SEQUENCE VARIATION IN THE CD36 GENE AND ITS RELATIONSHIP WITH PLASMA HDL CHOLESTEROL LEVELS

by

Sarah Caitlin Hughes

B.A., Malone University, 2007

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

2010

UNIVERSITY OF PITTSBURGH

Graduate School of Public Health

This thesis was presented

by

Sarah Caitlin Hughes

It was defended on

April 13th, 2010

and approved by

Committee Member:
F. Yeşim Demirci, M.D.
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

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

SEQUENCE VARIATION IN THE CD36 GENE AND ITS RELATIONSHIP WITH PLASMA HDL-CHOLESTEROL LEVELS

Sarah Caitlin Hughes, M.S.

University of Pittsburgh, 2010

Heart disease (HD) is a primary public health concern, with HD being one of the leading causes of death every year in the United States. Many risk factors influence HD, including lipid levels, and studies have shown that higher levels of plasma high density lipoprotein (HDL) cholesterol have a protective effect against HD. Recent genome-wide linkage scans have associated a locus on chromosome 7, harboring *CD36*, as being involved in components of the metabolic syndrome, including HDL-C levels. Therefore, identifying variation in this gene affecting HDL-C levels is of great public health importance. The "common variant-common disease" hypothesis has been tested by a limited number of studies through common SNP genotyping with inconsistent results. To date, no studies to our knowledge have evaluated *CD36* using the "rare variant-common disease" hypothesis. The aim of this study was to further evaluate the role of common and rare variation in *CD36* by sequencing individuals having extremely low and high HDL-cholesterol levels in two populations, U.S. Non-Hispanic Whites (NHWs), and African Blacks. In our initial sequence analysis, 343 variants were identified in *CD36*, 168 of which were previously unreported in the SeattleSNPs database. According to preliminary analysis of the sequencing data, our findings support the associations of three SNPs with HDL-C levels reported in the literature. No striking difference was noticed between the distribution of rare variants between high and low HDL-C groups. We identified four common variants (MAF≥5%) in our sequencing data from our small sample that displayed statistically significant differences in MAF between the low and high HDL-C groups but have not been confirmed yet by genotyping in the

entire NHW and Black populations while thirteen common variants had *p*-values between 5-10%, which may be statistically significant due to the small sample size. To date, screening data was compiled for the entire NHWs and Black samples for a total of nineteen common variants. None of these variants displayed a significant p-value in our entire NHW and Black samples. Additional variants identified in sequencing remain to be screened in the entire NHW and Black samples.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to begin by thanking Dr. M. Ilyas Kamboh for the opportunity to participate in this exciting research during my graduate career. I have learned more than I could have ever imagined during the course of my project, and appreciate you allowing my to be a part of your lab.

I also would like to thank Dr. F. Yeşim Demirci for the countless hours of guidance and instruction that she has provided me during the course of my project. I have enjoyed working with you so very much, and appreciate the patience with which you have educated and mentored me over the past two years.

Thank you to Dr. David Finegold, whose love of genetics and clinical work has been such an inspiration to me. Thank you for making your classes such a wonderful environment for learning, and also for your advice and guidance whenever I have needed it. I would also like to thank Amy Dressen and Dr. Candace Kammerer for all of their assistance on the statistical analysis of this project.

Thank you to my coworkers and fellow students in Dr. Kamboh's lab, you have made these past two years so much fun. Thank you to Jatinder Singh, Dr. Oussama Khalifa, Dilek Pirim, and Zahra Rahman for all of their assistance with my project. Thank you to Lauren Burns, my fellow Genetic Counseling student and lab coworker, who I have always been able to go to with my joys and frustrations as we have gone through the past two years of classes and lab

work. Thank you to Ryan Minster for always being there to answer my questions with such kindness and patience.

I would also like to thank Ms. Elizabeth Gettig and Dr. Robin Grubs, the co-directors of the Genetic Counseling Program, for the education and support that have given me throughout my graduate training and my thesis work. Thank you to all of my classmates in the Genetic Counseling Program, all of you have become my family over the past two years, and your friendship means the world to me.

I would also like to thank my family. My mom and dad have been there as a never-ending source of support over the past two years, always sharing in my joys and my frustrations- no matter how trivial the frustrations may be! Thank you to my husband, Tim, for all of your love and encouragement, especially during my final thesis preparations- I don't know where I'd be without you. Thank you to my brother and his wonderful fiancé Jamie for all the love and laughter we've shared, as well as all my friends who have always been able to lean on.

1.0 BACKGROUND AND SIGNIFICANCE

1.1 CARDIOVASCULAR DISEASE AND CHOLESTEROL LEVELS

Over 16 million people in the United States are living with coronary heart disease (CHD), with a total of 80 million American adults living with one or more types of cardiovascular disease (CVD) (Lloyd-Jones et al., 2009). This translates to approximately 1 in 3 American adults living with CVD, with the estimated direct and indirect cost of CVD in the U.S for 2009 being $475.3 billion (Lloyd-Jones et al., 2009). Mortality data from the National Center for Health Statistics (NCHS) indicates that CVD was the underlying cause of 1 out of every 2.8 deaths in 2005, and CVD has accounted for more deaths than any other major cause of death since 1900, aside from the year 1918 (NCHS, 2007).

Many risk factors have been identified that influence the risk to develop CHD. These risk factors include excessive alcohol consumption, abnormal serum cholesterol levels, body mass index, diabetes, smoking, lack of regular physical activity, consumption of less than five servings of fruits and vegetables per day, psychosocial index, and hypertension (Lloyd-Jones et al., 2009; American Heart Association, 2003). It is important to note that blood pressure and hypertension are particularly important factors in black populations because compared to whites, blacks develop hypertension earlier in life and their average blood pressures are much higher – resulting in a 1.5-times greater risk of CHD death (Lloyd-Jones et al., 2009).

Genetic factors, such as genes influencing lipid metabolism, make a significant impact on the lipid profile of an individual. This can result in a risk factor for CHD if the lipid profile is abnormal. Studies have shown that family history of a parent or sibling with CVD increases the risk of CVD two-fold, and further data has found that anywhere from 50-80% of the variation of the lipid profile is under genetic control (Berg et al., 1987; Boes et al., 2009; Lloyd-Jones et al, 2004; Murabito et al., 2005). This information serves to illustrate the importance of understanding the genetic influences on lipid metabolism.

When considering the lipid profile, cholesterol levels are of particular importance due to the fact that the levels of both low-density lipoprotein (LDL) cholesterol and high-density lipoprotein (HDL) cholesterol have been found to impact the risk for CVD. According to the American Heart Association (2009), total cholesterol should ideally be less than 200mg/dL, with 200-239 mg/dL being borderline high risk and over 240 mg/dL being high risk. LDL-C levels of 129 mg/dL and less are considered optimal/near optimal, 130-189 mg/dL are considered borderline high/high, and anything above 190 mg/dL is considered very high. When measuring HDL-C, the average male range is from 40 to 50 mg/dL and the average female range is from 50 to 60 mg/dL. Low levels (less than 40 mg/dL for men, less than 50 mg/dL for women) are a risk factor for CVD, and HDL-C of 60 mg/dL or higher gives some protection against CVD.

Data from 2001–2004 showed the serum total crude mean cholesterol level in adults was 201 mg/dL for men and 203 mg/dL for women, and the age-adjusted prevalence of high LDL cholesterol adults from 1999–2004 was 25.3% (Lloyd-Jones et al., 2009). As of 2006 the percentages of individuals with high LDL-C were 31% in non-Hispanic white (NHW) males, 33.7% in NHW females, and 36.2 % in non-Hispanic black (NHB) males, and 27.4% in NHB females (Lloyd-Jones et al., 2009).

According to the NCHS, the mean level of HDL-C for American adults from 2005-2006 was 54.6 mg/dL (Lloyd-Jones et al., 2009). In individuals with HDL-C levels less than 40 mg/dL, which is a known risk factor for CHD, the percentages was 24.9% for NHW males, 6.5% for NHW females, 13.5% for NHB males, and 6.1% for NHB females (Lloyd-Jones et al., 2009). It is important to note the difference between the sexes, with males having a significantly higher proportion of individuals with low HDL-C compared to females.

In addition to the evidence supporting the role of a favorable LDL-C level to prevent CVD, studies also show the importance of HDL-C levels in influencing CVD and CHD risk in both men and women (Boes et al., 2009; Gordon et al., 1977). The Framingham Heart Study showed that HDL-C levels are inversely correlated with the risk for CVD, with high levels of HDL being preventative against CVD (Castelli et al., 1986). This inverse relationship has been quantified, with each 1 mg/dL increase in HDL-C being associated with a 2-3% decrease in the risk for developing CVD (Boes et al., 2009; Lewington et al., 2007).

HDL-C is influenced by a number of environmental factors, which can either raise HDL levels or lower them. Factors that are associated with lower HDL-C levels are BMI and smoking. Factors that are associated with higher HDL-C levels include physical activity, healthy diet, and estrogen (evidenced by the fact that women have higher HDL-C levels than men). In addition to environmental factors, the many genes that are involved in the structure, metabolism, and production of HDL may have an effect on CVD risk as well.

