAGTTTCAC ATATCCCTGT GGCTTCCAAA AAGCTAAGCT CAGACAGGGA [2287];rs13306174
116,198,821  AAATGGCAGG TTGAGGCACC CCCACCATCA TCCAGTCTGC AGCTCAGAGC
116,198,771  TGGAGCAGAG GGGCCACACA GGAGACGGGG CCTCATGAAT TGCTCTCTGT
116,198,721  TACCACCCAG GAGCCAGGGC TGAGGTCAGT GCTGACCAGG TGGCCACGGT Exon 2 rs5092(2104)
116,198,671  GATGTGGGAC TACTTCAGCC AGCTGAGCAA CAATGCCAAG GAGGCCGTGG
116,198,621  AACATCTCCA GAAATCTGAA CTCACCCAGC AACTCAAGTA AGAGGGACTA Intron 2
116,198,571  CAGTGTGCGG TGGTGACGGG GAATTCTTAA AGGCCATGCA ATGTACTGGC rs5093(1994);rs2239013(1993)
116,198,521  AAGGGTTGAG CTTAGAGACA GGAGCCCTGA GCTTAGGATA CCCACTGCCC [1948];rs13306177
116,198,471  TGCCACTAAC TGGCCGGGCC TCTGAACCTA GGATCCACAT ATGTAAACCG rs5094(1853)
116,198,421  GAAGTTTGGA CCGAATAATC CCTGCCATGT CCTTTTGCTT TGACGTTCTA rs13306179;rs5095(1853)
116,198,371  GAGTTTGACA AATGGCCACA TCCTATCATT CAGGCTCATG GAAGAGAGGG
116,198,321  AGGGAGGAAA ATGTCACGTG AGCTGATTTC TAATACGTTT CAGAAAGACA [1743];rs5096(1735)
116,198,271  GGCCCCAGTG GAATCAAGGG GAGGGAGGTG GGAATATTTG GGAGGCCCCT
116,198,221  GGGCACAGGC AAGGAAAGCA GCACCTTGTG CCACTGGAAG ACCCCAGCAG
116,198,171  AGGTCAAGAA GACAACATTG TGTTACACAA TGTGATCCTA TGGCCCAGAA
116,198,121  CACTCCCTCT GGGAAGGACC TCAAAGTCCC ACCCTCTGCA GACAAGGAGG rs2234667
116,198,071  GGAAAGCAAA CTGCTGGAGG TGACATGGTG GGTAGATTCT GAGACAAACT rs5098(1453)
116,198,021  ATGTGGGAGA TCCTGAGATA GAAATTCAGC ATCGTAACTT AGTCTGTGAC
116,197,971  ACCCATCCTC TCCAATCTGC ACCACCATAG GGAGGGTGAA CTCGGTACCT [1371]
116,197,921  CTGAGCACTC ACCTGTCCTA GCACGTGTGC ATAAGGCGAG TGGTATACAA rs5100(1334);[1326]
116,197,871  GCAGACAAAG TCTTGCCGTG TAAATGCCAA ATGTAACGTG GCCTCCTTGT [1274]
116,197,821  GCCCTTCCCC ACAGTGCCCT CTTCCAGGAC AAACTTGGAG AAGTGAACAC Exon 3 rs13306173
116,197,771  TTACGCAGGGT GACCTGCAGA AGAAGCTGGT GCCCTTTGCC ACCGAGCTGC rs5101(1198);rs5102;rs5103(1192);rs12721042(1183)
116,197,721  ATGAACGCCT GGCCAAGGAC TCGGAGAAAC TGAAGGAGGA GATTGGGAAG
116,197,671  GAGCTGGAGG AGCTGAGGGC CCGGCTGCTG CCCCATGCCA ATGAGGTGAG rs6413456
116,197,621  CCAGAAGATC GGGGACAACC TGCGAGAGCT TCAGCAGCGC CTGGAGCCCT [1033]
116,197,571  ACGCGGACCA GCTGCGCACC CAGGTCAGCA CGCAGGCCGA GCAGCTGCGG [952];rs5104(974);rs2234668(964)
116,197,521  CGCCAGCTGA CCCCCTACGC ACAGCGCATG GAGAGAGTGC TGCGGGAGAA [945];(933)
116,197,471  CGCCGACAGC CTGCAGGCCT CGCTGAGGCC CCACGCCGAC GAGCTCAAGG
116,197,421  CCAAGATCGA CCAGAACGTG GAGGAGCTCA AGGGACGCCT TACGCCCTAC
116,197,371  GCTGACGAAT TCAAAGTCAA GATTGACCAG ACCGTGGAGG AGCTGCGCCG [755]
116,197,321  CAGCCTGGCT CCCTATGCTC AGGACACGCA GGAGAAGCTC AACCACCAGC
116,197,271  TTGAGGGCCT GACCTTCCAG ATGAAGAAGA ACGCCGAGGA GCTCAAGGCC
116,197,221  AGGATCTCGG CCAGTGCCGA GGAGCTGCGG CAGAGGCTGG CGCCCTTGGC rs5105[634];rs2238008
116,197,171  CGAGGACGTG CGTGGCAACC TGAGGGGCAA CACCGAGGGG CTGCAGAAGT rs1042372;rs5106(568)
116,197,121  CACTGGCAGA GCTGGGTGGG CACCTGGACC AGCAGGTGGA GGAGTTCCGA rs5107;[520]
116,197,071  CGCCGGGTGG AGCCCTACGG GGAAAACTTC AACAAAGCCC TGGTGCAGCA rs5108
116,197,021  GATGGAACAG CTCAGGCAGA AACTGGGCCC CCATGCGGGG GACGTGGAAG [422]*;rs5109(406)
116,196,971  GCCACTTGAG CTTCCTGGAG AAGGACCTGA GGGACAAGGT CAACTCCTTC [357]
116,196,921  TTCAGCACCT TCAAGGAGAA AGAGAGCCAG GACAAGACTC TCTCCCTCCC rs675(315)
116,196,871  TGAGCTGGAG CAACAGCAGG AACAGCAGCA GGAGCAGCAG CAGGAGCAGG [288];rs5030782;rs5110(274)
116,196,821  TGCAGATGCT GGCCCCTTTG GAGAGCTGAg ctgcccctgg tgcactggcc
116,196,771  ccaccctcgt ggacacctgc cctgccctgc cacctgtctg tctgtctgtc (165)
116,196,721  ccaaagaagt tctggtatga acttgaggac acatgtccag tgggaggtga rs12721040(120)
116,196,671  gaccacctct caatattcaa taaagctgct gagaatctag cctc
```

Figure 11 Continued

3.2 DISTRIBUTION OF *APOA1* AND *APOA4* VARIANTS IN HIGH AND LOW HDL GROUPS

3.2.1 *APOA1*

3.2.1.1 Non-Hispanic Whites

Of 34 variants identified in NHWs, 11 had MAF >5%, 13 had MAF 1-5%, and 10 had MAF <1%. All 10 new variants had a MAF of ≤1%. No statistically significant difference was identified when comparing the allele frequencies between the high HDL and low HDL groups for any of the 34 variants in this small sequencing sample set (Tables 18 and 19). Of 23 rare variants, 7 were present only in the high HDL group versus 5 only in the low HDL group. Of 3 exonic variants identified in NHWs, 2 were present only in the low HDL group versus 1 only in the high HDL group. Of 2 nonsynonymous variants identified in NHWs, 1 was present only in the low HDL group and 1 only in the high HDL group. Of the 48 individuals with low HDL

levels, 5 (10.42%) had rare variants unique to the low HDL group. Of the 47 individuals with high HDL levels 7 (14.89%) had rare variants unique to the high HDL group.

Table 18. Distribution of *APOA1* Variants in High and Low HDL Groups in NHWs

	Non-Hispanic Whites*	
Rare (MAF<5%)	**High HDL (n=47)**	**Low HDL (n=48)**
533C>T	0.021	0.010
689C>T	-	0.010
959G>C	0.011	-
1049T>A	0.021	-
1407insT	0.011	-
1507T>C	0.011	-
1549C>T	-	0.010
1749T>C	0.043	0.031
2077G>A	0.043	0.031
2198T>G	0.043	0.031
2373T>C	0.053	0.031
2376A>T	0.021	-
2652C>A**	-	0.010
3220G>A	0.053	0.031
3307C>A	0.011	0.010
3431G>A	0.032	0.021
3769A>C	0.011	-
3959G>T	0.011	-
4151G>C	-	0.010
4283C>T	-	0.010
4284G>A	0.043	0.021
4693T>G	0.053	0.031
5131C>T	0.011	0.010
Common (MAF≥5%)		
206A>C	**0.149**	**0.198**
1128G>T	0.196	0.188
1308C>T	0.250	0.219
1546A>G	0.370	0.375
1598T>G	**0.096**	**0.167**
1620A>G	0.191	0.177
3368G>A	0.149	0.146
3613G>A	0.096	0.115
3714G>A	0.096	0.115
4050G>A	0.340	0.323
4443C>T	0.191	0.188

* The locations and nucleotide changes are based on the reverse strand sequence used in the SeattleSNPs database.
** Suspicious variants with low sequence quality.
Novel variants are hi-lighted.

Table 19. Allele Frequencies of *APOA1* Variants in High and Low HDL Groups in NHWs

206 A>C						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
AA	33	(70.21)	31	(64.58)	64	67.37
AC	14	(29.79)	15	(31.25)	29	30.53
CC	0	(0.00)	2	(4.17)	2	2.11
	47		48		95	
A	0.851		0.802		0.826	
C	0.149		0.198		0.174	
Z	0.894					
p	0.371	*test for allele frequencies				

533 C>T						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	45	(95.74)	47	(97.92)	92	96.84
CT	2	(4.26)	1	(2.08)	3	3.16
TT	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
C	0.979		0.990		0.984	
T	0.021		0.010		0.016	
Z	0.599					
p	0.549	*test for allele frequencies				

689 C>T						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	47	(100.00)	47	(97.92)	94	98.95
CT	0	(0.00)	1	(2.08)	1	1.05
TT	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
C	1.000		0.990		0.995	
T	0.000		0.010		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

959 G>C						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	46	(97.87)	48	(100.00)	94	98.95
GC	1	(2.13)	0	(0.00)	1	1.05
CC	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
G	0.989		1.000		0.995	
C	0.011		0.000		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

1049 T>A						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
TT	45	(95.74)	48	(100.00)	93	97.89
TA	2	(4.26)	0	(0.00)	2	2.11
AA	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
T	0.979		1.000		0.989	
A	0.021		0.000		0.011	
Z	1.430					
p	0.153	*test for allele frequencies				

1128 G>T						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	28	(60.87)	32	(66.67)	60	63.83
GT	18	(39.13)	14	(29.17)	32	34.04
TT	0	(0.00)	2	(4.17)	2	2.13
	46		48		94	
G	0.804		0.813		0.809	
T	0.196		0.188		0.191	
Z	0.142					
p	0.887	*test for allele frequencies				

1308 C>T						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	24	(52.17)	30	(62.50)	54	57.45
CT	21	(45.65)	15	(31.25)	36	38.30
TT	1	(2.17)	3	(6.25)	4	4.26
	46		48		94	
C	0.750		0.781		0.766	
T	0.250		0.219		0.234	
Z	0.506					
p	0.613	*test for allele frequencies				

1407 ins T						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
WW	46	(97.87)	48	(100.00)	94	98.95
WI	1	(2.13)	0	(0.00)	1	1.05
II	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
W	0.989		1.000		0.995	
II	0.011		0.000		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

Table 19 (Continued)

1507 T>C	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
TT	46	(97.87)	48	(100.00)	94	98.95
TC	1	(2.13)	0	(0.00)	1	1.05
CC	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
T	0.989		1.000		0.995	
C	0.011		0.000		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

1546 A>G	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
AA	18	(39.13)	19	(39.58)	37	39.36
AG	22	(47.83)	22	(45.83)	44	46.81
GG	6	(13.04)	7	(14.58)	13	13.83
	46		48		94	
A	0.630		0.625		0.628	
G	0.370		0.375		0.372	
Z	0.077					
p	0.939	*test for allele frequencies				

1549 C>T	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	47	(100.00)	47	(97.92)	94	98.95
CT	0	(0.00)	1	(2.08)	1	1.05
TT	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
C	1.000		0.990		0.995	
T	0.000		0.010		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

1598 T>G	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
TT	38	(80.85)	34	(70.83)	72	75.79
TG	9	(19.15)	12	(25.00)	21	22.11
GG	0	(0.00)	2	(4.17)	2	2.11
	47		48		95	
T	0.904		0.833		0.868	
G	0.096		0.167		0.132	
Z	1.458					
p	0.145	*test for allele frequencies				

1620 A>G	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
AA	29	(61.70)	33	(68.75)	62	65.26
AG	18	(38.30)	13	(27.08)	31	32.63
GG	0	(0.00)	2	(4.17)	2	2.11
	47		48		95	
A	0.809		0.823		0.816	
G	0.191		0.177		0.184	
Z	0.256					
p	0.798	*test for allele frequencies				

1749 T>C	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
TT	43	(91.49)	45	(93.75)	88	92.63
TC	4	(8.51)	3	(6.25)	7	7.37
CC	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
T	0.957		0.969		0.963	
CC	0.043		0.031		0.037	
Z	0.413					
p	0.680	*test for allele frequencies				

2077 G>A	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	43	(91.49)	45	(93.75)	88	92.63
GA	4	(8.51)	3	(6.25)	7	7.37
AA	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
G	0.957		0.969		0.963	
A	0.043		0.031		0.037	
Z	0.413					
p	0.680	*test for allele frequencies				

2198 T>G	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
TT	43	(91.49)	45	(93.75)	88	92.63
TG	4	(8.51)	3	(6.25)	7	7.37
GG	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
T	0.957		0.969		0.963	
G	0.043		0.031		0.037	
Z	0.413					
p	0.680	*test for allele frequencies				

Table 19 (Continued)

2373 T>C						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
TT	42	(89.36)	45	(93.75)	87	91.58
TC	5	(10.64)	3	(6.25)	8	8.42
CC	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
T	0.947		0.969		0.958	
C	0.053		0.031		0.042	
Z	0.752					
p	0.452	*test for allele frequencies				

2376 A>T						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
AA	45	(95.74)	48	(100.00)	93	97.89
AT	2	(4.26)	0	(0.00)	2	2.11
TT	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
A	0.979		1.000		0.989	
T	0.021		0.000		0.011	
Z	1.430					
p	0.153	*test for allele frequencies				

2652* C>A						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	47	(100.00)	47	(97.92)	94	98.95
CA	0	(0.00)	1	(2.08)	1	1.05
AA	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
C	1.000		0.990		0.995	
A	0.000		0.010		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

3220 G>A						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	42	(89.36)	45	(93.75)	87	91.58
GA	5	(10.64)	3	(6.25)	8	8.42
AA	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
G	0.947		0.969		0.958	
A	0.053		0.031		0.042	
Z	0.752					
p	0.452	*test for allele frequencies				

3307 C>A						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	46	(97.87)	47	(97.92)	93	97.89
CA	1	(2.13)	1	(2.08)	2	2.11
AA	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
C	0.989		0.990		0.989	
A	0.011		0.010		0.011	
Z	0.015					
p	0.988	*test for allele frequencies				

3368 G>A						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	33	(70.21)	36	(75.00)	69	72.63
GA	14	(29.79)	10	(20.83)	24	25.26
AA	0	(0.00)	2	(4.17)	2	2.11
	47		48		95	
G	0.851		0.854		0.853	
A	0.149		0.146		0.147	
Z	0.060					
p	0.952	*test for allele frequencies				

3431 G>A						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	44	(93.62)	46	(95.83)	90	94.74
GA	3	(6.38)	2	(4.17)	5	5.26
AA	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
G	0.968		0.979		0.974	
A	0.032		0.021		0.026	
Z	0.476					
p	0.634	*test for allele frequencies				

3613 G>A						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	38	(80.85)	39	(81.25)	77	81.05
GA	9	(19.15)	7	(14.58)	16	16.84
AA	0	(0.00)	2	(4.17)	2	2.11
	47		48		95	
G	0.904		0.885		0.895	
A	0.096		0.115		0.105	
Z	0.424					
p	0.672	*test for allele frequencies				

Table 19 (Continued)

3714 G>A	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	38	(80.85)	39	(81.25)	77	81.05
GA	9	(19.15)	7	(14.58)	16	16.84
AA	0	(0.00)	2	(4.17)	2	2.11
	47		48		95	
G	0.904		0.885		0.895	
A	0.096		0.115		0.105	
Z	0.424					
p	0.672	*test for allele frequencies				

3769 A>C	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
AA	46	(97.87)	48	(100.00)	94	98.95
AC	1	(2.13)	0	(0.00)	1	1.05
CC	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
A	0.989		1.000		0.995	
C	0.011		0.000		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

3959 G>T	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	46	(97.87)	48	(100.00)	94	98.95
GT	1	(2.13)	0	(0.00)	1	1.05
TT	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
G	0.989		1.000		0.995	
T	0.011		0.000		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

4050 G>A	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	20	(42.55)	24	(50.00)	44	46.32
GA	22	(46.81)	17	(35.42)	39	41.05
AA	5	(10.64)	7	(14.58)	12	12.63
	47		48		95	
G	0.660		0.677		0.668	
A	0.340		0.323		0.332	
Z	0.256					
p	0.798	*test for allele frequencies				

4151 G>C	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	47	(100.00)	47	(97.92)	94	98.95
GC	0	(0.00)	1	(2.08)	1	1.05
CC	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
G	1.000		0.990		0.995	
C	0.000		0.010		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

4283 C>T	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	47	(100.00)	47	(97.92)	94	98.95
CT	0	(0.00)	1	(2.08)	1	1.05
TT	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
C	1.000		0.990		0.995	
T	0.000		0.010		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

4284 G>A	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	43	(91.49)	46	(95.83)	89	93.68
GA	4	(8.51)	2	(4.17)	6	6.32
AA	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
G	0.957		0.979		0.968	
A	0.043		0.021		0.032	
Z	0.855					
p	0.393	*test for allele frequencies				

4443 C>T	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	29	(61.70)	32	(66.67)	61	64.21
CT	18	(38.30)	14	(29.17)	32	33.68
TT	0	(0.00)	2	(4.17)	2	2.11
	47		48		95	
C	0.809		0.813		0.811	
T	0.191		0.188		0.189	
Z	0.070					
p	0.944	*test for allele frequencies				

Table 19 (Continued)

4693 T>G						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
TT	42	(89.36)	45	(93.75)	87	91.58
TG	5	(10.64)	3	(6.25)	8	8.42
GG	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
T	0.947		0.969		0.958	
G	0.053		0.031		0.042	
Z	0.752					
p	0.452	*test for allele frequencies				

5131 C>T						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	46	(97.87)	47	(97.92)	93	97.89
CT	1	(2.13)	1	(2.08)	2	2.11
TT	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
C	0.989		0.990		0.989	
T	0.011		0.010		0.011	
Z	0.015					
p	0.988	*test for allele frequencies				

3.2.1.2 Blacks

Of 37 variants identified in Blacks, 20 had MAF >5%, 10 had MAF 1-5%, and 7 had MAF <1%. Of the 15 new variants, 1 had MAF >5%, 7 had MAF 1-5%, and 7 had MAF <1%. No statistically significant difference was identified when comparing the allele frequencies between the high HDL and low HDL groups for any of the 37 variants in this small sequencing sample set (Tables 20 and 21). Of 17 rare variants, 5 were present only in the high HDL group versus 4 only in the low HDL group. Of 3 exonic variants identified in Blacks, 1 was present only in the low HDL group versus 1 only in the high HDL group. Of 2 nonsynonymous variants identified in Blacks, 1 was present only in the low HDL group and 1 only in the high HDL group. Of the 47 individuals with low HDL levels, 2 (4.26%) had rare variants unique to the low HDL group. Of the 48 individuals with high HDL levels 6 (12.5%) had rare variants unique to the high HDL group.

Table 20. Distribution of *APOA1* Variants in High and Low HDL Groups in Blacks

Blacks*		
Rare (MAF<5%)	**High HDL (n=47)**	**Low HDL (n=48)**
338A>G	0.021	0.043
386G>A	0.010	-
477C>T	0.010	0.021
656C>T	-	0.011
894G>A	0.021	0.011
1143G>T	0.010	-
1308C>T	0.022	-
1407delT	0.031	0.032
1965T>C	0.031	0.032
2120C>A	0.021	-
2215C>A	-	0.011
2626G>C	0.010	0.021
2880C>G	-	0.011
3867G>T	0.010	-
4208C>T	-	0.011
4987T>C	0.042	0.011
5066G>T	0.010	0.011
Common (MAF≥5%)		
206C>A	0.351	0.348
631A>G	0.490	0.478
1049T>A	0.128	0.128
1128G>T	0.104	0.085
1546G>A	0.073	0.085
1598T>G	0.104	0.064
1620G>A	0.302	0.298
2373C>T	0.406	0.372
2376A>T	0.125	0.128
3220G>A	0.448	0.426
3368A>G	0.344	0.372
3543C>T	0.052	0.064
3613G>A	0.125	0.074
3714G>A	0.117	0.074
4050G>A	0.448	0.424
4284G>A	0.448	0.426
4443C>T	0.125	0.117
4732C>A	0.096	0.117
4807C>T	0.073	0.064
5055A>T	0.073	0.064

* The locations and nucleotide changes are based on the reverse strand sequence used in the SeattleSNPs database. Novel variants are hi-lighted.

Table 21. Allele Frequencies of *APOA1* Variants in High and Low HDL Groups in Blacks

206 C>A	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	20	(42.55)	19	(41.30)	39	41.94
AC	21	(44.68)	22	(47.83)	43	46.24
AA	6	(12.77)	5	(10.87)	11	11.83
	47		46		93	
C	0.649		0.652		0.651	
A	0.351		0.348		0.349	
Z	0.046					
p	0.963	*test for allele frequencies				

338 A>G	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
AA	45	(95.74)	43	(91.49)	88	93.62
AG	2	(4.26)	4	(8.51)	6	6.38
GG	0	(0.00)	0	(0.00)	0	0.00
	47		47		94	
A	0.979		0.957		0.968	
G	0.021		0.043		0.032	
Z	0.831					
p	0.406	*test for allele frequencies				

386 G>A	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	47	(97.92)	47	(100.00)	94	98.95
GA	1	(2.08)	0	(0.00)	1	1.05
AA	0	(0.00)	0	(0.00)	0	0.00
	48		47		95	
G	0.990		1.000		0.995	
A	0.010		0.000		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

477 C>T	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	47	(97.92)	45	(95.74)	92	96.84
CT	1	(2.08)	2	(4.26)	3	3.16
TT	0	(0.00)	0	(0.00)	0	0.00
	48		47		95	
C	0.990		0.979		0.984	
T	0.010		0.021		0.016	
Z	0.599					
p	0.549	*test for allele frequencies				

631 A>G	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
AA	13	(27.08)	10	(21.74)	23	24.47
AG	23	(47.92)	28	(60.87)	51	54.26
GG	12	(25.00)	8	(17.39)	20	21.28
	48		46		94	
A	0.510		0.522		0.516	
G	0.490		0.478		0.484	
Z	0.155					
p	0.877	*test for allele frequencies				

656 C>T	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	48	(100.00)	46	(97.87)	94	98.95
CT	0	(0.00)	1	(2.13)	1	1.05
TT	0	(0.00)	0	(0.00)	0	0.00
	48		47		95	
C	1.000		0.989		0.995	
T	0.000		0.011		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

894 G>A	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	46	(95.83)	46	(97.87)	92	96.84
GA	2	(4.17)	1	(2.13)	3	3.16
AA	0	(0.00)	0	(0.00)	0	0.00
	48		47		95	
G	0.979		0.989		0.984	
A	0.021		0.011		0.016	
Z	0.566					
p	0.571	*test for allele frequencies				

1049 T>A	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
TT	36	(76.60)	35	(74.47)	71	75.53
TA	10	(21.28)	12	(25.53)	22	23.40
AA	1	(2.13)	0	(0.00)	1	1.06
	47		47		94	
T	0.872		0.872		0.872	
A	0.128		0.128		0.128	
Z	0.000					
p	1.000	*test for allele frequencies				

Table 21 (Continued)

1128 G>T	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	38	(79.17)	39	(82.98)	77	81.05
GT	10	(20.83)	8	(17.02)	18	18.95
TT	0	(0.00)	0	(0.00)	0	0.00
	48		47		95	
G	0.896		0.915		0.905	
T	0.104		0.085		0.095	
Z	0.449					
p	0.653	*test for allele frequencies				

1143 G>T	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	47	(97.92)	47	(100.00)	94	98.95
GT	1	(2.08)	0	(0.00)	1	1.05
TT	0	(0.00)	0	(0.00)	0	0.00
	48		47		95	
G	0.990		1.000		0.995	
T	0.010		0.000		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

1308 C>T	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	44	(95.65)	44	(100.00)	88	97.78
CT	2	(4.35)	0	(0.00)	2	2.22
TT	0	(0.00)	0	(0.00)	0	0.00
	46		44		90	
C	0.978		1.000		0.989	
T	0.022		0.000		0.011	
Z	1.430					
p	0.153	*test for allele frequencies				

1407 del T	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
WW	45	(93.75)	44	(93.62)	89	93.68
WD	3	(6.25)	3	(6.38)	6	6.32
DD	0	(0.00)	0	(0.00)	0	0.00
	48		47		95	
W	0.969		0.968		0.968	
D	0.031		0.032		0.032	
Z	0.026					
p	0.979	*test for allele frequencies				

1546 G>A	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	41	(85.42)	39	(82.98)	80	84.21
AG	7	(14.58)	8	(17.02)	15	15.79
AA	0	(0.00)	0	(0.00)	0	0.00
	48		47		95	
G	0.927		0.915		0.921	
A	0.073		0.085		0.079	
Z	0.311					
p	0.756	*test for allele frequencies				

1598 T>G	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
TT	38	(79.17)	41	(87.23)	79	83.16
TG	10	(20.83)	6	(12.77)	16	16.84
GG	0	(0.00)	0	(0.00)	0	0.00
	48		47		95	
T	0.896		0.936		0.916	
G	0.104		0.064		0.084	
Z	1.006					
p	0.314	*test for allele frequencies				

1620 G>A	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	22	(45.83)	22	(46.81)	44	46.32
AG	23	(47.92)	22	(46.81)	45	47.37
AA	3	(6.25)	3	(6.38)	6	6.32
	48		47		95	
G	0.698		0.702		0.700	
A	0.302		0.298		0.300	
Z	0.063					
p	0.950	*test for allele frequencies				

1965 T>C	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
TT	45	(93.75)	44	(93.62)	89	93.68
TC	3	(6.25)	3	(6.38)	6	6.32
CC	0	(0.00)	0	(0.00)	0	0.00
	48		47		95	
T	0.969		0.968		0.968	
C	0.031		0.032		0.032	
Z	0.026					
p	0.979	*test for allele frequencies				

Table 21 (Continued)

2120 C>A	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	46	(95.83)	47	(100.00)	93	97.89
CA	2	(4.17)	0	(0.00)	2	2.11
AA	0	(0.00)	0	(0.00)	0	0.00
	48		47		95	
C	0.979		1.000		0.989	
A	0.021		0.000		0.011	
Z	1.429					
p	0.153	*test for allele frequencies				

2215 C>A	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	48	(100.00)	46	(97.87)	94	98.95
CA	0	(0.00)	1	(2.13)	1	1.05
AA	0	(0.00)	0	(0.00)	0	0.00
	48		47		95	
C	1.000		0.989		0.995	
A	0.000		0.011		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

2373 C>T	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	18	(37.50)	15	(31.91)	33	34.74
TC	21	(43.75)	29	(61.70)	50	52.63
TT	9	(18.75)	3	(6.38)	12	12.63
	48		47		95	
C	0.594		0.628		0.611	
T	0.406		0.372		0.389	
Z	0.480					
p	0.632	*test for allele frequencies				

2376 A>T	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
AA	37	(77.08)	35	(74.47)	72	75.79
AT	10	(20.83)	12	(25.53)	22	23.16
TT	1	(2.08)	0	(0.00)	1	1.05
	48		47		95	
A	0.875		0.872		0.874	
T	0.125		0.128		0.126	
Z	0.055					
p	0.956	*test for allele frequencies				

2626 G>C	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	47	(97.92)	45	(95.74)	92	96.84
GC	1	(2.08)	2	(4.26)	3	3.16
CC	0	(0.00)	0	(0.00)	0	0.00
	48		47		95	
G	0.990		0.979		0.984	
C	0.010		0.021		0.016	
Z	0.599					
p	0.549	*test for allele frequencies				

2880 C>G	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	48	(100.00)	46	(97.87)	94	98.95
CG	0	(0.00)	1	(2.13)	1	1.05
GG	0	(0.00)	0	(0.00)	0	0.00
	48		47		95	
C	1.000		0.989		0.995	
G	0.000		0.011		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

3220 G>A	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	16	(33.33)	13	(27.66)	29	30.53
GA	21	(43.75)	28	(59.57)	49	51.58
AA	11	(22.92)	6	(12.77)	17	17.89
	48		47		95	
G	0.552		0.574		0.563	
A	0.448		0.426		0.437	
Z	0.311					
p	0.756	*test for allele frequencies				

3368 A>G	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
AA	22	(45.83)	16	(34.04)	38	40.00
AG	19	(39.58)	27	(57.45)	46	48.42
GG	7	(14.58)	4	(8.51)	11	11.58
	48		47		95	
A	0.656		0.628		0.642	
G	0.344		0.372		0.358	
Z	0.411					
p	0.681	*test for allele frequencies				

Table 21 (Continued)

3543 C>T						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	43	(89.58)	41	(87.23)	84	88.42
CT	5	(10.42)	6	(12.77)	11	11.58
TT	0	(0.00)	0	(0.00)	0	0.00
	48		47		95	
C	0.948		0.936		0.942	
T	0.052		0.064		0.058	
Z	0.346					
p	0.729	*test for allele frequencies				

3613 G>A						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	37	(77.08)	40	(85.11)	77	81.05
GA	10	(20.83)	7	(14.89)	17	17.89
AA	1	(2.08)	0	(0.00)	1	1.05
	48		47		95	
G	0.875		0.926		0.900	
A	0.125		0.074		0.100	
Z	1.168					
p	0.243	*test for allele frequencies				

3714 G>A						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	37	(78.72)	40	(85.11)	77	81.91