1.2 HDL-CHOLESTEROL METABOLISM

HDL is a mixture of lipoprotein particles that range in density from 1.063-1.21 g/mL based on the lipid composition (Tsompanidi et al., 2009). The particle is composed of an esterified cholesterol and triglyceride core, surrounded by an ampipathic layer of free cholesterol, apolipoproteins, and phospholipids (Boes et al., 2004). The main protein component of HDL is apolipoprotein A1 (apoA-1), which is involved in the biogenesis and function of HDL (Tsompanidi et al., 2009). Serum HDL-C levels are influenced by a complex set of interactions, and variation in the components in these interactions can have a major impact on HDL and its properties. Some of these components include apoprotein concentrations; function of enzymes, transport proteins, receptors; other lipoproteins and their clearance from plasma (Boes et al., 2009; Cavelier et al., 2006; Tsompanidi et al., 2009).

As described above, studies have found that the level of serum HDL-C can act as both a protection against and risk factor for CVD. The protective role of HDL-C has yet to be completely defined; however, it is clear that HDL-C particles exhibit multiple antiatherogenic effects (Kontush and Chapman, 2006). The first hypothesis is that reverse cholesterol transport is the bases of these antiatherogenic effects. Reverse cholesterol transport is a process in which HDL-C removes cholesterol from peripheral cells (such as macrophages) and delivers this cholesterol to the liver or other tissues in need of large amounts of cholesterol. This is an important process that relieves the peripheral cells from cholesterol burden and prevents the excess accumulation of cholesterol in the arterial wall (Tsompanidi et al., 2009). Reverse cholesterol transport is accomplished using three routes (Figure 1). In the first route, large HDL-C particles with multiple copies of apoE can be taken up by the liver via the LDL-C receptor (Bruce et al., 1998). In the second route, accumulated cholesteryl esters from HDL-C are

selectively taken up by the liver via the SR-BI receptor that is expressed primarily in the liver and steroidogenic tissues (Acton et al., 1996). In the third route, cholesteryl esters are transferred by the cholesteryl ester transfer protein (CETP) from HDL-C to triglyceride-rich lipoproteins (Bruce et al., 1998). These processes are summarized in Figure 1 from Bruce, Chouinard, and Tall (1998).

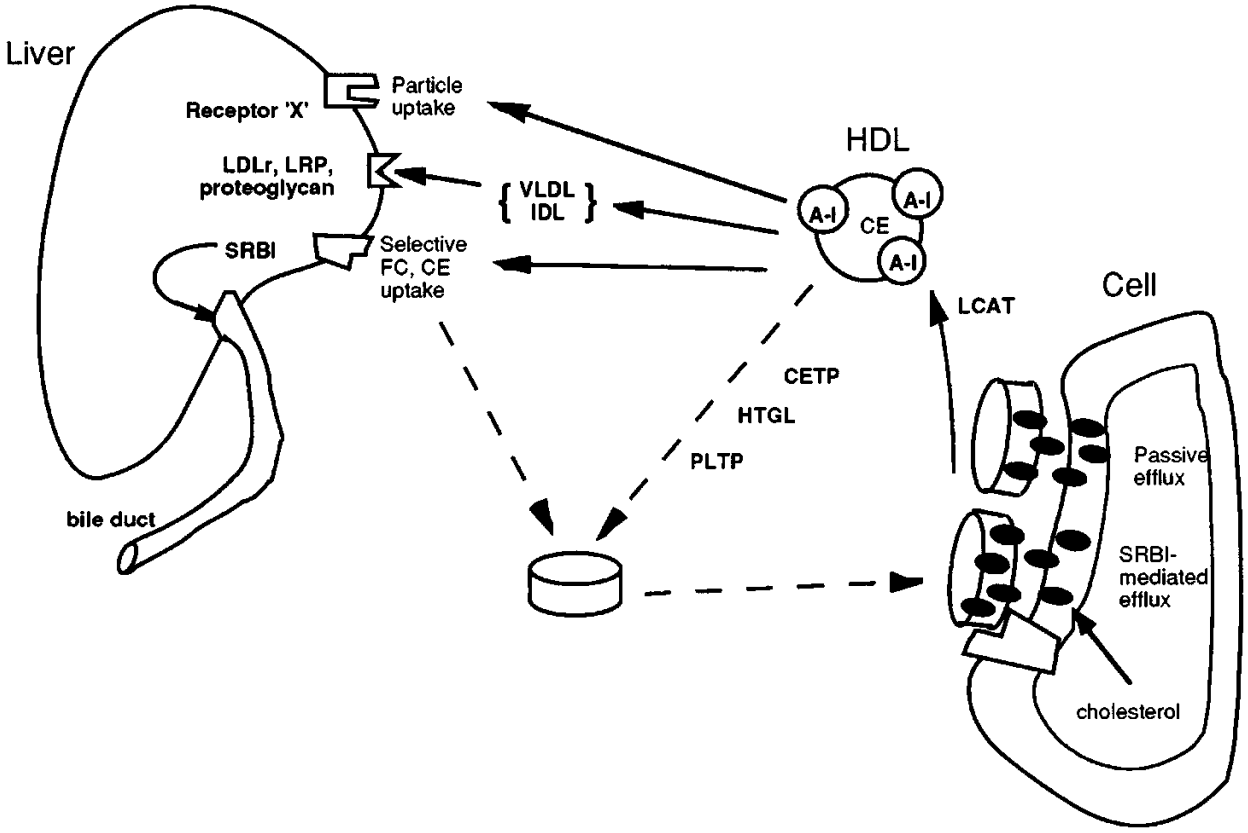


Figure 1. The three pathways of reverse cholesterol transport. Solid lines represent cholesterol movement, and dashed lines represent the regeneration of nascent HDL-C from Bruce et al. (1998)

Aside from reverse cholesterol transport, there are several other ways in which HDL is hypothesized to have antiatherogenic effects. HDL-C has been found to have anti-oxidative properties due to associated anti-oxidative enzymes, which provide direct or indirect protection against oxidation of LDL-C (Tsompanidi et al., 2009). HDL-C also protects against inflammation, a large component in the pathogenesis of atherosclerosis, by expressing anti-inflammatory activity in different pathways (Boes et al., 2009; Tsompanidi et al., 2009). HDL-C has been reported to stimulate the activity of endothelial nitric oxide synthase (eNOS), which diminishes endothelial dysfunction that contributes to the development of atherogenesis (Nofer et al., 2004). Another property of HDL-C contributing to atheroprotection is its ability to stimulate SR-BI-dependent endothelial cell migration, which initiates recruitment of endothelial progenitor cells into the intimal layer of the artery at the site of endothelial injury (Tso and Martinic, 2009).

1.3 GENETIC DETERMINANTS OF HDL-CHOLESTEROL LEVELS

Studies on the heritability of the lipid profile, as well as the heritability of HDL-C itself, estimate that HDL-C is under considerable genetic control with heritability estimates of up to 80% (Boes et al., 2009). Many of the studies to identify what genetic components were related to HDL-C levels were performed using twin studies (which predict heritability of HDL-C to be 50%), family studies, and linkage analysis, with 50 genes being associated with HDL-C (Holleboom et al., 2008). Studies have also examined the heritability of other portions of the lipid profile, with the heritability estimate for LDL-C being ~40% and the heritability for triglycerides being ~36-80% (Krauss, 2008; Shea et al., 2009).

However, the more recent development of genome-wide association (GWA) studies has allowed researchers to travel outside the hypothesis-driven search for genetic factors influencing the lipid profile, which were often based on limited knowledge of the processes involved in lipid metabolism (Boes et al., 2009). These GWA studies allowed the detection of new genes which may be involved in lipid metabolism. Prior to 2009, eight GWA studies on HDL-C levels have been available (Willer et al., 2008; Kathiresan et al., 2008b; Kooner et al., 2008; Wallace et al., 2008; Heid et al., 2008; Kathiresan et al., 2007; Aulchenko et al., 2009; Chasman et al., 2008). These studies have implicated the following genes in the regulation of HDL-C levels, which had already been identified in prior functional and association studies: *CETP, LPL, LIPC, LIPG, ABCA1, LCAT* and the *APOA1/C3/A4/A5* gene cluster (Boes et al., 2009). However, these association only account for about 8-10% of the variation seen in HDL-C levels, which indicates that the majority of the additional genetic factors involved in HDL-C levels have not yet been identified (Boes et al., 2009; Willer et al., 2008). Several other genes have been implicated in affecting HDL-C levels, including *CD36*. Since the focus of this study is *CD36*, the following sections provide a brief overview of the molecular, biological, and genetic aspects of *CD36*.

1.4 MOLECULAR ASPECTS OF *CD36*

The *CD36* gene is located at 7q11.2, and consists of 15 exons spanning at least 32 kb (Figure 2) (Fernandez-Ruiz et al., 1993; Armesilla and Vega, 1994). The *CD36* protein is a cell-surface glycoprotein, which is composed of a single polypeptide chain that ranges from 78-88 kDa depending on the cell type, which results in a 50 kDa chain after post-translational glycosylation at 10 potential N-linked glycosylation sites (Alession et al., 1991; Oquendo et al., 1989; Gruarin

et al., 1997). The *CD36* glycoprotein is characterized by two hydrophobic transmembrane domains spanning amino acids 7-34 and 440-466, with one located at the N-terminus and the other located at the C-terminus (Oquendo et al., 1989; Gruarin et al., 1997; Daviet et al., 1995). *CD36* is also predicted to include two short cystoplasmic tails as well, which extend from amino acids 1-6 and 467-472 (Oquendo et al., 1989). The highly glycosylated central hydrophilic domain, which is rich in N-glycan acceptor sites and monoclonal antibody epitopes, seems to lie extracellularly. However, the exact topology of the molecule is still unclear (Gruarin et al., 1997; Tao et al., 1996; Pearce et al., 1994).