GA	9	(19.15)	7	(14.89)	16	17.02
AA	1	(2.13)		(0.00)	1	1.06
	47		47		94	
G	0.883		0.926		0.904	
A	0.117		0.074		0.096	
Z	0.994					
p	0.320	*test for allele frequencies				

3867 G>T						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	47	(97.92)	47	(100.00)	94	98.95
GT	1	(2.08)	0	(0.00)	1	1.05
TT	0	(0.00)	0	(0.00)	0	0.00
	48		47		95	
G	0.990		1.000		0.995	
T	0.010		0.000		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

4050 G>A						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	15	(31.25)	13	(28.26)	28	29.79
GA	23	(47.92)	27	(58.70)	50	53.19
AA	10	(20.83)	6	(13.04)	16	17.02
	48		46		94	
G	0.552		0.576		0.564	
A	0.448		0.424		0.436	
Z	0.332					
p	0.740	*test for allele frequencies				

4208 C>T						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	48	(100.00)	46	(97.87)	94	98.95
CT	0	(0.00)	1	(2.13)	1	1.05
TT	0	(0.00)	0	(0.00)	0	0.00
	48		47		95	
C	1.000		0.989		0.995	
T	0.000		0.011		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

4284 G>A						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	16	(33.33)	13	(27.66)	29	30.53
GA	21	(43.75)	28	(59.57)	49	51.58
AA	11	(22.92)	6	(12.77)	17	17.89
	48		47		95	
G	0.552		0.574		0.563	
A	0.448		0.426		0.437	
Z	0.311					
p	0.756	*test for allele frequencies				

4443 C>T						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	37	(77.08)	36	(76.60)	73	76.84
CT	10	(20.83)	11	(23.40)	21	22.11
TT	1	(2.08)	0	(0.00)	1	1.05
	48		47		95	
C	0.875		0.883		0.879	
T	0.125		0.117		0.121	
Z	0.169					
p	0.866	*test for allele frequencies				

Table 21 (Continued)

4732 C>A						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	39	(82.98)	36	(76.60)	75	79.79
CA	7	(14.89)	11	(23.40)	18	19.15
AA	1	(2.13)	0	(0.00)	1	1.06
	47		47		94	
C	0.904		0.883		0.894	
A	0.096		0.117		0.106	
Z	0.473					
p	0.636	*test for allele frequencies				

4807 C>T						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	41	(85.42)	41	(87.23)	82	86.32
CT	7	(14.58)	6	(12.77)	13	13.68
TT	0	(0.00)	0	(0.00)	0	0.00
	48		47		95	
C	0.927		0.936		0.932	
T	0.073		0.064		0.068	
Z	0.248					
p	0.804	*test for allele frequencies				

4987 T>C						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
TT	44	(91.67)	46	(97.87)	90	94.74
TC	4	(8.33)	1	(2.13)	5	5.26
CC	0	(0.00)	0	(0.00)	0	0.00
	48		47		95	
T	0.958		0.989		0.974	
C	0.042		0.011		0.026	
Z	1.350					
p	0.177	*test for allele frequencies				

5055 A>T						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
AA	41	(85.42)	41	(87.23)	82	86.32
AT	7	(14.58)	6	(12.77)	13	13.68
TT	0	(0.00)	0	(0.00)	0	0.00
	48		47		95	
A	0.927		0.936		0.932	
T	0.073		0.064		0.068	
Z	0.248					
p	0.804	*test for allele frequencies				

5066 G>T						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	47	(97.92)	46	(97.87)	93	97.89
GT	1	(2.08)	1	(2.13)	2	2.11
TT	0	(0.00)	0	(0.00)	0	0.00
	48		47		95	
G	0.990		0.989		0.989	
T	0.010		0.011		0.011	
Z	0.015					
p	0.988	*test for allele frequencies				

3.2.2 *APOA4*

3.2.2.1 Non-Hispanic Whites

Of 23 variants identified in NHWs, 11 had MAF >5%, 5 had MAF 1-5%, and 7 had MAF <1%. All 7 of the new variants had MAF MAF <1%. No statistically significant difference was identified when comparing the allele frequencies between the high HDL and low HDL groups for any of the 23 variants in this small sequence sample set (Tables 22 and 23). Of 12 rare variants, 4 were present only in the high HDL versus 4 only in the low HDL group. Of 13 exonic variants identified in NHWs, 3 were present only in the low HDL group versus 2 only in the high HDL group. Of 5 nonsynonymous variants identified in NHWs, 2 were present only in the low HDL group versus none only in the high HDL group. Of the 48 individuals with low HDL levels, 3 (6.25%) had rare variants unique to the low HDL group. Of the 47 individuals with high HDL levels 5 (10.64%) had rare variants unique to the high HDL group.

Table 22. Distribution of *APOA4* Variants in High and Low HDL Groups in NHWs

	Non-Hispanic Whites*	
Rare (MAF<5%)	**High HDL (n=47)**	**Low HDL (n=48)**
120G>A	0.011	0.031
422G>T**	0.011	-
520C>T	-	0.010
945G>A	-	0.010
952C>T	0.011	-
1033G>T	-	0.010
1853G>A	0.021	-
1993C>T	0.053	0.031
1994G>A	0.022	0.042
2287G>A	0.011	-
2978C>A	0.021	0.021
2984G>A	-	0.010
Common (MAF≥5%)		
165delACAG	0.511	0.479
274C>A	0.085	0.094
315T>A	0.202	0.188
964G>A	**0.043**	**0.073**
974T>C	0.149	0.177
1192A>G	0.064	0.042
1334A>G	0.415	0.406
1735A>G	0.415	0.406
1803A>G	0.202	0.188
2104T>C	0.213	0.219
2695C>G	0.085	0.052

* The locations and nucleotide changes are based on the reverse strand sequence used in the SeattleSNPs database.
** Suspicious variants with low sequence quality.
Novel variants are hi-lighted.

Table 23. Allele Frequencies of *APOA4* Variants in High and Low HDL Groups in NHWs

120 G>A	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	46	(97.87)	45	(93.75)	91	95.79
GA	1	(2.13)	3	(6.25)	4	4.21
AA	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
G	0.989		0.969		0.979	
A	0.011		0.031		0.021	
Z	0.997					
p	0.319	*test for allele frequencies				

165 del ACAG	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
DD	11	(23.40)	14	(29.17)	25	26.32
DW	24	(51.06)	22	(45.83)	46	48.42
WW	12	(25.53)	12	(25.00)	24	25.26
	47		48		95	
	0.489		0.521		0.505	
	0.511		0.479		0.495	
Z	0.434					
p	0.664	*test for allele frequencies				

274 C>A	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	39	(82.98)	39	(81.25)	78	82.11
CA	8	(17.02)	9	(18.75)	17	17.89
AA	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
C	0.915		0.906		0.911	
A	0.085		0.094		0.089	
Z	0.209					
p	0.835	*test for allele frequencies				

315 T>A	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
TT	29	(61.70)	31	(64.58)	60	63.16
TA	17	(36.17)	16	(33.33)	33	34.74
AA	1	(2.13)	1	(2.08)	2	2.11
	47		48		95	
	0.798		0.813		0.805	
	0.202		0.188		0.195	
Z	0.255					
p	0.799	*test for allele frequencies				

422* G>T	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	46	(97.87)	48	(100.00)	94	98.95
GT	1	(2.13)	0	(0.00)	1	1.05
TT	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
G	0.989		1.000		0.995	
T	0.011		0.000		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

520 C>T	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	47	(100.00)	47	(97.92)	94	98.95
CT	0	(0.00)	1	(2.08)	1	1.05
TT	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
C	1.000		0.990		0.995	
T	0.000		0.010		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

945 G>A	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	47	(100.00)	47	(97.92)	94	98.95
GA	0	(0.00)	1	(2.08)	1	1.05
AA	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
G	1.000		0.990		0.995	
A	0.000		0.010		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

952 C>T	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	46	(97.87)	48	(100.00)	94	98.95
CT	1	(2.13)	0	(0.00)	1	1.05
TT	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
C	0.989		1.000		0.995	
T	0.011		0.000		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

Table 23 (Continued)

964 G>A						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	43	(91.49)	41	(85.42)	84	88.42
GA	4	(8.51)	7	(14.58)	11	11.58
AA	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
G	0.957		0.927		0.942	
A	0.043		0.073		0.058	
Z	0.900					
p	0.368	*test for allele frequencies				

974 T>C						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
TT	33	(70.21)	33	(68.75)	66	69.47
TC	14	(29.79)	13	(27.08)	27	28.42
CC	0	(0.00)	2	(4.17)	2	2.11
	47		48		95	
T	0.851		0.823		0.837	
C	0.149		0.177		0.163	
Z	0.526					
p	0.599	*test for allele frequencies				

1033 G>T						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	47	(100.00)	47	(97.92)	94	98.95
GT	0	(0.00)	1	(2.08)	1	1.05
TT	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
G	1.000		0.990		0.995	
T	0.000		0.010		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

1192 A>G						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
AA	41	(87.23)	44	(91.67)	85	89.47
AG	6	(12.77)	4	(8.33)	10	10.53
GG	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
A	0.936		0.958		0.947	
G	0.064		0.042		0.053	
Z	0.683					
p	0.494	*test for allele frequencies				

1334 A>G						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
AA	16	(34.04)	18	(37.50)	34	35.79
AG	23	(48.94)	21	(43.75)	44	46.32
GG	8	(17.02)	9	(18.75)	17	17.89
	47		48		95	
A	0.585		0.594		0.589	
G	0.415		0.406		0.411	
Z	0.121					
p	0.904	*test for allele frequencies				

1735 A>G						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
AA	16	(34.04)	18	(37.50)	34	35.79
AG	23	(48.94)	21	(43.75)	44	46.32
GG	8	(17.02)	9	(18.75)	17	17.89
	47		48		95	
A	0.585		0.594		0.589	
G	0.415		0.406		0.411	
Z	0.121					
p	0.904	*test for allele frequencies				

1803 A>G						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
AA	29	(61.70)	31	(64.58)	60	63.16
AG	17	(36.17)	16	(33.33)	33	34.74
GG	1	(2.13)	1	(2.08)	2	2.11
	47		48		95	
A	0.798		0.813		0.805	
G	0.202		0.188		0.195	
Z	0.255					
p	0.799	*test for allele frequencies				

1853 G>A						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	45	(95.74)	48	(100.00)	93	97.89
GA	2	(4.26)	0	(0.00)	2	2.11
AA	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
G	0.979		1.000		0.989	
A	0.021		0.000		0.011	
Z	1.430					
p	0.153	*test for allele frequencies				

Table 23 (Continued)

1993 C>T						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	42	(89.36)	45	(93.75)	87	91.58
CT	5	(10.64)	3	(6.25)	8	8.42
TT	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
C	0.947		0.969		0.958	
T	0.053		0.031		0.042	
Z	0.752					
p	0.452	*test for allele frequencies				

1994 G>A						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	44	(95.65)	44	(91.67)	88	93.62
GA	2	(4.35)	4	(8.33)	6	6.38
AA	0	(0.00)	0	(0.00)	0	0.00
	46		48		94	
G	0.978		0.958		0.968	
A	0.022		0.042		0.032	
Z	0.783					
p	0.433	*test for allele frequencies				

2104 T>C						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
TT	28	(59.57)	30	(62.50)	58	61.05
TC	18	(38.30)	15	(31.25)	33	34.74
CC	1	(2.13)	3	(6.25)	4	4.21
	47		48		95	
T	0.787		0.781		0.784	
C	0.213		0.219		0.216	
Z	0.100					
p	0.920	*test for allele frequencies				

2287 G>A						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	46	(97.87)	48	(100.00)	94	98.95
GA	1	(2.13)	0	(0.00)	1	1.05
AA	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
G	0.989		1.000		0.995	
A	0.011		0.000		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

2695 C>G						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	39	(82.98)	43	(89.58)	82	86.32
CG	8	(17.02)	5	(10.42)	13	13.68
GG	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
C	0.915		0.948		0.932	
G	0.085		0.052		0.068	
Z	0.901					
p	0.367	*test for allele frequencies				

2978 C>A						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	45	(95.74)	46	(95.83)	91	95.79
CA	2	(4.26)	2	(4.17)	4	4.21
AA	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
C	0.979		0.979		0.979	
A	0.021		0.021		0.021	
Z	0.021					
p	0.983	*test for allele frequencies				

2984 G>A						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	47	(100.00)	47	(97.92)	94	98.95
GA	0	(0.00)	1	(2.08)	1	1.05
AA	0	(0.00)	0	(0.00)	0	0.00
	47		48		95	
G	1.000		0.990		0.995	
A	0.000		0.010		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

3.2.2.2 Blacks

Of 30 variants identified in Blacks, 12 had MAF >5%, 11 had MAF 1-5%, and 7 had MAF <1%. Of the 13 new variants, 2 had MAF >5%, 4 had MAF 1-5%, and 7 had MAF <1%. No statistically significant difference was identified when comparing the allele frequencies between the high HDL and low HDL groups for any of the 30 variants in this small sequence sample set (Tables 24 and 25). Of 17 rare variants, 3 were present only in the high HDL versus 5 only in the low HDL group. Of 12 exonic variants identified in Blacks 1 was present only in the high HDL group, and there was no low HDL group unique exonic variant. Of 4 nonsynonymous variants identified in Blacks, 1 was present only in the high HDL group versus none only in the low HDL group. Of the 47 individuals with low HDL levels, 4 (8.51%) had rare variants unique to the low HDL group. Of the 48 individuals with high HDL levels 4 (8.33%) had rare variants unique to the high HDL group.

Table 24. Distribution of *APOA4* Variants in High and Low HDL Groups in Blacks

	Blacks*	
Rare (MAF<5%)	**High HDL (n=47)**	**Low HDL (n=48)**
288ins12	0.021	0.011
568G>A	0.042	0.043
634G>A	0.010	0.032
755C>T	0.010	-
1274G>A	0.010	0.011
1371C>T	0.031	0.054
1453G>C	0.012	0.012
1743T>G	-	0.011
1948C>A	0.031	0.011
1993C>T	0.031	0.054
1994G>A	0.021	-
2327C>A**	0.011	-
2406C>G**	-	0.011
2685C>T**	-	0.011
2705C>T	-	0.011
2981C>T	0.031	0.043
3146G>A	-	0.011
Common (MAF≥5%)		
165delACAG	0.052	0.064
315T>A	0.073	0.074
357A>C	0.063	0.043
406C>A	0.135	0.117
974T>C	0.115	0.098
1198G>A	**0.372**	**0.478**
1326A>G	0.063	0.043
1334A>G	**0.448**	**0.359**
1735A>G	**0.448**	**0.352**
1803A>G	0.073	0.056
1853G>A	0.085	0.076
2104T>C	0.177	0.141
2645C>T	0.054	0.045

* The locations and nucleotide changes are based on the reverse strand sequence used in the SeattleSNPs database.
** Suspicious variants with low sequence quality.
Novel variants are hi-lighted.

Table 25. Allele Frequencies of *APOA4* Variants in High and Low HDL Groups in Blacks

165 del ACAG	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
WW	43	(89.58)	41	(87.23)	84	88.42
WD	5	(10.42)	6	(12.77)	11	11.58
DD		(0.00)		(0.00)	0	0.00
	48		47		95	
W	0.948		0.936		0.942	
D	0.052		0.064		0.058	
Z	0.346					
p	0.729	*test for allele frequencies				

288 ins CTGTTCCTGCTG	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
WW	46	(95.83)	46	(97.87)	92	96.84
WI	2	(4.17)	1	(2.13)	3	3.16
II		(0.00)		(0.00)	0	0.00
	48		47		95	
W	0.979		0.989		0.984	
I	0.021		0.011		0.016	
Z	0.566					
p	0.571	*test for allele frequencies				

315 T>A	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
TT	41	(85.42)	40	(85.11)	81	85.26
TA	7	(14.58)	7	(14.89)	14	14.74
AA		(0.00)		(0.00)	0	0.00
	48		47		95	
T	0.927		0.926		0.926	
A	0.073		0.074		0.074	
Z	0.041					
p	0.967	*test for allele frequencies				

357 A>C	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
AA	43	(89.58)	43	(91.49)	86	90.53
AC	4	(8.33)	4	(8.51)	8	8.42
CC	1	(2.08)		(0.00)	1	1.05
	48		47		95	
A	0.938		0.957		0.947	
C	0.063		0.043		0.053	
Z	0.617					
p	0.537	*test for allele frequencies				

406 C>A	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	35	(72.92)	36	(76.60)	71	74.74
CA	13	(27.08)	11	(23.40)	24	25.26
AA		(0.00)		(0.00)	0	0.00
	48		47		95	
C	0.865		0.883		0.874	
A	0.135		0.117		0.126	
Z	0.382					
p	0.702	*test for allele frequencies				

568 G>A	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	44	(91.67)	43	(91.49)	87	91.58
GA	4	(8.33)	4	(8.51)	8	8.42
AA		(0.00)		(0.00)	0	0.00
	48		47		95	
G	0.958		0.957		0.958	
A	0.042		0.043		0.042	
Z	0.030					
p	0.976	*test for allele frequencies				

634 G>A	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	47	(97.92)	44	(93.62)	91	95.79
GA	1	(2.08)	3	(6.38)	4	4.21
AA		(0.00)		(0.00)	0	0.00
	48		47		95	
G	0.990		0.968		0.979	
A	0.010		0.032		0.021	
Z	1.030					
p	0.303	*test for allele frequencies				

755 C>T	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	47	(97.92)	47	(100.00)	94	98.95
CT	1	(2.08)		(0.00)	1	1.05
TT		(0.00)		(0.00)	0	0.00
	48		47		95	
C	0.990		1.000		0.995	
T	0.010		0.000		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

Table 25 (Continued)

974 T>C						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
TT	37	(77.08)	37	(80.43)	74	78.72
TC	11	(22.92)	9	(19.57)	20	21.28
CC		(0.00)		(0.00)	0	0.00
	48		46		94	
T	0.885		0.902		0.894	
C	0.115		0.098		0.106	
Z	0.373					
p	0.709	*test for allele frequencies				

1198 G>A						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	20	(42.55)	11	(23.91)	31	33.33
GA	19	(40.43)	26	(56.52)	45	48.39
AA	8	(17.02)	9	(19.57)	17	18.28
	47		46		93	
G	0.628		0.522		0.575	
A	0.372		0.478		0.425	
Z	1.469					
p	0.142	*test for allele frequencies				

1274 G>A						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	47	(97.92)	45	(97.83)	92	97.87
GA	1	(2.08)	1	(2.17)	2	2.13
AA		(0.00)		(0.00)	0	0.00
	48		46		94	
G	0.990		0.989		0.989	
A	0.010		0.011		0.011	
Z	0.030					
p	0.976	*test for allele frequencies				

1326 A>G						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
AA	43	(89.58)	42	(91.30)	85	90.43
AG	4	(8.33)	4	(8.70)	8	8.51
GG	1	(2.08)		(0.00)	1	1.06
	48		46		94	
A	0.938		0.957		0.947	
G	0.063		0.043		0.053	
Z	0.584					
p	0.559	*test for allele frequencies				

1334 A>G						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
AA	16	(33.33)	18	(39.13)	34	36.17
AG	21	(43.75)	23	(50.00)	44	46.81
GG	11	(22.92)	5	(10.87)	16	17.02
	48		46		94	
A	0.552		0.641		0.596	
G	0.448		0.359		0.404	
Z	1.252					
p	0.210	*test for allele frequencies				

1371 C>T						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	45	(93.75)	41	(89.13)	86	91.49
CT	3	(6.25)	5	(10.87)	8	8.51
TT		(0.00)		(0.00)	0	0.00
	48		46		94	
C	0.969		0.946		0.957	
T	0.031		0.054		0.043	
Z	0.781					
p	0.435	*test for allele frequencies				

1453 G>C						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	42	(97.67)	42	(97.67)	84	97.67
GC	1	(2.33)	1	(2.33)	2	2.33
CC		(0.00)		(0.00)	0	0.00
	43		43		86	
G	0.988		0.988		0.988	
C	0.012		0.012		0.012	
Z	0.000					
p	1.000	*test for allele frequencies				

1735 A>G						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
AA	16	(33.33)	18	(40.91)	34	36.96
AG	21	(43.75)	21	(47.73)	42	45.65
GG	11	(22.92)	5	(11.36)	16	17.39
	48		44		92	
A	0.552		0.648		0.598	
G	0.448		0.352		0.402	
Z	1.330					
p	0.183	*test for allele frequencies				

Table 25 (Continued)

1743 T>G						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
TT	48	(100.00)	45	(97.83)	93	98.94
TG		(0.00)	1	(2.17)	1	1.06
GG		(0.00)		(0.00)	0	0.00
	48		46		94	
T	1.000		0.989		0.995	
G	0.000		0.011		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

1803 A>G						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
AA	41	(85.42)	40	(88.89)	81	87.10
AG	7	(14.58)	5	(11.11)	12	12.90
GG		(0.00)		(0.00)	0	0.00
	48		45		93	
A	0.927		0.944		0.935	
G	0.073		0.056		0.065	
Z	0.484					
p	0.628	*test for allele frequencies				

1853 G>A						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	39	(82.98)	39	(84.78)	78	83.87
GA	8	(17.02)	7	(15.22)	15	16.13
AA		(0.00)		(0.00)	0	0.00
	47		46		93	
G	0.915		0.924		0.919	
A	0.085		0.076		0.081	
Z	0.226					
p	0.821	*test for allele frequencies				

1948 C>A						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	45	(93.75)	45	(97.83)	90	95.74
CA	3	(6.25)	1	(2.17)	4	4.26
AA		(0.00)		(0.00)	0	0.00
	48		46		94	
C	0.969		0.989		0.979	
A	0.031		0.011		0.021	
Z	0.980					
p	0.327	*test for allele frequencies				

1993 C>T						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	45	(93.75)	41	(89.13)	86	91.49
CT	3	(6.25)	5	(10.87)	8	8.51
TT		(0.00)		(0.00)	0	0.00
	48		46		94	
C	0.969		0.946		0.957	
T	0.031		0.054		0.043	
Z	0.781					
p	0.435	*test for allele frequencies				

1994 G>A						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	46	(95.83)	46	(100.00)	92	97.87
GA	2	(4.17)		(0.00)	2	2.13
AA		(0.00)		(0.00)	0	0.00
	48		46		94	
G	0.979		1.000		0.989	
A	0.021		0.000		0.011	
Z	1.429					
p	0.153	*test for allele frequencies				

2104 T>C						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
TT	32	(66.67)	33	(71.74)	65	69.15
TC	15	(31.25)	13	(28.26)	28	29.79
CC	1	(2.08)		(0.00)	1	1.06
	48		46		94	
T	0.823		0.859		0.840	
C	0.177		0.141		0.160	
Z	0.672					
p	0.502	*test for allele frequencies				

2327* C>A						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	46	(97.87)	46	(100.00)	92	98.92
CA	1	(2.13)		(0.00)	1	1.08
AA		(0.00)		(0.00)	0	0.00
	47		46		93	
C	0.989		1.000		0.995	
A	0.011		0.000		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

Table 25 (Continued)

2406* C>G						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	48	(100.00)	46	(97.87)	94	98.95
CG		(0.00)	1	(2.13)	1	1.05
GG		(0.00)		(0.00)	0	0.00
	48		47		95	
C	1.000		0.989		0.995	
G	0.000		0.011		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

2645 C>T						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	41	(89.13)	40	(90.91)	81	90.00
CT	5	(10.87)	4	(9.09)	9	10.00
TT		(0.00)		(0.00)	0	0.00
	46		44		90	
C	0.946		0.955		0.950	
T	0.054		0.045		0.050	
Z	0.274					
p	0.784	*test for allele frequencies				

2685* C>T						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	48	(100.00)	45	(97.83)	93	98.94
CT		(0.00)	1	(2.17)	1	1.06
TT		(0.00)		(0.00)	0	0.00
	48		46		94	
C	1.000		0.989		0.995	
T	0.000		0.011		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

2705 C>T						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	48	(100.00)	45	(97.83)	93	98.94
CT		(0.00)	1	(2.17)	1	1.06
TT		(0.00)		(0.00)	0	0.00
	48		46		94	
C	1.000		0.989		0.995	
T	0.000		0.011		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

2981 C>T						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
CC	45	(93.75)	42	(91.30)	87	92.55
CT	3	(6.25)	4	(8.70)	7	7.45
TT		(0.00)		(0.00)	0	0.00
	48		46		94	
C	0.969		0.957		0.963	
T	0.031		0.043		0.037	
Z	0.441					
p	0.659	*test for allele frequencies				

3146 G>A						
	High HDL		Low HDL		TOTAL	
	n	(%)	*n*	(%)	*n*	(%)
GG	47	(100.00)	45	(97.83)	92	98.92
GA		(0.00)	1	(2.17)	1	1.08
AA		(0.00)		(0.00)	0	0.00
	47		46		93	
G	1.000		0.989		0.995	
A	0.000		0.011		0.005	
Z	1.005					
p	0.315	*test for allele frequencies				

3.3 LD AND TAGGER ANALYSES OF *APOA1* AND *APOA4* VARIANTS

SNPs that are in close proximity to one another along the chromosome can be inherited together, or in linkage disequilibrium (LD). SNPs in LD are compiled into haplotypes, which are identified by TagSNPs. TagSNPs can be identified within a group of SNPs using Tagger analysis. Identifying TagSNPs for a given haplotype reduces the number of SNPs needed for genotype screening by eliminating redundant analysis. LD and Tagger analysis was used to identify TagSNPs amongst the variants in *APOA1* and *APOA4*.