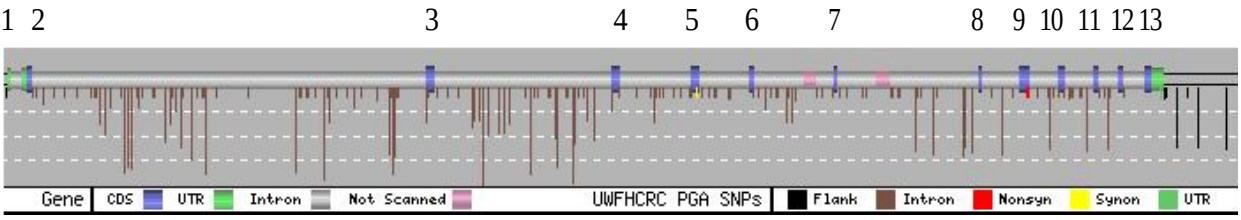


Figure 2. Physical Map of the CD36 gene from the SeattleSNP database. The legend is located at the bottom of the figure, and exons are labeled above.

There are ten cysteine amino acids in the *CD36* sequence, with six of them (amino acids 243, 272, 311, 313, 322 and 333) being clustered in the C-terminal portion of the extracellular domain. The remaining four cysteine amino acids are distributed equally at the N-terminus (amino acids 3 and 7) and the C-terminus (amino acids 464 and 466) (Gruarin et al., 1997). A total of 11 exons (4-13 and part of 14) are thought to encode this extracellular domain, and the large number of exons encoding this domain may be a consequence of the multiple interactions

in which *CD36* seems to be involved (4-20). The large amount of exons encoding the extracellular domain also suggests that the polypeptide chain may be organized into discrete portions that may act independently or in combination to create independent structural and/or functional domains (Armesilla and Vega, 1994).

The *CD36* protein also undergoes posttranslational modification. Intrachain disulphide bonds are formed in the endoplasmic reticulum, and these bonds appear to be essential for maturation, intracellular transport and structural biogenesis of the protein (Gruarin et al., 1997). Palmitoylation also occurs, which involves the covalent, post-translational attachment of the 16-carbon saturated fatty acid palmitate through a thioester linkage to cysteine residues (Tao et al., 1996). Four cysteine residues at the very N and C termini of the protein have been identified as the sites of palmitoylation. (Tao et al., 1996) The possible role of this lipid modification is in the function of *CD36* as a lipid receptor, binding oxidized low density lipoprotein (oxLDL), fatty acids, and anionic phospholipids (Tao et al., 1996).

1.5 BIOLOGICAL FUNCTION OF *CD36*

CD36 is a class B scavenger receptor that is expressed by erythrocyte precursors, mature monocytes, platelets, microvascular endothelial cells, adipose tissue, liver tissue, and mammary epithelial cells (Thorne et al., 2000; Greenwalt et al., 1992). It is a receptor for thromposondin, collagen, apoptotic cells, *P. falciparum*-infected red blood cells, and can trigger the activation of monocytes and platelets (Endemann et al., 1993; Xiaowei et al., 2004).

CD36 is also known to act as a receptor for fatty acids (FA), LDL-C, HDL-C, and very low density lipoprotein (VLDL) (Calvo et al., 1998). *CD36* has been identified as possibly being involved in scavenging OxLDL from the blood via Kuppfer cells in the liver (Endemann et al., 1993). Because *CD36* has been found to be an OxLDL receptor that likely mediates a response in the initial stages of lipoprotein oxidation, *CD36* may be implicated in atherogenesis due to its possible involvement in foam cell formation (Endemann et al., 2007). *CD36* may also be involved in processes that result in the differentiation of monocytes and the accumulation and adhesion of macrophages, monocytes, endothelial cells, and platelets in atherosclerotic lesions (Nagy et al., 1998; Tontonoz et al., 1998; Masuda and Ross, 1990; Endemenn et al., 1993). Due to these properties of *CD36*, it has been hypothesized that a reduction in *CD36* activity is protective against atherosclerosis.

However, despite this hypothesis that CD36 may be protective against CVD, other research has linked *CD36* to alterations in the lipid profile that have the potential to increase CVD risk. Studies have shown the common haplotypes in *CD36* are associated with abnormalities in serum FA and triglyceride (TG) levels, and an increased risk for CAD (Ma et al., 2004). Miyaoka et al. (2001) showed that complete *CD36* deficiency may be associated with an abnormal lipid profile, including elevated fasting plasma TG and LDL-C levels, in combination with reduced HDL-C levels (Miyaoka et al., 2001 and Yanai et al., 2000). *CD36* activity has also been shown to correlate with HDL-C levels. Madden et al. (2008) showed that common *CD36* polymorphisms impact HDL-C levels after fish oil supplementation, while Love-Gregory et al. (2008) linked SNPs in the *CD36* gene with HDL-C concentrations.

Physiological studies of *CD36* have also implicated the gene in metabolic syndrome and diabetes. Love-Gregory et al. (2008) found that certain variants in the *CD36* gene are associated with metabolic syndrome, a cluster of risk factors that increase susceptibility to CVD and type 2 diabetes. In a Dutch population, a promoter SNP was more common in subjects who had diabetes, and other studies have linked *CD36* with diabetes, insulin resistance, and low adiponectin in diabetic subjects (Corpeleijn et al., 2006; Leprêtre et.al, 2004).

1.6 *CD36* POLYMORPHISMS

A total of 187 variants in the *CD36* gene have been reported and are summarized in the Seattle SNP Database for populations of African (n=24) and European (n=23) descent. Twenty of these are insertion/deletion polymorphisms and 167 are substitutions, with 5 exonic and 182 intronic. Of these variants, 167 are present in the African population and 80 are present in the European population. The African population has 70 SNPs with a minor allele frequency less than 5% (MAF<5%), and the European population has 19 SNPs with a MAF<5%.

Gelhaus et. al (2001) sequenced *CD36* in 12 individuals from West Africa and identified 24 variants, 21 of them being novel. They identified 5 variants in the promoter region of the gene (-220A>G, -144G>T, -53G>T, -50G>T, -2A>C), 12 variants in exonic regions (158C>A, 656G>A, 809A>G, 985_986insAAT, 1101T>C, 1264T>G, 1712_1715delAGTA, 1910G>T, 1934A>G, 2065C>T, 2512A>C, 2549G>A), and 7 variants in intronic regions (990+7_8insAGTA, 1107+29_30CT>TC, 1296-119T>C, 1414-13C>A, 1415-150A>C, 1489-121A>T, 1489-5_6insAT). These variants included three single-nucleotide substitutions causing non-conservative amino acid exchanges E123K, T174A, and I271T, and the E123K variant was

located within the putative ligand-binding domain for oxidized low density lipoprotein, while the other substitutions resided outside any of the binding sites for reaction partners mapped on CD36 so far. They also found a three base pair insertion resulting in the addition of an asparagine residue (N232-233ins). Five additional SNPs were located in the promoter region, with -144G-->T, -53G-->T, and -2A-->G altering putative binding sites for the transcription factors purine factor (PuF), phorbol ester-responsive element AP-2, and CCAAT/enhancer-binding protein. A G-->T exchange at position -50 appears to introduce a new recognition site for PuF (Gelhaus et al., 2001)

1.7 *CD36* GENOTYPE ASSOCIATIONS WITH PLASMA HDL-CHOLESTEROL

To date, GWA studies have not identified *CD36* as a major gene associated with HDL levels. However, several genomewide linkage scans have linked a region of chromosome 7 (7q11.2–7q21.11) with components of the MetS, including insulin resistance and dyslipidemia (Love-Gregory et al., 2008). This region harbors the *CD36* gene, which has prompted researchers to further investigate the relationship of common SNPs and haplotypes with the lipid profile and other components of MetS (Love-Gregory et al., 2008; Madden et al., 2008; Goyenechea et al., 2008; Ma et al., 2004).

To our knowledge, four studies have examined the association of *CD36* SNPs with plasma HDL-C levels in non-diabetic Caucasians, healthy middle-aged male, African American, and Spanish populations. These studies examined 42 SNPs, and several SNPs revealed an association with HDL-C levels. Table 1 summarizes these associations with HDL-C and the *p*-value identified, and information from these studies on associations with LDL-C and TG levels

is included as well. These studies only examined a select set of SNPs and did not examine the entire genetic variation of the *CD36* gene in relationship to plasma lipid levels, which is why we undertook this comprehensive study.