The *APOA1* and *APOA4* genes located in close vicinity on chromosome 11, within 12.5Kb distance. Therefore, pairwise LD and Tagger analysis was done for the *APOA1* and *APOA4* genes together to assess both intergenic and intragenic correlations. LD and Tagger analysis was limited to variants with a MAF >5%. A r^2 cutoff of 0.9 was used to assess high LD. A striking difference in LD was observed between the two populations.

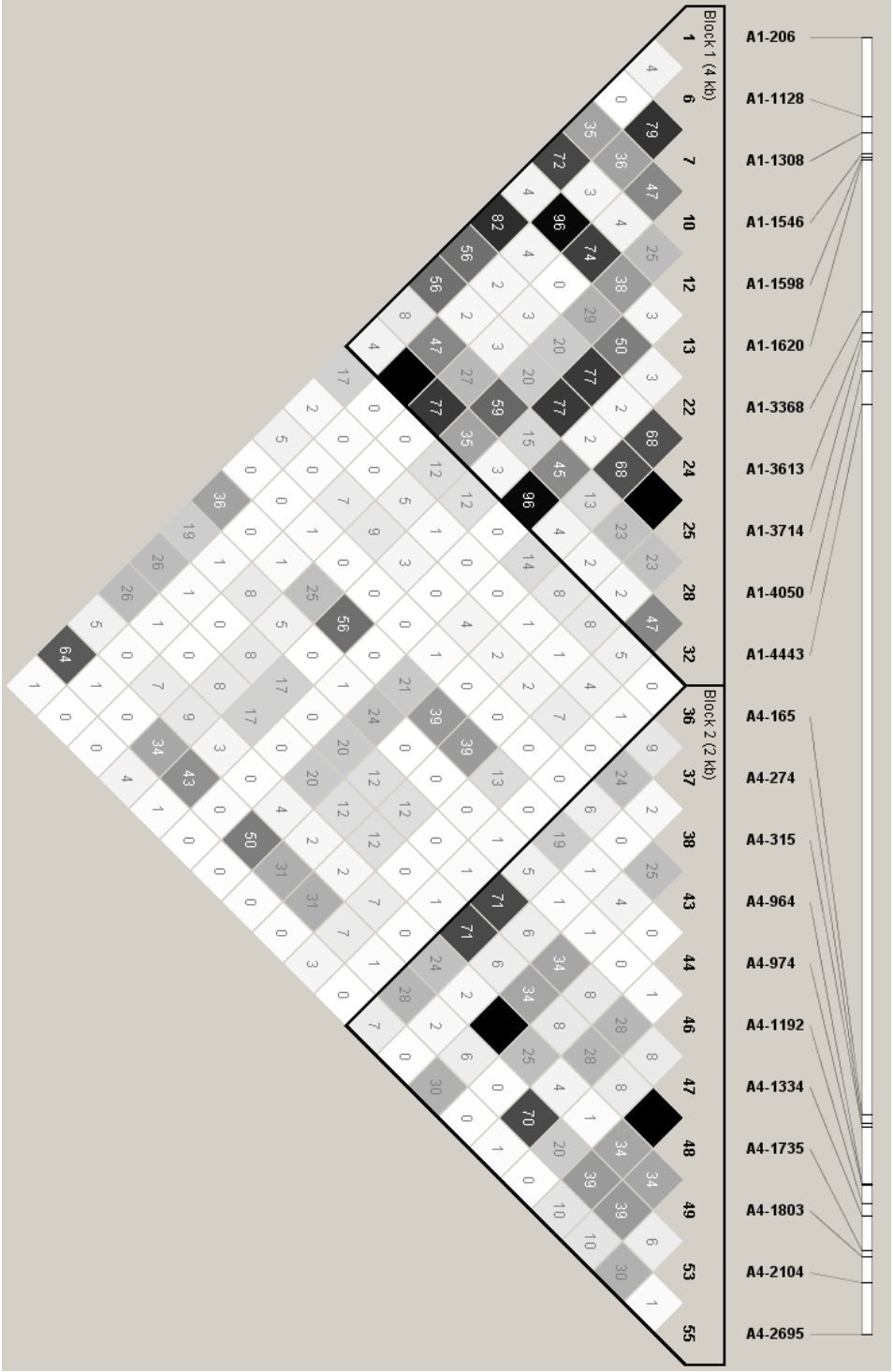
3.3.1 Non-Hispanic Whites

High correlation between *APOA1* and *APOA4* variants was not observed (the highest observed r^2 was 0.64 for any *APOA1/APOA4* variant pairs), although some strong correlations were present within each gene. Tagger analysis identified a total of 8 Bins for *APOA1* and 9 Bins for *APOA4*. A total of 22 common variants were captured in 17 Bins, 11 from *APOA1* and 11 from *APOA4*. Of the total 17 Bins from the two genes, pre-made TaqMan assays (Applied Biosystems) were

available for at least one variant in 5 Bins (Underlined in Table 26). Of those 5 available TaqMan assays two were for intronic *APOA1* variants (rs5070 and rs5072), two for exonic *APOA4* variants (rs5104, rs5092), and 1 for intronic *APOA4* variant (rs5100). The remaining Bins will be evaluated using either custom TaqMan assays or the Sequenome® iPLEX genotyping array.

Table 26. Tagger Results for NHWs

BIN	Gene	Location of Variants Captured	rs Numbers
1	APOA1	4443, 1620, 1128	rs670, rs12721028, rs11216153
2	APOA1	3613, 3714	rs5072, rs2070665
3	APOA1	4050	rs5070
4	APOA1	3368	rs7116797
5	APOA1	206	rs7123454
6	APOA1	1598	rs10750098
7	APOA1	1308	rs12721030
8	APOA1	1546	rs525028
9	APOA4	1735, 1334	rs5096, rs5100
10	APOA4	315, 1803	rs675, rs5095
11	APOA4	2104	rs5092
12	APOA4	2695	rs5090
13	APOA4	964	rs2234668
14	APOA4	165	rs9282602
15	APOA4	1192	rs5103
16	APOA4	274	rs5110
17	APOA4	974	rs5104



Color Scheme for r^2	
$r^2 = 0$	White
$0 < r^2 < 1$	Shades of Grey
$r^2 = 1$	Black

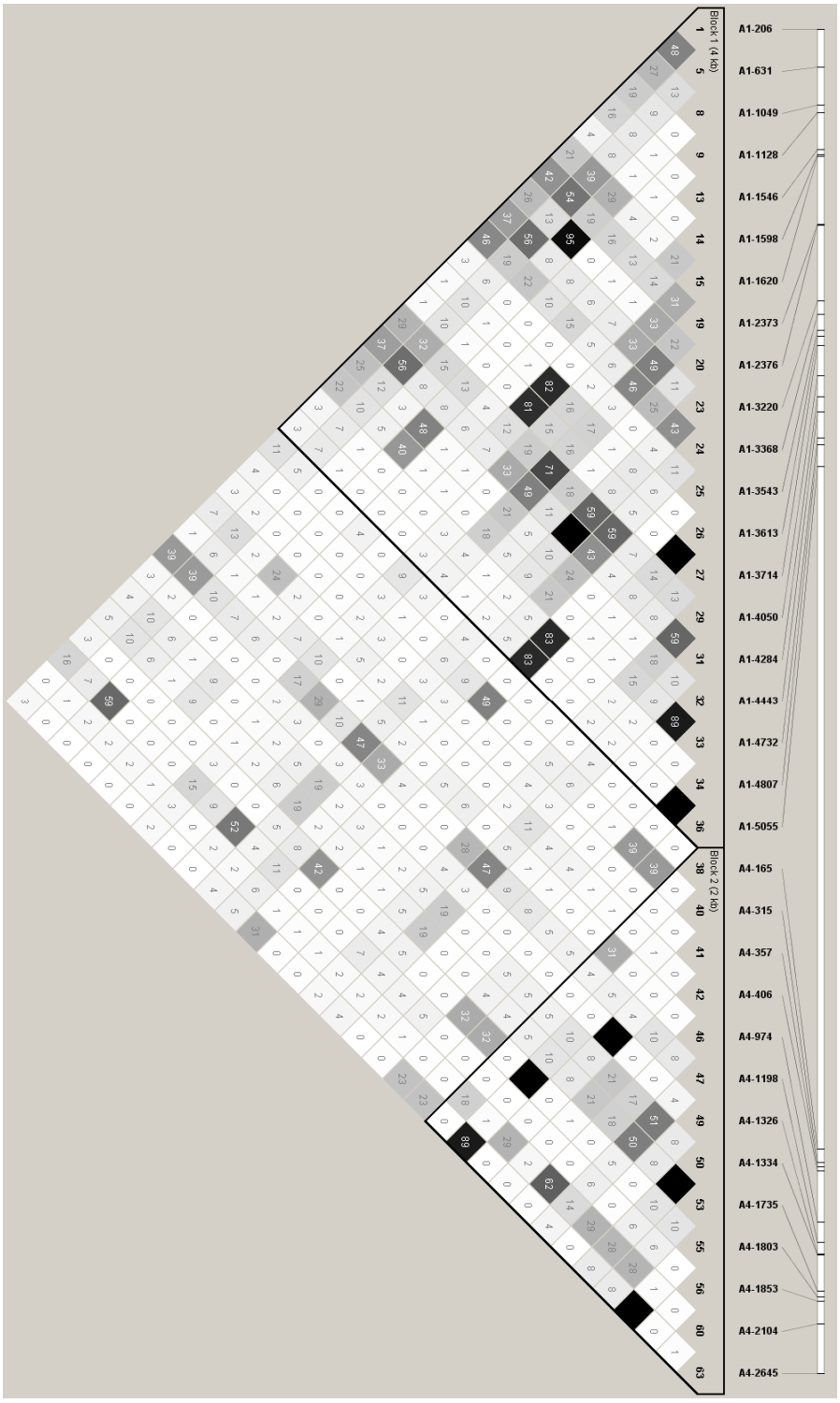
Figure 12. LD Analysis for NHWs.

3.3.2 Blacks

High correlation between *APOA1* and *APOA4* variants was not observed (the highest observed r^2 was 0.59 for any *APOA1/APOA4* variant pairs), although some strong correlations were present within each gene. Tagger analysis identified a total of 16 Bins for *APOA1* and 9 Bins for *APOA4*. A total of 33 common variants were captured in 25 Bins, 20 from *APOA1* and 13 from *APOA4*. Of the total 25 Bins from the two genes, pre-made TaqMan assays (Applied Biosystems) were available for at least one variant in 6 Bins (Underlined in Table 27). Additionally, a pre-made TaqMan assay was available and ordered for one rare variant in *APOA4* (not shown in Table 27: rs5106). Of those 7 available TaqMan assays two were for intronic *APOA1* variants (rs5070 and rs5072), four for exonic *APOA4* variants (rs5104, rs5092, rs5106, rs5109), and 1 for intronic *APOA4* variant (rs5100). The remaining Bins will be evaluated using either custom TaqMan assays or the Sequenome® iPLEX genotyping array.