Table 1. Prior Genotype Associations

SNP	Study	Association p-value			
		Total chol	**LDL-C**	**HDL-C**	**TG**
rs2151916	Goyenechea (2008)	0.03	0.02	0.01	
rs3211938	Love-Gregory (2008)			> 0.0002	0 .006
rs10499859	Love-Gregory (2008)			0.00028	
rs13438282	Love-Gregory (2008)			0.0038	
rs1358337	Love-Gregory (2008)			0.00066	
rs1054516	Love-Gregory (2008)			0.003	
rs1049654	Love-Gregory (2008)			0.0028	
rs3211909	Love-Gregory (2008)			0.022	
rs3211849	Love-Gregory (2008)			0.029	
rs3211913	Love-Gregory (2008)			0.039	
rs13246513	Love-Gregory (2008)			0.038	
rs3173798	Love-Gregory (2008)			0.033	
rs3211870	Love-Gregory (2008)			0.0079	
rs3211842	Love-Gregory (2008)			0.00074	
rs9784998	Love-Gregory (2008)			0.015	
rs3211810	Love-Gregory (2008)			0.044	
rs3211868	Love-Gregory (2008)			0.011	
-31118	Madden (2008)		<0.05	<0.05	
25444	Madden (2008)			<0.01	
30294	Madden (2008)			<0.01	
-33137	Madden (2008)			<0.001	
-33137	Ma (2004)				0.06

1.8 SPECIFIC AIMS

Due to the importance of HDL-C to cardiovascular health and CHD, determining the sequence variation in the *CD36* gene that affect HDL-C levels in the general population would increase our understanding of HDL-C levels and increase our ability to detect and predict individuals at an increased risk for CHD based on HDL-C levels. The objective of this study was to evaluate the genetic variation in the *CD36* gene in relation to HDL-C levels in two well-characterized samples of 788 African blacks from Benin City, Nigeria and 624 U.S. Non-Hispanic Whites (NHWs) from Colorado. Sequencing of the CD36 gene in individuals from our population with extreme HDL-C levels (upper and lower fifth percentiles) will allow to test the "common trait- rare variant" hypothesis, while screening common tagSNPs in our entire population will allow us to test the "common trait- common variant" hypothesis.

1.8.1 Aim 1

Resequence the *CD36* gene in individuals with HDL-C in the upper 5th percentile (47 NHWs and 48 Blacks) and in the lower 5th percentile (48 NHWs and 47 Blacks) to identify common and rare variants.

1.8.2 Aim 2

Screen common *CD36* tagSNPs in the entire NHW and Black population (individuals with varying levels of HDL-cholesterol) in order to determine the distribution of common variants throughout our population.

1.8.3 Aim 3

Determine the association of both the rare and common *CD36* variants identified in Aim 1 and Aim 2 with plasma HDL-C levels in our NHW and Blacks populations.

2.0 SUBJECTS AND METHODS

2.1 SUBJECTS

2.1.1 Study Samples

Samples from the Non-Hispanic White (NHW) population were taken from the San Luis Valley Southern Colorado Diabetes Study. All subjects who were included in the current study were normoglycemic, and a more thorough description of the population can be found in Rewers et al. (1993) and Hamman et al. (1989). Samples from the African Black population were recruited in Benin City, Nigeria as a part of a study on CHD risk factors in Blacks. Subjects from Benin City were recruited from Junior and Senior staff in government (representing different salary grades). Demographic and health information was gathered from participants, and this detailed information about the study population can be found in Bunker et al. (1995, 1996).

In order to measure the fasting total serum cholesterol, the esterase-oxidase method was used, and then HDL-C was determined enzymatically after dextran sulfate magnesium precipitation (Harris et al., 1998; Richmond et al., 1973). Table 2 summarizes the population characteristic data for the subjects used in this study. 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 2. Population characteristic data (mean SD) of entire NHW and Black populations (Harris et al., 1998)

Variable	*NHW (*n=623)		*Blacks* (n=788)	
	Men	Women	Men	Women
Age (years)	52.9 ± 0.6	52.4 ± 0.6	42.5± 0.4	38.7± 0.4
Sex	295	328	495	293
BMI (kg/m2)	26.2 ± 0.20	24.8 ± 0.2	22.0 ± 0.2	24.3± 0.3
LDL (mg/dl)	139.8± 2.0	134.7± 2.0	104.7 ± 0.1	117.3 ± 0.1
HDL (mg/dl)	43.9 ± 0.6	56.3 ± 0.7	45.9 ± 0.6	50.6 ± 0.7
Triglycerides (mg/dl)	147.6 ± 4.1	128.2 ± 2.9	77.8 ± 1.7	62.8 ± 1.4
Total Cholesterol (mg/dl)	213.7 ± 2.2	217.7 ± 2.1	167.1 ± 1.6	181.3 ± 2.2

2.1.2 Subset of the Study Population used for Sequencing

Subjects with serum HDL-C in the upper 5th percentile (47 NHW and 48 Black) and in the lower 5th percentile (48 NHW and 47 Black) were selected to be screened for common and rare variants by re-sequencing of DNA samples for the entire *CD36* gene. Of 95 NHW individuals, 47 were females (23 with high HDL-C levels and 24 with low HDL-C levels) and 48 were males (24 with high HDL-C levels and 24 with low HDL-C levels). Of 95 Black individuals, 48 were females (24 with high HDL-C levels and 24 with low HDL-C levels) and 47 were males (24 with high HDL-C levels and 23 with low HDL-C levels). Table 3 shows a summary of the population characteristics including age, gender, BMI, LDL-C, HDL-C, total cholesterol, and triglyceride levels in NHWs and Blacks for both high HDL-C and low HDL-C subgroups.

Table 3. **Population characteristic data (mean SD) of Black and NHW samples used for DNA sequencing**

Variable	*NHW (*n=95)			*Blacks* (n=95)		
	High HDL(n=47)	Low HDL(n=48)	p-value	High HDL(n=47)	Low HDL(n=48)	p-value
Age (years)	55.45 ± 9.8	53.03 ± 10.54	0.25	41.26 ± 8.72	40.87 ± 7.16	0.8
Sex (M/F)	24/23	24/24	0.92	24/24	23/24	0.92
BMI (kg/m2)	23.17 ± 3.17	27.35 ± 3.90	<0.001	22.06 ± 4.71	23.91 ± 5.51	0.08
LDL (mg/dl)	126.84 ± 46.95	136.95 ± 41.28	0.28	112.55 ± 39.75	95.04 ± 28.28	0.02
HDL (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
Total Cholesterol (mg/dl)	227.34 ± 51.76	208.81 ± 44.65	0.06	201 ± 39.68	141.68 ±31.03	<0.001

2.2 DNA SEQUENCING

We used publicly available information from the SeattleSNPs database (http://pga.mbt.washington.edu/) to order the M13-tagged sequencing primers which produced 44 overlapping resequencing amplicons (Table 4). Due to polymerase chain reaction (PCR) related technical difficulties, we designed new primers for amplicons 14, 19, and 20 using Primer3 software (http://frodo.wi.mit.edu/primer3/). Uppercase letters in Table 4 indicate gene specific sequences, and lowercase letters indicate the M-13 sequence. The base pair length of each amplicon is located in parenthesis.

Table 4. *CD36* Polymerase chain reaction (PCR) primers

PCR Amplicon	Forward Primer	Reverse Primer
2 (863)	5'-tgtaaaacgacggccagtATGTCTTGCTGTTGATTTGTGAA-3'	5'-caggaaacagctatgaccTCGCATCATATAGAGTTGCAGTG-3'
3 (799)	5'-tgtaaaacgacggccagtTGGAGGTATTCTAATGCCAGTTG-3'	5'-caggaaacagctatgaccCAATCAACGTTTCTGATGAGTGA-3'
4 (798)	5'-tgtaaaacgacggccagtGAAGCTTCATATTGGAATCTTAGAAA-3'	5'-caggaaacagctatgaccCCATCCCACTATTAGATAAGCCC-3'
5 (885)	5'-tgtaaaacgacggccagtACAGAGGGCTGACTGTATTGTGT -3'	5'-caggaaacagctatgaccATCTTCACTGCATTTGGTAGCAT-3'
6 (1,007)	5'-tgtaaaacgacggccagtTGAGACTCTAGAATTGAATTGGAA-3'	5'-caggaaacagctatgaccTGAACTCTGACCTTTGAAACACA-3'
7 (683)	5'-tgtaaaacgacggccagtCTCAAGATGTCCAGTGAGTTATT-3'	5'-caggaaacagctatgaccGTACCATACATCTTGACCAATAT-3'
8 (777)	5'-tgtaaaacgacggccagtATTCAAACTCAAGGAGGTGGTAC-3'	5'-caggaaacagctatgaccCAACTCTATTAAGAATCACGGTC-3'
9 (493)	5'-tgtaaaacgacggccagtAGGGATGTCTCTGGTATCCTCAT-3'	5'-caggaaacagctatgaccGGTAATAAGCATAGTACCTGATA-3'
10 (836)	5'-tgtaaaacgacggccagtGGTTTCTTTGTTCCTGTTAGAGAA-3'	5'-caggaaacagctatgaccAACTATGTTGTGTTTGGCATGAA-3'
11 (1,035)	5'-tgtaaaacgacggccagtTTTAATAACGTAAGAACAACCCAAA-3'	5'caggaaacagctatgaccAAGTGGCCAACAATGAAATTAGA-3'
12 (952)	5'-tgtaaaacgacggccagtAAGTGTTCACGTTTATTGATCCC-3'	5'-caggaaacagctatgaccATGCTTCTGACGCTACCACTACT-3'
13 (962)	5'-tgtaaaacgacggccagtCACATACACATTAGCCAAGTGAGA-3'	5'-caggaaacagctatgaccGAGGATGTTGAAGCTCAAAGCTA-3'
14* (854)	5'-CCTTGAGAGGCACTTGATGA-3'	5'-GGAAGAATGCCCAGGTTAAA-3'
15 (914)	5'-tgtaaaacgacggccagtGTTTCTCAGAGCCTCAGTGTGAT-3'	5'-caggaaacagctatgaccGACAAGATATTGCTCTGTCACCC-3'
16 (767)	5'-tgtaaaacgacggccagtTGCTTCAGCTCAGGAGTTCAA-3'	5'-caggaaacagctatgaccTCAAGAGTTGGACACTTCAGAGG-3'
17 (996)	5'-tgtaaaacgacggccagtAAAGCGTCACTCTAAAGCTTGC-3'	5'-caggaaacagctatgaccCCTCAACTCTTTCATTCATTTGG-3'
18 (970)	5'-tgtaaaacgacggccagtACTCTGAAATATTCCTGCTGAGG-3'	5'-caggaaacagctatgaccCTCACCATGCCTGTTATTTCATT-3'
19* (985)	5'-TTCCCTCATGGATAATCACAAC-3'	5'-TTCCCTCATGGATAATCACAAC-3'
20* (808)	5'-AGTTCACTGCATCCTCAACCT-3'	5'-CTGAAGGAATTACAGCATCTTCA-3'
21 (965)	5'-tgtaaaacgacggccagtCTTTGATAGTGCATGTGTTGAGC-3'	5'-caggaaacagctatgaccCTGTCATATTTGAATGCCTGTGA-3'
22 (1,019)	5'-tgtaaaacgacggccagtTTTCTAGCCAACTTTGAATCCTC-3'	5'-caggaaacagctatgaccCAGGGTCTCTAGCAAACTAACGA-3'
23 (999)	5'-tgtaaaacgacggccagtATGCTACCATCTGCCGTACTTTA-3'	5'-caggaaacagctatgacc TATTGCCCACTGGTACAGCTACT-3'
24 (894)	5'-tgtaaaacgacggccagtCAGGAAGATGCTTAAGAAACAAG-3'	5'-caggaaacagctatgaccGAGGATGAGGAGGACTACGATTT-3'
25 (1,016)	5'-tgtaaaacgacggccagtACATCCACAGCACATCCTAATTC-3'	5'-caggaaacagctatgaccGACCAACTGTGGTAGTAACAGGG-3'
26 (830)	5'-tgtaaaacgacggccagtTAGGCTGCATCCCATATCTATCA-3'	5'-caggaaacagctatgaccGCCAAATGAAGTCATAGTCCAAC-3'
27 (1,127)	5'-tgtaaaacgacggccagtAAAGTTAGCCTAATGTTCACATCTCA-3'	5'-caggaaacagctatgaccGGTCTGTTCTATAGGTTGATGCC-3'

Table 4 Continued		
28 (814)	5’-tgtaaaacgacggccagtTTTCCCATACATATATTTCAGTACAACA-3’	5’-caggaaacagctatgaccTAATGGCTTAGGAAGCTGATTTG-3’
29 (798)	5’-tgtaaaacgacggccagtATATTCCAGTGGCATGACCTAAA-3’	5’-caggaaacagctatgaccTTGCCATGAGTTAAATCAACCTT-3’
30 (863)	5’-tgtaaaacgacggccagtGGAGTTGCAAAGCACTCCTAGTT-3’	5’-caggaaacagctatgaccACCTGTACCATTAATCATGTCGC-3’
31 (933)	5’-tgtaaaacgacggccagtGGCATCAGGTACATTGCAATAAG-3’	5’-caggaaacagctatgaccGTGCTGGGATTATAGACATGAGC-3’
32 (569)	5’-tgtaaaacgacggccagtTTGGATAAGGTGGTCAGAATGAG-3’	5’-caggaaacagctatgacc TTCCTAGTAGTTGAAAGCTTGCC-3’
33 (843)	5’-tgtaaaacgacggccagtGAGACATAGTCTGGCTCTGTTGC-3’	5’-caggaaacagctatgaccTATACCCAGCAACTACCATGGAC-3’
34 (895)	5’-tgtaaaacgacggccagtTGAGAGAGATTCTTGCTTATGGC-3’	5’-caggaaacagctatgaccGGCTGCTATTGTCAACAACAAAT-3’
35 (898)	5’-tgtaaaacgacggccagtATCAGCCATTAGGACAAATGAGA-3’	5’-caggaaacagctatgaccAAATCGAGTGGCAAATGATTAGA-3’
36 (897)	5’-tgtaaaacgacggccagtGGTTTCACCATGTAGGCCAG-3’	5’-caggaaacagctatgaccGGTACATTTCCATCGTTTACCAA-3’
37 (780)	5’-tgtaaaacgacggccagtGAGCCTTTACCACTACCCTTGAG-3’	5’-caggaaacagctatgaccTTCTTTGCATTTGCTGATGTCTA-3’
38 (1,017)	5’-tgtaaaacgacggccagtATTTGAATCCGACGTTAATCTGA-3’	5’-caggaaacagctatgaccCTTCATTTGGGTTTAATCCATCA-3’
39 (822)	5’-tgtaaaacgacggccagtGAAATCAACTGACATAATTCTTCCC-3’	5’-caggaaacagctatgaccACACACAATTTATTTGCCCAATC-3’
40 (799)	5’-tgtaaaacgacggccagtCCCAAATGAAGAAGAACATAGGA-3’	5’-caggaaacagctatgaccTGCAAATTGTAAAGTGAATCCAG-3’
41 (894)	5’-tgtaaaacgacggccagtGATTGGGCAAATAAATTGTGTGT-3’	5’-caggaaacagctatgaccACAGCTGCAAATACAAACCTCAT-3’
42 (841)	5’-tgtaaaacgacggccagtAAATCAAATTAGCAACAGCAACT-3’	5’-caggaaacagctatgaccCACCACACCAACACTGAGTAAGA-3’
43 (994)	5’-tgtaaaacgacggccagtGTGATAGGCAATTGAAGGGTTTA-3’	5’-caggaaacagctatgaccTCTTTCTTTAGCATGGTACTGGC-3’
44 (783)	5’-tgtaaaacgacggccagtTAAAGATGAATGAATGCCTGACC-3’	5’-caggaaacagctatgaccCTGACATCCAAGGATCATTAAGC-3’
45 (1,012)	5’-tgtaaaacgacggccagtCCTTCAATACCTGTCAGTAGCCT-3’	5’-caggaaacagctatgaccAGTGCCACCATTTCTTCAACTAA-3’

*Redesigned primer

These amplicons produced a ~30 kb genomic fragment (accession number AY095373) harboring the entire *CD36* gene as well as ~ 150bp of 5' flanking region and ~ 1830 bps of 3' flanking region. All amplicons were sequenced in both the forward and reverse directions. The PCR reaction and cycling conditions are listed in Table 5.

Table 5. PCR reaction and cycling conditions

PCR Reaction (total volume of 25 µL)		PCR Conditions
DNA	3.0 µL	1. 95º C for 5 minutes 2. 95º C for 45 seconds 3. 58-60º C for 45 seconds 4. 72º C for 1 minute - Repeat steps 2-4 for 40 cycles 5. 72º C for 10 minutes 6. Cool to 4º C
dH_20	11.75-13.25 µL	
10x BufferGold	2.5 µL	
$MgCl_2$ (25 mM)	1.5-3.0 µL	
dNTPs (1.25 mM)	3.8 µL	
Forward Primer (20 mM)	0.4µL	
Reverse Primer (20 mM)	0.4 µL	
TaqPolymerase (5U/ µL)	0.15 µL	

Gel electrophoreses was performed to confirm the successful amplification of PCR products prior to sequencing using the Invitrogen™ E-Gel® 96 2% with SYBR® Safe precaste gels. For samples that failed in the initial amplification, re-amplification and confirmation was performed using regular 2% agarose gels. The amplified samples were sent to a commercial lab for automated fluorescence-based cycle sequencing and capillary electrophoresis on ABI 3730x1DNA Analyzers (Genomic Services of Agencourt Bioscience Corporation, Beverly, MD). Sequencher version 4.8 (Gene Codes Corporation, Ann Arbor, MI), and Variant Reporter version 1.0 (Applied Biosystems, Foster City, CA) were used to analyze sequencing data.

2.3 GENOTYPING

Premade Taqman SNP genotyping assays were ordered for 19 SNPs, of which 18 were common (MAF≥5%) in at least one population (NHWs or Blacks). The SNPs were primarily selected from the list provided in the paper by Love-Gregory et al. (2008) and/or the information available through HapMap (www.hapmap.org). We initially ordered only available SNPs in order to being genotyping while we waited for sequencing to be completed. Table 6 provides the assay IDs of the genotyped SNPs and study samples in which the SNPs were screened. We also selected some SNPs located within ~44 kb 5' further upstream (*) or ~3kb 3' further downstream (**) from our targeted region for sequencing, corresponding to the largest *CD36* sequence available in Genbank (NC_000007.13) that harbors additional alternative noncoding exons. Figure 3 shows the full size of the CD36 gene, showing the targeted region for sequencing in our study and the areas not sequenced in our study but containing TaqMan SNPS that were genotyped in our entire population.