Table 27. Tagger Results for Blacks

BIN	Gene	Location of Variants Captured	rs Numbers
1	APOA1	3714, 3613	rs2070665, rs5072
2	APOA1	2376, 1049	rs5081, rs1263162
3	APOA1	3220, 4284	rs5076, rs5069
4	APOA1	4807, 5055	rs12691374, -
5	APOA1	631	rs7948159
6	APOA1	206	rs7123454
7	APOA1	3543	rs5073
8	APOA1	1620	rs12721028
9	APOA1	1128	rs11216153
10	APOA1	3368	rs7116797
11	APOA1	4050	rs5070
12	APOA1	1598	rs10750098
13	APOA1	4732	rs12718467
14	APOA1	4443	rs670
15	APOA1	2373	rs12718436
16	APOA1	1546	rs525028
17	APOA4	315, 2645, 1803	rs675, rs5091, rs5095
18	APOA4	357, 1326	-, -
19	APOA4	1334, 1735	rs5100, rs5096
20	APOA4	1853	rs5094
21	APOA4	974	rs5104
22	APOA4	165	rs9282602
23	APOA4	1198	rs5101
24	APOA4	406	rs5109
25	APOA4	2104	rs5092



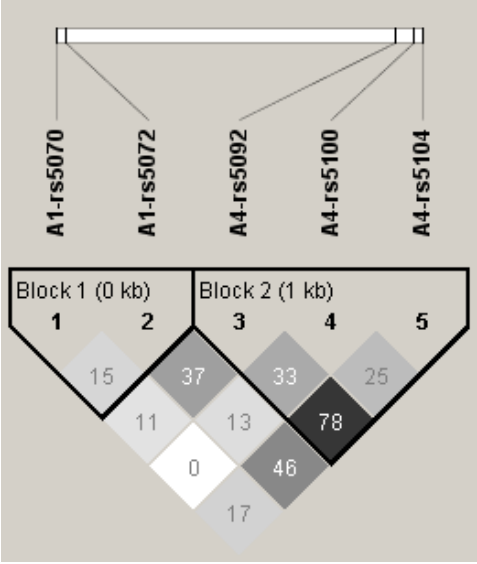
Color Scheme for r^2	
$r^2 = 0$	White
$0 < r^2 < 1$	Shades of Grey
$r^2 = 1$	Black

Figure 13. LD Analysis for Blacks

3.4 GENOTYPING OF ENTIRE NHW AND BLACK SAMPLES USING AVALIABLE TAQMAN SNP GENOTYPING ASSAYS

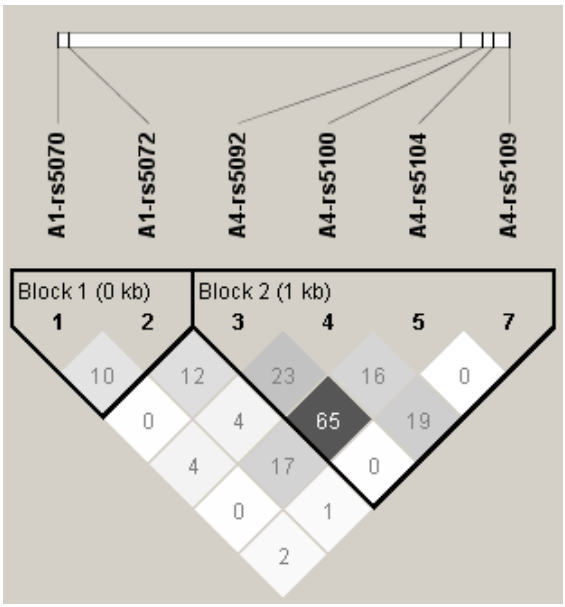
3.4.1 LD Analysis of the Variants Screened in Entire NWH and Black Samples

For the variants that were screened using the available TaqMan assays in the entire samples, the LD analysis was repeated (Figures 14 an 15) and the LD patterns were found to be similar to those observed in the subsets of the populations used for sequencing. A striking difference in LD was not observed between the two populations.



Color Scheme for r^2	
$r^2 = 0$	White
$0 < r^2 < 1$	Shades of Grey
$r^2 = 1$	Black

Figure 14. LD Analysis of the Variants Screened in the Entire NHW Samples.



Color Scheme for r^2	
$r^2 = 0$	White
$0 < r^2 < 1$	Shades of Grey
$r^2 = 1$	Black

Figure 15. LD Analysis of the Variants Screened in the Entire Black Samples.

3.4.2 Association Analysis of the Variants Screened in the Entire NWH and Black Samples for their Effects on Plasma HDL levels

A total of 7 variants screened in entire NHW and Black samples (only 5 were present in NHWs, all 7 were present in Blacks) were analyzed for their relation to plasma HDL levels separately in males and females within each ethnic group. The Tables 28 and 29 show the genotype counts, adjusted mean HDL levels (fore each genotype) and adjusted p-values (under the additive model) for each variant.

Although some modest or marginal p-values were observed, the associations were not consistent or strong enough to survive multiple testing correction in either of the populations.

Table 28. Genotype Distribution, Mean HDL Levels, and Adjusted p-values for five *APOA1* and *APOA4* Variants in NHWs

	NHW Males			**NWH Females**		
***APOA1*- rs5070**	**G/G[138]**	**G/A[127]**	**A/A[25]**	**G/G[167]**	**G/A[129]**	**A/A[28]**
HDL-C	44.02±0.87	43.84±0.91	43.41±2.05	56.90±1.07	57.24±1.21	54.68±2.60
	p^a=0.56			p^b=0.56		
***APOA1*- rs5072**	**G/G[252]**	**G/A[39]**	**A/A[3]**	**G/G[284]**	**G/A[41]**	**A/A[0]**
HDL-C	44.07±0.64	44.47±1.62	30.01±5.86	56.98±0.81	55.98±2.14	n/a
	p^a=0.16			p^b=0.53		
***APOA4*- rs5092**	**T/T[213]**	**T/C[76]**	**C/C[4]**	**T/T[226]**	**T/C[94]**	**C/C[6]**
HDL-C	44.44±0.70	43.31±1.17	36.43±5.10	56.14±0.91	57.79±1.41	67.23±5.61
	p^a=0.08			p^b=0.12		
***APOA4*- rs5100**	**A/A[123]**	**A/G[128]**	**G/G[43]**	**A/A[142]**	**A/G[143]**	**G/G[42]**
HDL-C	44.95±0.92	43.83±0.90	41.67±1.55	55.67±1.14	56.95±1.14	60.74±2.10
	p^a=0.04			**p^b=0.05**		
***APOA4*- rs5104**	**T/T[222]**	**T/C[68]**	**C/C[2]**	**T/T[248]**	**T/C[75]**	**C/C[3]**
HDL-C	44.43±0.68	43.04±1.24	34.26±7.22	56.73±0.87	56.93±1.58	58.01±7.96
	p^a=0.12			p^b=0.96		

p^a-values for log transformed HDL levels, adjusted for "BMI" under the additive model
p^b-values for log transformed HDL levels, adjusted for "age, smoking status, and BMI" under the additive model

Table 29. Genotype Distribution, Mean HDL Levels, and Adjusted p-values for five *APOA1* and *APOA4* Variants in Blacks

	Black Males			**Black Females**		
APOA1- rs5070	**G/G[92]**	**G/A[193]**	**A/A[89]**	**G/G[72]**	**G/A[136]**	**A/A[66]**
HDL-C	47.57±1.23	45.35±0.85	47.74±1.25	48.56±1.48	52.20±1.08	52.24±1.54
	p^a=0.88			**p^b=0.05**		
APOA1- rs5072	**G/G[309]**	**G/A[71]**	**A/A[1]**	**G/G[220]**	**G/A[58]**	**A/A[4]**
HDL-C	46.08±0.68	46.60±1.41	43.22±11.91	50.63±0.84	52.51±1.65	58.56±6.25
	p^a=0.77			p^b=0.12		
APOA4- rs5092	**T/T[297]**	**T/C[75]**	**C/C[6]**	**T/T[210]**	**T/C[64]**	**C/C[2]**
HDL-C	45.66±0.69	48.11±1.38	45.95±4.86	51.49±0.85	50.71±1.56	45.36±8.76
	p^a=0.15			p^b=0.51		
APOA4- rs5100	**A/A[154]**	**A/G[169]**	**G/G[51]**	**A/A[120]**	**A/G[127]**	**G/G[29]**
HDL-C	45.55±0.96	45.76±0.92	48.83±1.67	51.73±1.13	50.80±1.10	50.04±2.35
	p^a=0.15			p^b=0.47		
APOA4- rs5104	**T/T[317]**	**T/C[54]**	**C/C[4]**	**T/T[227]**	**T/C[41]**	**C/C[1]**
HDL-C	45.78±0.66	48.25±1.61	48.53±5.93	51.08±0.83	52.95±1.98	44.99±12.55
	p^a=0.13			p^b=0.48		
APOA4- rs5106	**G/G[353]**	**G/A[25]**	**A/A[0]**	**G/G[262]**	**G/A[16]**	**A/A[0]**
HDL-C	46.31±0.63	44.75±2.37	n/a	51.38±0.77	47.49±3.12	n/a
	p^a=0.48			p^b=0.24		
APOA4- rs5109	**C/C[310]**	**C/A[72]**	**A/A[1]**	**C/C[233]**	**C/A[45]**	**A/A[3]**
HDL-C	46.43±0.67	45.06±1.40	32.25±11.88	51.17±0.82	51.54±1.87	45.36±7.25
	p^a=0.25			p^b=0.79		

p^a-values for log transformed HDL levels, adjusted for "BMI" under the additive model
p^b-values for log transformed HDL levels, adjusted for "age, smoking status, and BMI" under the additive model

4.0 DISCUSSION

Apoa-I is the major apolipoprotein in HDL particles; many studies have provided evidence of the athroprotective role of ApoA-I.[32] *APOA1* mutations have been correlated with Mendelian disease, however further study is needed to associate common and rare variants in *APOA1* with complex genetic disease.[16]

While the exact function of apoA-IV is not known, it has a number of proposed functions, including involvement in the assembly and secretion of chylomicrons and the reverse cholesterol transport system.[81] Some studies have associated variation of the *APOA4* gene with changes in lipid levels, while others have not observed this same pattern.[79,83,90] The *APOA4* gene clearly requires further study due to the lack of consensus about its biological function and the correlation of variation within this gene with lipid levels.

This study aimed to evaluate the role of *APOA1* and *APOA4* genetic variation by sequencing a subset of samples from healthy individuals with HDL levels in the 5th and 95th percentiles. The purpose of this was to detect both the rare and common variants in both genes in NHW and Black populations. Through sequence analysis and detection of these variants both the "common variant-common disease" and "rare variant-common disease" hypotheses were tested.

The common variant hypothesis has been extensively evaluated through the candidate gene approach and GWAS. The earlier candidate gene studies reported association of the

APOA1/APOC3/APOA4/APOA5 gene cluster with HDL and triglyceride levels, however the results were not consistent.[45,79] In recent GWAS this gene cluster showed some association with triglyceride levels; most of the genetic variants with significance were in *APOA5*, intergenic regions, or other genes residing near this cluster.[18-25] The two GWAS that reported association with HDL levels and this gene cluster again included SNPs from neighboring genes (*ZNF259* and *BUD13*), but not SNPs from *APOA1* or *APOA4*.[21,23]

The rare variant hypothesis was less frequently addressed in the literature, although it is more likely to be addressed in the near future with the advent of Next Gen sequencing, and the decreasing cost of sequencing technology. Cohen *et al.*[27] used sequencing to analyze the coding regions of three genes (*APOA1*, *ABCA1*, and *LCAT*), and reported that individuals with low HDL had significantly more nonsynonymous variants than individuals with high HDL levels. However, most of the variants they identified were from *ABCA1* or *LCAT*, and only a few were from *APOA1*. This study used the same approach for sample population selection (sequencing individuals with HDL levels in the 5th and 95th percentile), however, in this study the genes were completely sequenced to document all variants and their association with HDL levels, rather than just the coding regions as in Cohen *et al*.[27]

SeattleSNPs also completely sequenced both *APOA1* and *APOA4*; however, there are some inherent reasons for differences between sequence variants reported in the SeattleSNPs database and sequence variants identified in this study. A total of 24 African-American individuals, 24 European individuals, and 24 non-Hispanic European-American individuals, unselected for HDL-cholesterol levels, were sequenced in the SeattleSNPs study.[78] A total of 95 African individuals and 95 NHWs, who were selected based on their extreme HDL-cholesterol levels, were sequenced in this study. The SeattleSNPs study sequenced African-American

individuals, whereas African samples were sequenced in this study. In African-American samples there is a greater likelihood of admixture from other ethnic groups as compared to African samples. Different software tools were used to analyze the sequence data in the SeattleSNPs study versus this study. Also, different primers were used for sequencing (*APOA1*).

For sequence analysis the data collected in this study was compared with the data from the SeattleSNPs database, published in Fullerton *et al.*[78] A total of 31 sequence variants in *APOA1* were reported in the SeattleSNPs database (Tables 2-4, section 1.4.3). A total of 54 sequence variants in *APOA1* were identified in this study (Table 16, section 3.1.1). A total of 24 sequence variants in *APOA4* were reported in the SeattleSNPs database (Tables 5-6, section 1.5.2). A total of 43 sequence variants were identified in this study (Table 17, section 3.1.2).