Table 6. TaqMan SNP genotyping assays

CD36 Reference SNP ID	Position	Taqman Assay Type	Location	Assay ID	Population
rs3211822	3157	Pre-Made	Intron	C__30605020_10	Both
rs3211842	7167	Pre-Made	Intron	C__26572331_10	Both
rs3211881	13094	Pre-Made	Intron	C___1803783_10	Both
rs1924	15932	Pre-Made	Intron	C__26572328_10	Both
rs3211908	18463	Pre-Made	Intron	C__31374621_10	Both
rs3173804	24426	Pre-Made	Intron	C___1803772_10	Both
rs1527483	26076	Pre-Made	Intron	C___8315330_10	Both
rs1405747	26512	Pre-Made	Intron	C__22275240_10	Both
rs3211956	28375	Pre-Made	Intron	C__27519229_10	Both
rs1334511		Pre-Made	Intron*	C___8314966_10	Both
rs1537593		Pre-Made	Intron*	C___1803815_10	Both
rs9641866		Pre-Made	Intron	C__30118565_10	Both
rs1194182		Pre-Made	Exon*	C___8315074_10	Both
rs17154155		Pre-Made	Intron*	C___1803841_10	Both
rs10499858		Pre-Made	Intron*	C__29685580_20	Blacks
rs1049654		Pre-Made	Exon*	C___8314408_1_	Both
rs1194181		Pre-Made	Intron*	C___8315073_20	Both
rs4731642		Pre-Made	Intron*	C__28030011_10	NHW
rs7755		Pre-Made	Exon**	C___8315318_10	Both

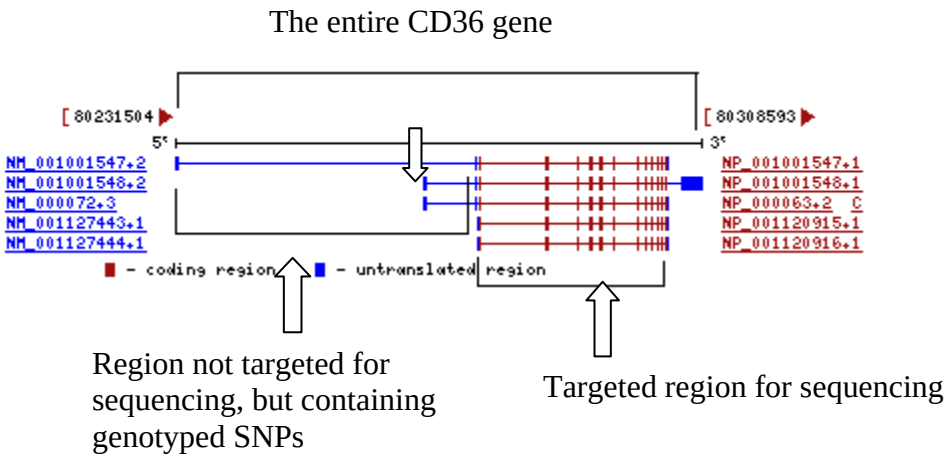


Figure 3. CD36 and its splice forms

2.3.1 Taqman Analysis

Genotyping of the sample populations was done using the TaqMan procedure, which involves the amplification of the product and endpoint reading using the ABI Prism 7900HT Sequence Detection System (Applied Biosystems). TaqMan Genotyping Master Mix and Assay is added to 384-well plates containing dried whole genome amplified DNA. The Genotyping Assay contains sequence-specific forward and reverse primers, one TaqMan minor groove binder (MGB) probe labeled at the 5' end with VIC dye and one TaqMan MGB probe labeled at the 5' end with FAM dye to detect the alternative alleles. On the 3' end of the probes, a nonfluorescent quencher (NFQ) is attached. PCR amplification is done using a PTC-200 MJThermal Cycler (Biorad) or a GeneAmp 9700 (Applied Biosystems), and the endpoint fluorescence reading is done on an ABI Prism 7900HT instrument. The cycling conditions are displayed in Table 7.

Table 7. TaqMan thermal cycler conditions

TaqMan Reaction (total volume of 5 µL)		PCR Conditions
dH_20	2.74 µL	1. 95º C for 10 minutes 2. 95º C for 45 seconds 3. 60º C for 1 minute -repeat steps 2-3 49x
Master Mix	2.5 µL	
Assay Primer	0.06 µL	

Each TaqMan MGB probe anneals to the target sequence harboring the SNP of interest during the annealing step. Then, AmpliTaq Gold DNA polymerase cleaves the probes that hybridize to the target sequence. During this process, the reporter dye is separated from the quencher dye, releasing fluorescence. Fluorescence is suppressed if probes do not hybridize to the target sequence and the reporter dye does not separate from the quencher dye. Because of the selective annealing of the TaqMan MGB probes, discrimination of alleles is possible. This process is shown in Figure 4.

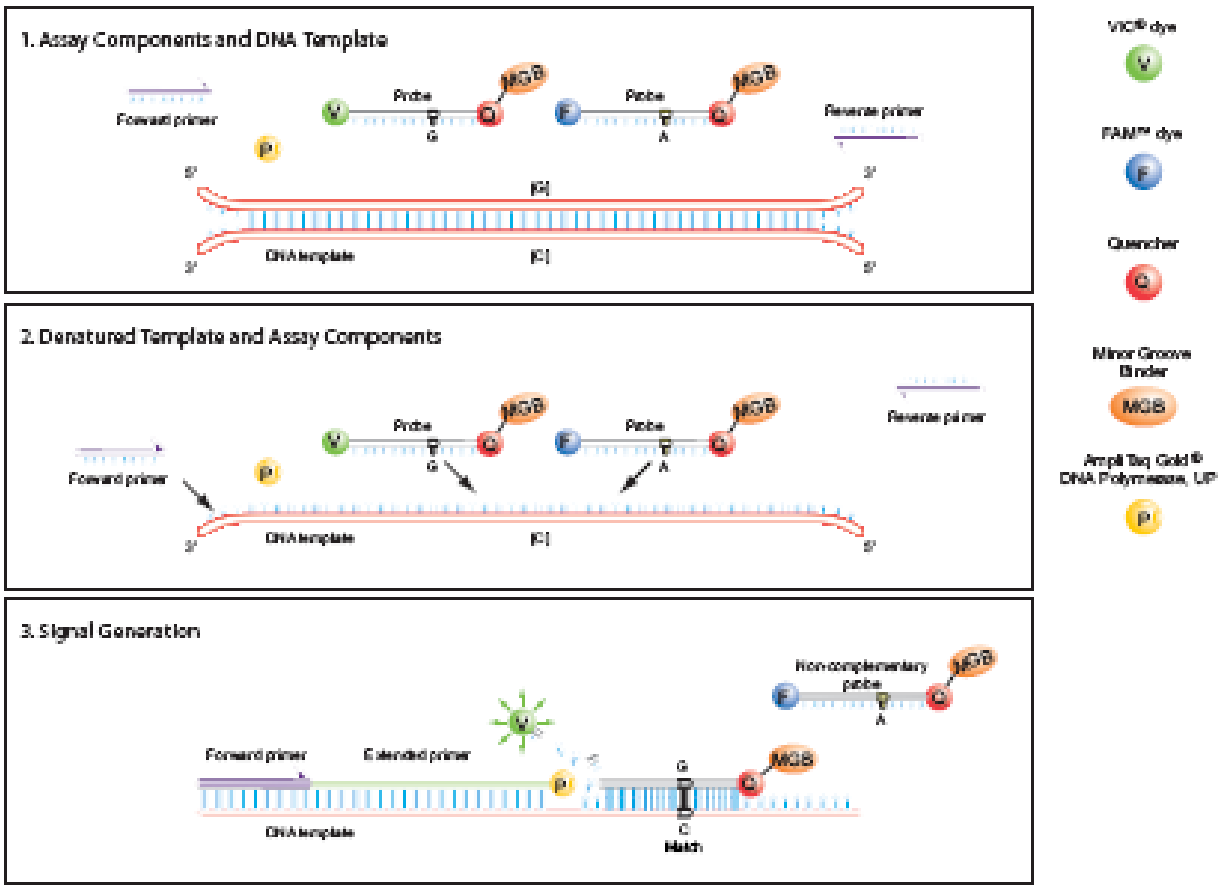


Figure 4. Illustration of the Taqman process for genotyping. Applied Biosystems, 2007.

2.4 STATISTICAL METHODS

Allele frequencies, concordance of genotype distributions with Hardy-Weinberg equilibrium (HWE), and linkage disequilibrium (LD) patterns were determined using the Haploview version 4.1 (http://www.broad.mit.edu/mpg/haploview/). χ^2 test was used to compare the allele frequencies of the SNPs with a MAF $\geq$ 5% between high and low HDL -C groups among the subjects included in sequencing. For those SNPs that were genotyped in the entire population, 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 HDL-C levels. The HDL-C levels were transformed (using either log or square root transformation) to reduce the effects of non-normality. The significant covariates were identified using stepwise regression in both directions. The R statistical software package (version 2.3.1, http://www.r-project.org) and Statistical Analysis Software (SAS) were used to perform all computations. Two genetic models were used for data analysis, the additive and dominant models. A *p*-value of less than 0.05 under one of these models was considered as suggestive evidence of association.

3.0 RESULTS

3.1 DNA RESEQUENCING

Resequencing of the *CD36* gene from 190 NHW chromosomes (94 in high HDL group and 96 in low HDL group) and 190 African black chromosomes (94 in high HDL group and 96 in low HDL group) revealed a total of 343 variants. Of those 343 variants, 46 were insertion or deletions and the remaining 297 were single nucleotide substitutions. Out of 343 variants identified in our study, 1 was located in the promoter region, 305 were located in introns, 18 were in exons, and 21 were in the 3' flanking region. Of the 343 variants identified in this study, only 175 were previously reported in the SeattleSNPs database. Of the total 343 variants, 131 were present in NHWs, 281 were present in Blacks, and 69 were shared by both populations. Of these shared variants, 8 had not been previously reported in the SeattleSNPs database.

3.1.1 Non-Hispanic Whites

A total of 131 variants were identified in the NHWs, of which 78 had been previously reported in the SeattleSNP database. 34 of these variants had a MAF>1%, 52 had a MAF of 1-5%, and 44 had a MAF ≥5%. We identified 119 variants located in introns, 5 variants located in exons, and 7 variants located in the 3' flanking region. 15 of the variants were indels, with the number of bases affected in NHW ranging from 1-29. Two of these indels are located in exons; however

they are located in the 3'UTR and not in coding sequence. Of the other three exonic variants, one was located in exon 5 and resulted in a synonymous change, one was located in exon 11 and resulted in a glycine→valine change, and one was located in exon 12 and resulted in a tyrosine→histidine change. Table 8 is a summary of the *CD36* variants identified in our NHW population sample, and the given locations are based on the DNA reference sequence provided in the SeattleSNPs database (accession number AY095373, total number of exons..13). The refSNP ID number is located in the table for those SNPs previously reported in the SeattleSNPs database, and those variants that were also identified in our black sample are highlighted in yellow.

Table 8. *CD36* variants identified in our study for the NHW population

CD36 Variant	Base Change	refSNP ID	Location	Amino Acid Change	MAF	HWE (p-value)
271_273	del3		Intron 1		0.037	1.000
861	T>C	rs1527463	Intron 2		0.016	1.000
1024	A>G	rs3211809	Intron 2		0.016	1.000
2389	delT	rs3211815	Intron 2		0.047	1.000
2521	G>A	rs3211816	Intron 2		0.389	0.735
2638	T>G	rs3211817	Intron 2		0.048	1.000
2688	A>G		Intron 2		0.005	1.000
2996	C>A	rs3211820	Intron 2		0.405	0.677
3049	G>A		Intron 2		0.011	1.000
3094	G>A	rs3211821	Intron 2		0.453	0.946
3157	G>A	rs3211822	Intron 2		0.404	0.753
3304	C>G	rs3212000	Intron 2		0.043	0.288
3349_3350	insA	rs3211823	Intron 2		0.011	1.000
3691	G>A	rs3211825	Intron 2		0.016	1.000
3991	A>C	rs3211827	Intron 2		0.389	0.735
4108	G>A		Intron 2		0.005	1.000
4134	T>G	rs3211828	Intron 2		0.047	1.000
4249	C>T	rs3211830	Intron 2		0.053	1.000
4366	A>T	rs997906	Intron 2		0.395	0.613
4595	G>A		Intron 2		0.005	1.000
4648	C>A	rs3211834	Intron 2		0.405	0.677
4990	C>T	rs3211838	Intron 2		0.043	1.000
4993	G>A	rs3211839	Intron 2		0.441	1.000
5497	A>G		Intron 2		0.005	1.000
6146	T>C		Intron 2		0.005	1.000
6652	G>T		Intron 2		0.005	1.000
7167	G>A	rs3211842	Intron 2		0.437	1.000
7306	G>A	rs3212001	Intron 2		0.026	1.000
7664	G>A	rs3212002	Intron 2		0.005	1.000
7854	A>G	rs3211849	Intron 2		0.442	1.000
7988	A>C	rs3211851	Intron 2		0.048	1.000
8595	T>A	rs3211855	Intron 2		0.016	1.000
8639	G>A		Intron 2		0.005	1.000
9136	A>T		Intron 2		0.005	1.000
9347	C>G		Intron 2		0.005	1.000

Table 8 Continued

9473	C>T	rs1054516	Intron 2		0.452	0.864
9534	C>T	rs1054517	Intron 2		0.436	1.000
9600	A>G	rs1133344	Intron 2		0.065	0.630
9616	T>C	rs3212003	Intron 2		0.018	1.000
9663	G>A	rs3212004	Intron 2		0.005	1.000
9786	delA		Intron 2		0.036	1.000
9794_9799	del6		Intron 2		0.058	1.000
10279	C>T		Intron 2		0.005	1.000
10381	T>C	rs3173798	Intron 2		0.047	1.000
10876	A>G	rs3211864	Intron 3		0.063	1.000
11249	T>G		Intron 3		0.005	1.000
11440	T>C		Intron 3		0.016	1.000
11472	C>A	rs3211867	Intron 3		0.043	1.000
11554	T>C	rs3211868	Intron 3		0.048	1.000
11618	A>G		Intron 3		0.005	1.000
11684	T>A	rs3211869	Intron 3		0.043	1.000
11741	C>T	rs3211870	Intron 3		0.437	1.000
11890	A>G	rs3211871	Intron 3		0.388	0.812
11904	G>C		Intron 3		0.005	1.000
12132	T>C		Intron 3		0.388	0.812
12143_12144	del2		Intron 3		0.389	0.994
12145	G>A	rs3211873	Intron 3		0.039	1.000
12272_12274	del3	rs3211874	Intron 3		0.379	0.960
12403	C>G	rs3211875	Intron 3		0.039	1.000
12642	G>A		Intron 3		0.016	1.000
12775	A>G	rs3211876	Intron 3		0.026	1.000
12919	G>A	rs1358337	Intron 3		0.437	1.000
12936	A>T		Intron 3		0.005	1.000
13094	T>A	rs3211879	Intron 3		0.068	0.712
13279	C>T		Intron 3		0.006	1.000
13388	A>G	rs3211881	Intron 3		0.062	0.566
13456	A>G		Intron 3		0.011	1.000
13528_13530	del3	rs3211882	Intron 3		0.396	0.787
13577	T>A	rs3211883	Intron 3		0.06	0.544
13936	C>T	rs3211885	Intron 3		0.437	1.000
13973	G>A	rs3211886	Intron 3		0.389	0.735
14299	A>G	rs3173799	Intron 3		0.053	1.000
14455	A>T	rs3173800	Intron 3		0.389	0.735
14510	G>A		Intron 3		0.005	1.000
14903	G>A	rs3211892	Intron 3		0.016	1.000

Table 8 Continued

15075	A>G		Intron 4		0.005	1.000
15833	C>T	rs3212008	Intron 4		0.005	1.000
15932	G>A	rs1924	Intron 4		0.048	1.000
16377	A>G		Intron 4		0.016	1.000
16385	C>T	rs3212009	Intron 4		0.011	1.000
16824	G>T		Intron 4		0.005	1.000
16983	G>A	rs5956	**Exon 5 (Synonymous)**	**Proline-Proline**	0.042	0.285
17274	A>T		Intron 5		0.005	1.000
17282_17823	insT		Intron 5		0.005	1.000
17641	G>A		Intron 5		0.016	1.000
17743	A>G	rs3211905	Intron 5		0.016	1.000
18137	delA		Intron 5		0.016	1.000
18463	C>T	rs3211908	Intron 6		0.032	1.000
18662	T>C	rs3211909	Intron 6		0.016	1.000
18724	A>C		Intron 6		0.016	1.000
18726	A>C		Intron 6		0.022	1.000
18966	G>A	rs3211912	Intron 6		0.016	1.000
19151	A>G	rs3211913	Intron 6		0.01	1.000
19228	A>C	rs3211914	Intron 6		0.042	1.000
19307_19335	del29	rs3211915	Intron 6		0.468	1.000
19678	T>G		Intron 6		0.016	1.000
19810	G>A		Intron 6		0.016	1.000
20630	C>T	rs3212013	Intron 7		0.021	1.000
20758	C>G		Intron 7		0.011	1.000
20759	T>C		Intron 7		0.005	1.000
21084	C>T		Intron 7		0.005	1.000
21110	C>T	rs3211922	Intron 7		0.016	1.000
22296	C>G	rs3211928	Intron 7		0.437	0.990
22749	C>T	rs3211931	Intron 7		0.426	1.000
22830	C>T	rs3212015	Intron 7		0.005	1.000
22907	G>C		Intron 7		0.005	1.000
23463	T>C	rs3211932	Intron 7		0.426	1.000
23533	G>A	rs3173802	Intron 7		0.426	1.000
23689	G>A	rs3173803	Intron 7		0.426	1.000
24071	A>T		Intron 8		0.006	1.000
24104	A>G		Intron 8		0.006	1.000
24331	T>C		Intron 8		0.005	1.000
24426	T>A	rs3173804	Intron 8		0.426	1.000
25575_25576	insA		Intron 9		0.016	1.000
25580	T>C	rs3173805	Intron 9		0.426	1.000

Table 8 Continued

26076	G>A	rs1527483	Intron 10		0.037	1.000
26512	C>A	rs1405747	Intron 10		0.426	1.000
26669	G>T		**Exon 11 (Non-synonymous)**	**Glycine>Valine**	0.005	1.000
26822	A>C		Intron 11		0.005	1.000
27003	C>T	rs3211952	Intron 11		0.426	1.000
27309	T>C		**Exon 12(Non-synonymous)**	**Tyrosine>Histidine**	0.005	1.000
27411	G>T		Intron 12		0.026	1.000
28080_28083	del4		**Exon- 3'UTR**		0.021	1.000
28314_28329	del 16	rs3212018	**Exon- 3'UTR**		0.163	0.487
28375	T>G	rs3211956	3' flanking		0.037	1.000
28412	G>A		3' flanking		0.005	1.000
28572	G>T		3' flanking		0.011	1.000
28685	A>G	rs3211958	3' flanking		0.426	1.000
29112	C>G		3' flanking		0.005	1.000
29225	A>G	rs3211960	3' flanking		0.426	1.000
29894_29902	del 9	rs3044712	3' flanking		0.426	1.000