A total of 25 sequence variants in *APOA1* were reported in the SeattleSNPs database for the two European populations (Tables 3&4, section 1.4.3). A total of 34 sequence variants in *APOA1* were identified in this study in NHWs (Table 16, section 3.1.1). A total of 25 sequence variants in *APOA1* were reported in the SeattleSNPs database for the African American population (Table 2, section 1.4.3). A total of 37 sequence variants in *APOA1* were identified in this study in Blacks (Table 16, section 3.1.1). A total of 18 sequence variants in *APOA4* were reported in the SeattleSNPs database for the two European populations (Tables 6&7, section 1.5.2). A total of 23 sequence variants were identified in this study in NHWs (Table 17, section 3.1.2). A total of 18 sequence variants in *APOA4* were reported in the SeattleSNPs database for the African American population (Table 5, section 1.5.2). A total of 30 sequence variants were identified in this study in Blacks (Table 17, section 3.1.2).

One coding variant in *APOA1* was reported by Fullerton *et al.* and the SeattleSNPs database data. This variant was present in both the African-American and European-American

populations. The sequence variant was nonsynonymous (Ala>Thr).[78] Five coding variants in *APOA1* were identified in this study (Table 16, section 3.1.1). Two of the 5 variants were present in NHWs and 3 of the 5 in Blacks. Four of the 5 variants were nonsynonymous. Twelve coding variants in *APOA4* were reported by Fullerton *et al.*[78] and the SeattleSNPs database data. Eight of the 12 variants were present in the African-American population and 7 of the 12 variants were reported in the European populations. Six of the 12 variants were nonsynonomous. Seventeen coding variants in *APOA4* were identified in this study (Table 17, section 3.1.2). Eleven of the 17 variants were present in NHWs and 11 of the 17 in Blacks. Seven of the 17 variants were nonsynonymous.

Table 30 is a list of sequence variants in *APOA1* and *APOA4* reported in the SeattleSNPs database that were not identified in this study. Identification of these variants would be expected given the fact that the NHW and Black sample sizes are 2-4 times larger in this study than in the SeattleSNPs study. Therefore, it is possible that those variants listed below that were seen in only a single individual in one population represent sequence artifacts as they have not been confirmed using another technology. Variants that were reported in more than one individual or population may have been absent in this study due to differences in selection criteria (sequencing of only individuals with HDL levels in the 5th and 95th percentile in this study).

Table 30. Unique Sequence Variants in the SeattleSNPs Database

Gene	SeattleSNP Location	rs Number	MAF JD-Pop	MAF ND-Pop	MAF RD-Pop
APOA1	1541	rs127211029	0.03	-	-
APOA1	1717	rs12718461	-	-	0.02
APOA1	3766	rs12718465	0.10	-	0.09
APOA1	4245	rs12712032	0.02	-	-
APOA4	933	rs12721043	-	0.09	0.02
APOA4	1183	rs12721042	0.02	-	-
APOA4	2511	rs12721041	0.05	-	-

Novel variants identified in this study (not previously reported in publicly available databases) are listed in Tables 16 and 17 in sections 3.1.1 and 3.1.2, respectively. Suspicious variants with low sequence quality (denoted in each table) are to be confirmed in future analysis. This study had a larger sequencing sample size than the SeattleSNPs study which may have contributed to the number of sequencing variants. Additionally, this study sequenced individuals with HDL levels in the 5th and 95th percentile whereas the SeattleSNPs database did not select for any risk-factor trait. It is possible that some of the novel variants seen in this study are unique to this group.

Fullerton *et al.*[78] reported a higher variability among African Americans as compared to Europeans for both *APOA1* and *APOA4*. In this study, a higher number of sequence variants were also identified in Blacks versus NHWs for both genes.

Fullerton *et al*.[78] observed that *APOA4* had many more coding region variants than the other genes in the *APOA1/APOC3/APOA4/APOA5* gene cluster. This same conclusion can be made when comparing the number of coding variants observed in *APOA4* versus *APOA1* in this study: 5 coding variants were identified in *APOA1* versus 17 in *APOA4* (Tables 16 and 17 in sections 3.1.1&3.1.2, respectively).

According to preliminary analysis of sequence data for *APOA1* and *APOA4*, no striking difference was noticed between the distribution of rare variants between high and low HDL groups in either population. For sequencing variants in *APOA1*: for NHWs, 5 out of 48 (10.42%) individuals with low HDL levels had rare variants unique to the low group versus 7 out of 47 (14.89%) individuals with high HDL levels with rare variants unique to the high group; for Blacks, 2 out of 47 (4.26%) individuals with low HDL levels had rare variants unique to the low group versus 6 out of 48 (12.5%) individuals with high HDL levels with rare variants unique to

the high group. For sequencing variants in *APOA4*: for NHWs, 3 out of 48 (6.25%) individuals with low HDL levels had rare variants unique to the low group versus 5 out of 47 (10.64%) individuals with high HDL levels with rare variants unique to the high group; for Blacks, 4 out of 47 (8.51%) individuals with low HDL levels had rare variants unique to the low group versus 4 out of 48 (8.33%) individuals with high HDL levels with rare variants unique to the high group. Overall, when individuals with rare variants are compared between high and low HDL groups the numbers were similar or slightly higher in the high HDL group.

Differences in MAF between low and high HDL groups have been observed for some common variants in the sequencing data and have not yet been confirmed by genotyping in the entire population. These variants include: 206 (rs7123454) and 1598 (rs10750098) in NHWs in *APOA1* (in bold in Table 18); 964 (rs2234668) in NHWs in *APOA4* (in bold in Table 22); 1198 (rs5101) and 1735 (rs5096) in Blacks in *APOA4* (in bold in Table 24). None of these variants have been previously associated with variation in HDL-cholesterol levels in the literature. *APOA4* variant 1334 (rs5100) showed a difference in MAF between low and high HDL groups in Blacks in the sequencing data (in bold in Table 24). This variant had not been previously associated with variation in HDL-cholesterol levels in the literature. This variant was genotyped in the entire NHW and Black population in this study.

Thus far screening data has been compiled for the entire NHW and Black population for a total of seven variants: 2 for *APOA1* (rs5070 and rs5072), and 5 in *APOA4* (rs5092, rs5100, rs5104, rs5106, and rs5109). All 7 variants were present in the Black population; five were present in NHWs (rs5070, rs5072, rs5092, rs5100, and rs5104). Modest or marginal p-values were observed, however, none would maintain significance after multiple testing correction in either population. Some of the variants were investigated in the literature with inconsistent

results including: rs5070, rs5092, and rs5104.[45,79] Inconsistencies in the literature and a lack of statistically significant association with HDL levels in this study may be due to population size; variants associated with a small effect on HDL levels may only be statistically significant with a larger population size. Furthermore, additional variants identified in sequencing (both novel and those previously reported in the publicly available databases) remain to be screened in the entire NHW and Black population.

5.0 CONCLUSION

Heart disease is a major public health concern, and decreased HDL-cholesterol levels are a major risk factor for heart disease. In previous candidate gene studies and GWAS, *APOA1* and *APOA4* have been associated with variation in HDL-cholesterol levels with inconsistent results. This study supported this paradigm. The common variants that were genotyped in the entire population had only modest or marginal p-values that would not maintain significance after multiple testing correction. However, additional common variants remain to be screened in the entire population; in some of these variants differences were observed between high and low HDL groups in preliminary sequence data.

Further data collection and analysis is necessary to better understand the significance of these variants. Additional studies of *APOA1*and *APOA4* with larger population sizes are needed to analyze variants that may only have a small effect on HDL-cholesterol levels. Further studies of rare variation in this and other genes are also required to better understand genetics of HDL-cholesterol in relation to the rare allele hypothesis.

BIBLIOGRAPHY

1. Dastani, Z., Engert, J.C., Genest, J. & Marcil, M. Genetics of high-density lipoproteins. *Curr Opin Cardiol* **21**, 329-35 (2006).
2. Hegele, R.A. Plasma lipoproteins: genetic influences and clinical implications. *Nat Rev Genet* **10**, 109-21 (2009).
3. Kung, H.-C., Hoyert, D.L., Xu, J. & Murphy, S.L. Deaths: Final Data for 2005. *National Vital Statistics Reports* **56**(2008).
4. Rosamond, W. et al. Heart disease and stroke statistics--2008 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation* **117**, e25-146 (2008).
5. Greenland, P. et al. Major risk factors as antecedents of fatal and nonfatal coronary heart disease events. *JAMA* **290**, 891-897 (2003).
6. Yusuf, S. et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet* **364**, 937-52 (2004).
7. Barter, P.J. et al. Antiinflammatory properties of HDL. *Circ Res* **95**, 764-72 (2004).
8. Lewington, S. et al. Blood cholesterol and vascular mortality by age, sex, and blood pressure: a meta-analysis of individual data from 61 prospective studies with 55,000 vascular deaths. *Lancet* **370**, 1829-39 (2007).
9. Castelli, W.P. et al. Incidence of coronary heart disease and lipoprotein cholesterol levels. The Framingham Study. *JAMA* **256**, 2835-8 (1986).
10. Gordon, D.J. et al. High-density lipoprotein cholesterol and cardiovascular disease. Four prospective American studies. *Circulation* **79**, 8-15 (1989).
11. Assmann, G., Cullen, P. & Schulte, H. The Munster Heart Study (PROCAM). Results of follow-up at 8 years. *Eur Heart J* **19 Suppl A**, A2-11 (1998).
12. Goode, E.L., Cherny, S.S., Christian, J.C., Jarvik, G.P. & de Andrade, M. Heritability of longitudinal measures of body mass index and lipid and lipoprotein levels in aging twins. *Twin Res Hum Genet* **10**, 703-11 (2007).
13. Kronenberg, F. et al. Segregation analysis of HDL cholesterol in the NHLBI Family Heart Study and in Utah pedigrees. *Eur J Hum Genet* **10**, 367-74 (2002).
14. Perusse, L. et al. Familial resemblance of plasma lipids, lipoproteins and postheparin lipoprotein and hepatic lipases in the HERITAGE Family Study. *Arterioscler Thromb Vasc Biol* **17**, 3263-9 (1997).
15. Wang, X. & Paigen, B. Genetics of variation in HDL cholesterol in humans and mice. *Circ Res* **96**, 27-42 (2005).

16. Pollex, R.L. & Hegele, R.A. Genetic determinants of plasma lipoproteins. *Nat Clin Pract Cardiovasc Med* **4**, 600-9 (2007).
17. Talmud, P.J. & Humphries, S.E. Genetic polymorphisms, lipoproteins and coronary artery disease risk. *Curr Opin Lipidol* **12**, 405-9 (2001).
18. Aulchenko, Y. et al. Genome-wide association study in 16 European population cohorts: Major loci influencing lipid levels and coronary heart risk. *Nat. Genet.* (2008).
19. Chasman, D. et al. Genetic loci associated with plasma concentration of low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, triglycerides, apolipoprotein AI, and apolipoprotein B among 6382 white women in genome-wide analysis with replication. *Circulation: Cardiovascular Genetics* **1**, 21-30 (2008).
20. Heid, I. et al. A genome-wide association analysis of HDL-cholesterol in the population-based KORA Study sheds new light on intergenic regions. *Circulation: Cardiovascular Genetics* **1**, 10-20 (2008).
21. Kathiresan, S. et al. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat Genet* **40**, 189-97 (2008).
22. Kooner, J.S. et al. Genome-wide scan identifies variation in MLXIPL associated with plasma triglycerides. *Nat Genet* **40**, 149-51 (2008).
23. Sabatti, C. et al. Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. *Nat Genet* **41**, 35-46 (2009).
24. Wallace, C. et al. Genome-wide association study identifies genes for biomarkers of cardiovascular disease: serum urate and dyslipidemia. *Am J Hum Genet* **82**, 139-49 (2008).
25. Willer, C.J. et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet* **40**, 161-9 (2008).
26. Yuan, G., Wang, J. & Hegele, R.A. Heterozygous familial hypercholesterolemia: an underrecognized cause of early cardiovascular disease. *CMAJ* **174**, 1124-9 (2006).
27. Cohen, J.C. et al. Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science* **305**, 869-72 (2004).
28. Benn, M. et al. Common and rare alleles in apolipoprotein B contribute to plasma levels of low-density lipoprotein cholesterol in the general population. *J Clin Endocrinol Metab* **93**, 1038-45 (2008).
29. Cohen, J.C. et al. Multiple rare variants in NPC1L1 associated with reduced sterol absorption and plasma low-density lipoprotein levels. *Proc Natl Acad Sci U S A* **103**, 1810-5 (2006).
30. Romeo, S. et al. Rare loss-of-function mutations in ANGPTL family members contribute to plasma triglyceride levels in humans. *J Clin Invest* **119**, 70-9 (2009).
31. Brewer, H.B., Jr. et al. The amino acid sequence of human APOA-I, an apolipoprotein isolated from high density lipoproteins. *Biochem Biophys Res Commun* **80**, 623-30 (1978).
32. Scanu, A.M. & Edelstein, C. HDL: bridging past and present with a look at the future. *FASEB J* **22**, 4044-54 (2008).
33. Scanu, A.M., Edelstein, C. & Gordon, J.I. Apolipoproteins of human plasma high density lipoproteins. Biology, biochemistry and clinical significance. *Clin Physiol Biochem* **2**, 111-22 (1984).

34. Segrest, J.P. et al. The amphipathic helix in the exchangeable apolipoproteins: a review of secondary structure and function. *J Lipid Res* **33**, 141-66 (1992).
35. Ajees, A.A., Anantharamaiah, G.M., Mishra, V.K., Hussain, M.M. & Murthy, H.M. Crystal structure of human apolipoprotein A-I: insights into its protective effect against cardiovascular diseases. *Proc Natl Acad Sci U S A* **103**, 2126-31 (2006).
36. Binder, C.J. et al. Innate and acquired immunity in atherogenesis. *Nat Med* **8**, 1218-26 (2002).
37. Gotsman, I., Gupta, R. & Lichtman, A.H. The influence of the regulatory T lymphocytes on atherosclerosis. *Arterioscler Thromb Vasc Biol* **27**, 2493-5 (2007).
38. Luoma, P.V. Gene activation, apolipoprotein A-I/high density lipoprotein, atherosclerosis prevention and longevity. *Pharmacol Toxicol* **81**, 57-64 (1997).
39. Rader, D.J. Molecular regulation of HDL metabolism and function: implications for novel therapies. *J Clin Invest* **116**, 3090-100 (2006).
40. Yancey, P.G. et al. Importance of different pathways of cellular cholesterol efflux. *Arterioscler Thromb Vasc Biol* **23**, 712-9 (2003).
41. Zannis, V.I., Chroni, A. & Krieger, M. Role of apoA-I, ABCA1, LCAT, and SR-BI in the biogenesis of HDL. *J Mol Med* **84**, 276-94 (2006).
42. Benoit, P. et al. Somatic gene transfer of human ApoA-I inhibits atherosclerosis progression in mouse models. *Circulation* **99**, 105-10 (1999).
43. Moore, R.E. et al. Apolipoprotein A-I deficiency results in markedly increased atherosclerosis in mice lacking the LDL receptor. *Arterioscler Thromb Vasc Biol* **23**, 1914-20 (2003).
44. Navab, M., Anantharamaiah, G.M., Reddy, S.T. & Fogelman, A.M. Apolipoprotein A-I mimetic peptides and their role in atherosclerosis prevention. *Nat Clin Pract Cardiovasc Med* **3**, 540-7 (2006).
45. Brown, C.M. et al. The contribution of individual and pairwise combinations of SNPs in the APOA1 and APOC3 genes to interindividual HDL-C variability. *J Mol Med* **84**, 561-72 (2006).
46. Kamboh, M.I., Aston, C.E., Nestlerode, C.M., McAllister, A.E. & Hamman, R.F. Haplotype analysis of two APOA1/MspI polymorphisms in relation to plasma levels of apo A-I and HDL-cholesterol. *Atherosclerosis* **127**, 255-62 (1996).
47. Shioji, K. et al. An association analysis between ApoA1 polymorphisms and the high-density lipoprotein (HDL) cholesterol level and myocardial infarction (MI) in Japanese. *J Hum Genet* **49**, 433-9 (2004).
48. Francis, M.C. & Frohlich, J.J. Coronary artery disease in patients at low risk--apolipoprotein AI as an independent risk factor. *Atherosclerosis* **155**, 165-70 (2001).
49. Lusis, A.J., Mar, R. & Pajukanta, P. Genetics of atherosclerosis. *Annu Rev Genomics Hum Genet* **5**, 189-218 (2004).
50. Weisgraber, K.H., Bersot, T.P., Mahley, R.W., Franceschini, G. & Sirtori, C.R. A-Imilano apoprotein. Isolation and characterization of a cysteine-containing variant of the A-I apoprotein from human high density lipoproteins. *J Clin Invest* **66**, 901-7 (1980).
51. Franceschini, G., Sirtori, C.R., Capurso, A., 2nd, Weisgraber, K.H. & Mahley, R.W. A-IMilano apoprotein. Decreased high density lipoprotein cholesterol levels with significant lipoprotein modifications and without clinical atherosclerosis in an Italian family. *J Clin Invest* **66**, 892-900 (1980).

52. Utermann, G., Feussner, G., Franceschini, G., Haas, J. & Steinmetz, A. Genetic variants of group A apolipoproteins. Rapid methods for screening and characterization without ultracentrifugation. *J Biol Chem* **257**, 501-7 (1982).
53. Breslow, J.L. Apolipoprotein genetic variation and human disease. *Physiol Rev* **68**, 85-132 (1988).
54. Strobl, W. et al. Apolipoprotein A-I (Glu 198----Lys): a mutant of the major apolipoprotein of high-density lipoproteins occurring in a family with dyslipoproteinemia. *Pediatr Res* **24**, 222-8 (1988).
55. von Eckardstein, A. et al. Apolipoprotein A-I variants. Naturally occurring substitutions of proline residues affect plasma concentration of apolipoprotein A-I. *J Clin Invest* **84**, 1722-30 (1989).
56. Van Allen, M.W., Frohlich, J.A. & Davis, J.R. Inherited predisposition to generalized amyloidosis. Clinical and pathological study of a family with neuropathy, nephropathy, and peptic ulcer. *Neurology* **19**, 10-25 (1969).
57. Nichols, W.C., Gregg, R.E., Brewer, H.B., Jr. & Benson, M.D. A mutation in apolipoprotein A-I in the Iowa type of familial amyloidotic polyneuropathy. *Genomics* **8**, 318-23 (1990).
58. Norum, R.A. et al. Familial deficiency of apolipoproteins A-I and C-III and precocious coronary-artery disease. *N Engl J Med* **306**, 1513-9 (1982).
59. Schaefer, E.J., Heaton, W.H., Wetzel, M.G. & Brewer, H.B., Jr. Plasma apolipoprotein A-1 absence associated with a marked reduction of high density lipoproteins and premature coronary artery disease. *Arteriosclerosis* **2**, 16-26 (1982).
60. Ladias, J.A., Kwiterovich, P.O., Jr., Smith, H.H., Karathanasis, S.K. & Antonarakis, S.E. Apolipoprotein A1 Baltimore (Arg10----Leu), a new ApoA1 variant. *Hum Genet* **84**, 439-45 (1990).
61. Funke, H. et al. A frameshift mutation in the human apolipoprotein A-I gene causes high density lipoprotein deficiency, partial lecithin: cholesterol-acyltransferase deficiency, and corneal opacities. *J Clin Invest* **87**, 371-6 (1991).
62. Matsunaga, T. et al. Apolipoprotein A-I deficiency due to a codon 84 nonsense mutation of the apolipoprotein A-I gene. *Proc Natl Acad Sci* **88**, 2793-2797 (1991).
63. Soutar, A.K. et al. Apolipoprotein AI mutation Arg-60 causes autosomal dominant amyloidosis. *Proc Natl Acad Sci U S A* **89**, 7389-93 (1992).
64. Ng, D.S., Leiter, L.A., Vezina, C., Connelly, P.W. & Hegele, R.A. Apolipoprotein A-I Q[-2]X causing isolated apolipoprotein A-I deficiency in a family with analphalipoproteinemia. *J Clin Invest* **93**, 223-9 (1994).
65. Nakata, K. et al. Autosomal dominant hypoalphalipoproteinemia due to a completely defective apolipoprotein A-I gene. *Biochem Biophys Res Commun* **196**, 950-5 (1993).
66. Romling, R. et al. A nonsense mutation in the apolipoprotein A-I gene is associated with high-density lipoprotein deficiency and periorbital xanthelasmas. *Atherosclerosis Thromb* **14**(1994).
67. Booth, D.R. et al. Hereditary hepatic and systemic amyloidosis caused by a new deletion/insertion mutation in the apolipoprotein AI gene. *J Clin Invest* **97**, 2714-21 (1996).
68. Huang, W. et al. A novel homozygous missense mutation in the apo A-I gene with apo A-I deficiency. *Arterioscler Thromb Vasc Biol* **18**, 389-96 (1998).

69. Yamakawa-Kobayashi, K. et al. Frequent occurrence of hypoalphalipoproteinemia due to mutant apolipoprotein A-I gene in the population: a population-based survey. *Hum Molec Genet* **8**, 331-336 (1999).
70. Hamidi Asl, L. et al. Hereditary amyloid cardiomyopathy caused by a variant apolipoprotein A1. *Am J Pathol* **154**, 221-7 (1999).
71. Hamidi Asl, K., Liepnieks, J.J., Nakamura, M., Parker, F. & Benson, M.D. A novel apolipoprotein A-1 variant, Arg173Pro, associated with cardiac and cutaneous amyloidosis. *Biochem Biophys Res Commun* **257**, 584-8 (1999).
72. Obici, L. et al. The new apolipoprotein A-I variant leu(174) --> Ser causes hereditary cardiac amyloidosis, and the amyloid fibrils are constituted by the 93-residue N-terminal polypeptide. *Am J Pathol* **155**, 695-702 (1999).
73. Lachmann, H.J. et al. Misdiagnosis of hereditary amyloidosis as AL (primary) amyloidosis. *N Engl J Med* **346**, 1786-91 (2002).
74. Pisciotta, L. et al. Recurrent mutations of the apolipoprotein A-I gene in three kindreds with severe HDL deficiency. *Atherosclerosis* **167**, 335-45 (2003).
75. Hovingh, G.K. et al. A novel apoA-I mutation (L178P) leads to endothelial dysfunction, increased arterial wall thickness, and premature coronary artery disease. *J Am Coll Cardiol* **44**, 1429-35 (2004).
76. Dastani, Z. et al. A novel nonsense apolipoprotein A-I mutation (apoA-I(E136X)) causes low HDL cholesterol in French Canadians. *Atherosclerosis* **185**, 127-36 (2006).
77. Esperon, P., Raggio, V., Stoll, M., Vital, M. & Alallon, W. A new APOA1 mutation with severe HDL-cholesterol deficiency and premature coronary artery disease. *Clin Chim Acta* **388**, 222-4 (2008).
78. Fullerton, S.M. et al. The effects of scale: variation in the APOA1/C3/A4/A5 gene cluster. *Hum Genet* **115**, 36-56 (2004).
79. Knoblauch, H. et al. Haplotypes and SNPs in 13 lipid-relevant genes explain most of the genetic variance in high-density lipoprotein and low-density lipoprotein cholesterol. *Hum Mol Genet* **13**, 993-1004 (2004).
80. Pearson, K. et al. Structure of human apolipoprotein A-IV: a distinct domain architecture among exchangeable apolipoproteins with potential functional implications. *Biochemistry* **43**, 10719-29 (2004).
81. Tubb, M.R., Silva, R.G.D., Fang, J., Tso, P. & Davidson, W.S. A Three-dimensional Homology Model of Lipid-free Apolipoprotein A-IV Using Cross-linking and Mass Spectrometry. *The Journal of Biological Chemestry* **283**, 17314-17323 (2008).
82. Lohse, P., Kindt, M.R., Rader, D.J. & Brewer, H.B., Jr. Genetic polymorphism of human plasma apolipoprotein A-IV is due to nucleotide substitutions in the apolipoprotein A-IV gene. *J Biol Chem* **265**, 10061-4 (1990).
83. Cendoroglo, M.S. et al. Association of apo A-IV 360 (Gln --> His) polymorphism with plasma lipids and lipoproteins: the Framingham Offspring Study. *Atherosclerosis* **179**, 169-75 (2005).
84. Boerwinkle, E., Visvikis, S. & Chan, L. Two polymorphisms for amino acid substitutions in the APOA4 gene. *Nucleic Acids Res* **18**, 4966 (1990).
85. Wong, W.M. et al. Apolipoprotein AIV gene variant S347 is associated with increased risk of coronary heart disease and lower plasma apolipoprotein AIV levels. *Circ Res* **92**, 969-75 (2003).

86. Wong, W.M. et al. The APOA4 T347S variant is associated with reduced plasma TAOS in subjects with diabetes mellitus and cardiovascular disease. *J Lipid Res* **45**, 1565-71 (2004).
87. Miltiadous, G. et al. Genetic polymorphisms of the apolipoprotein A-IV in a Greek population and their relation to plasma lipid and lipoprotein levels. *Clin Genet* **62**, 208-13 (2002).
88. Lohse, P., Kindt, M.R., Rader, D.J. & Brewer, H.B., Jr. Human plasma apolipoproteins A-IV-0 and A-IV-3. Molecular basis for two rare variants of apolipoprotein A-IV-1. *J Biol Chem* **265**, 12734-9 (1990).
89. Lohse, P., Kindt, M.R., Rader, D.J. & Brewer, H.B., Jr. Three genetic variants of human plasma apolipoprotein A-IV. apoA-IV-1(Thr347----Ser), apoA-IV-0(Lys167----Glu,Gln360----His), and apoA-IV-3(Glu165----Lys). *J Biol Chem* **266**, 13513-8 (1991).
90. Deeb, S.S., Nevin, D.N., Iwasaki, L. & Brunzell, J.D. Two novel apolipoprotein A-IV variants in individuals with familial combined hyperlipidemia and diminished levels of lipoprotein lipase activity. *Hum Mutat* **8**, 319-25 (1996).
91. Bunker, C.H. et al. Weight threshold and blood pressure in a lean black population. *Hypertension* **26**, 616-23 (1995).
92. Bunker, C.H. et al. Correlates of serum lipids in a lean black population. *Atherosclerosis* **123**, 215-25 (1996).
93. Hamman, R.F. et al. Methods and prevalence of non-insulin-dependent diabetes mellitus in a biethnic Colorado population. The San Luis Valley Diabetes Study. *Am J Epidemiol* **129**, 295-311 (1989).
94. Rewers, M. et al. Is the risk of coronary heart disease lower in Hispanics than in non-Hispanic whites? The San Luis Valley Diabetes Study. *Ethnicity Dis* **3**, 44-54 (1993).
95. Harris, M.R. et al. Racial differences in the distribution of a low density lipoprotein receptor-related protein (LRP) polymorphism and its association with serum lipoprotein, lipid and apolipoprotein levels. *Atherosclerosis* **137**, 187-95 (1998).
96. Richmond, W. Preparation and properties of a cholesterol oxidase from Nocardia sp. and its application to the enzymatic assay of total cholesterol in serum. *Clin Chem* **19**, 1350-6 (1973).
97. Warnick, G., Benderson, J. & Albers, J. Quantitation of high density lipoprotein subclasses after separation by dextran sulfate and magnesium precipitation. *Clin Chem* **28**, 1574 (1982).