3.1.2 Blacks

A total of 281 variants were identified in the Black population. Out of these variants, 143 had been previously reported in the SeattleSNPs database. 68 of these variants had a MAF<1%, 110 had a MAF of 1-5%, and 103 had a MAF ≥5%. We identified 1 variant located in the 5' flanking region, 250 variants located in introns, 13 variants located in exons, and 18 variants located in the 3' flanking region. 43 of the variants were indels, with the number of bases affected in Blacks ranging from 1-665. Four of these indels are located in exons; one of them is located in exon 3 and causes a frameshift, one causes a deletion of exon 4, and the other two are located in the 3'UTR and not in coding sequence. The following is a list of the remaining nine exonic variants: tryptophan→stop codon in exon 3, tyrosine→stop codon in exon 5, tyrosine→stop codon in exon 9, cysteine→phenylalanine in exon 9, tyrosine→aspartic acid in exon 10, tyrosine→phenylalanine in exon 10, arginine→tryptophan in exon 11, and two substitutions located in the non-coding 3' UTR region. Table 9 is a summary of the *CD36* variants identified in our Black population sample, and the given locations are based on the DNA reference sequence provided in the SeattleSNPs database (accession number AY095373, total number of exons..13). The refSNP ID number is located in the table for those SNPs previously reported in the SeattleSNPs database, and those variants that were also identified in blacks are highlighted in yellow.

Table 9. *CD36* Variants identified in our study for the Black population

CD36 Variant	Base Change	refSNP ID	Location	Amino Acid Change	MAF	HWE (p-value)
106	G>T	rs3211805	5' flanking		0.063	0.619
400	C>T		Intron 1		0.005	1.000
861	T>C	rs1527463	Intron 2		0.037	1.000
947_948	del2		Intron 2		0.084	1.000
949_950	insA		Intron 2		0.084	1.000
1021	T>C	rs3211808	Intron 2		0.068	0.609
1024	A>G	rs3211809	Intron 2		0.244	0.489
1466	G>T		Intron 2		0.005	1.000
1529	G>A		Intron 2		0.005	1.000
1547	T>G	rs3211810	Intron 2		0.147	0.730
1675	C>A		Intron 2		0.005	1.000
1848	T>C	rs3211811	Intron 2		0.058	0.528
2092	A>G		Intron 2		0.005	1.000
2162	A>G	rs3211812	Intron 2		0.058	0.528
2273	T>G	rs3211813	Intron 2		0.037	1.000
2298	C>A		Intron 2		0.026	1.000
2306	A>G	rs3211814	Intron 2		0.037	1.000
2389	delT	rs3211815	Intron 2		0.205	0.819
2521	G>A	rs3211816	Intron 2		0.100	0.737
2638	T>G	rs3211817	Intron 2		0.105	0.655
2652	A>C	rs3211818	Intron 2		0.084	1.000
2702	T>C		Intron 2		0.005	1.000
2856	A>G		Intron 2		0.005	1.000
2996	C>A	rs3211820	Intron 2		0.416	1.000
3079	C>G		Intron 2		0.032	1.000
3094	A>G	rs3211821	Intron 2		0.295	1.000
3157	G>A	rs3211822	Intron 2		0.400	1.000
3350	delA	rs3211823	Intron 2		0.196	**0.028**
3412	C>T	rs3211824	Intron 2		0.054	0.455
3505	C>G		Intron 2		0.005	1.000
3714_3715	del2	rs3211826	Intron 2		0.168	0.418
3835	delA		Intron 2		0.026	1.000
3852	G>A		Intron 2		0.005	1.000
3991	A>C	rs3211827	Intron 2		0.074	1.000
4039	C>T		Intron 2		0.005	1.000
4046	G>T		Intron 2		0.005	1.000
4133	T>G		Intron 2		0.005	1.000
4134	T>G	rs3211828	Intron 2		0.221	0.050
4249	C>T	rs3211830	Intron 2		0.116	0.506
4259	T>C	rs3211831	Intron 2		0.089	0.912
4266	C>T		Intron 2		0.005	1.000
4308	A>G		Intron 2		0.005	1.000

Table 9 Continued

4366	A>T	rs997906	Intron 2		0.095	0.823
4648	C>A	rs3211834	Intron 2		0.305	1.000
4752	C>A	rs3211835	Intron 2		0.016	1.000
4879	G>C	rs3211836	Intron 2		0.094	0.869
4990	C>T	rs3211838	Intron 2		0.200	0.164
4993	A>G	rs3211839	Intron 2		0.489	0.100
5081	C>T		Intron 2		0.005	1.000
5241	G>C	rs3211840	Intron 2		0.022	1.000
5290	G>A		Intron 2		0.006	1.000
5504	T>A		Intron 2		0.005	1.000
5913	delA		Intron 2		0.016	1.000
5950	C>T		Intron 2		0.005	1.000
6007	C>T	rs3211841	Intron 2		0.016	1.000
6238	A>G		Intron 2		0.005	1.000
6638	G>T		Intron 2		0.032	1.000
6877	G>A		Intron 2		0.005	1.000
7037_7038	insA		Intron 2		0.032	1.000
7046	G>C		Intron 2		0.005	1.000
7167	G>A	rs3211842	Intron 2		0.289	**0.023**
7175	A>G	rs3211843	Intron 2		0.106	0.536
7265	G>T	rs3211844	Intron 2		0.021	1.000
7351	T>A		Intron 2		0.005	1.000
7414	C>T		Intron 2		0.005	1.000
7430	C>A	rs3211845	Intron 2		0.054	0.450
7486	C>A	rs3211846	Intron 2		0.021	1.000
7663	C>T	rs3211848	Intron 2		0.038	1.000
7854	A>G	rs3211849	Intron 2		0.489	0.436
7947	G>A	rs3211850	Intron 2		0.063	1.000
7988	A>C	rs3211851	Intron 2		0.191	0.196
8152	G>A	rs3211852	Intron 2		0.021	1.000
8228	G>A		Intron 2		0.005	1.000
8414	A>T		Intron 2		0.011	1.000
8427	C>A		Intron 2		0.016	1.000
8583	C>T	rs3211854	Intron 2		0.043	1.000
8595	T>A	rs3211855	Intron 2		0.210	0.777
8796	A>G	rs3211856	Intron 2		0.055	1.000
8873	A>G		Intron 2		0.011	1.000
9045	T>A		Intron 2		0.032	1.000
9117	T>G		Intron 2		0.033	1.000
9152	delC	rs3211857	Intron 2		0.065	1.000
9291	G>A		Intron 2		0.005	1.000
9473	T>C	rs1054516	Intron 2		0.277	0.192
9474	G>A	rs3211858	Intron 2		0.056	1.000
9505	T>A	rs3211859	Intron 2		0.054	1.000
9534	C>T	rs1054517	Intron 2		0.261	0.139
9600	G>A	rs1133344	Intron 2		0.394	0.716
9699	T>C		Intron 2		0.021	1.000

Table 9 Continued

9786	delA		Intron 2		0.193	0.229
9794_9799	del6		Intron 2		0.051	1.000
10045	C>T		Intron 2		0.011	1.000
10103	T>C		Intron 2		0.005	1.000
10342	T>C		Intron 2		0.005	1.000
10381	T>C	rs3173798	Intron 2		0.189	0.204
10423_10424	del2	rs3211861	**Exon 3**	**Frameshift**	0.016	1.000
10465	G>A		**Exon 3 (Nonsense)**	**Tryptophan> Stop**	0.005	1.000
10654	C>A		Intron 3		0.005	1.000
10876	A>G	rs3211864	Intron 3		0.037	1.000
10975	G>A		Intron 3		0.005	1.000
11137	delG	rs3211865	Intron 3		0.069	1.000
11155	C>T	rs3211866	Intron 3		0.021	1.000
11472	C>A	rs3211867	Intron 3		0.300	0.143
11554	T>C	rs3211868	Intron 3		0.189	0.204
11640	G>A		Intron 3		0.011	1.000
11684	T>A	rs3211869	Intron 3		0.188	0.230
11741	C>T	rs3211870	Intron 3		0.371	0.601
11890	A>G	rs3211871	Intron 3		0.026	1.000
11943	C>T		Intron 3		0.021	1.000
12124	T>C	rs3211872	Intron 3		0.021	1.000
12125_12126	ins2		Intron 3		0.033	1.000
12132	T>C		Intron 3		0.072	1.000
12143_12144	del2		Intron 3		0.088	0.972
12145	G>A	rs3211873	Intron 3		0.162	0.745
12272_12274	del3	rs3211874	Intron 3		0.090	1.000
12403	C>G	rs3211875	Intron 3		0.101	0.910
12609	G>C		Intron 3		0.011	1.000
12691	A>T		Intron 3		0.016	1.000
12745	A>C		Intron 3		0.005	1.000
12919	G>A	rs1358337	Intron 3		0.463	0.532
12998	C>A	rs3211878	Intron 3		0.058	0.528
13094	T>A	rs3211879	Intron 3		0.058	1.000
13346	T>C	rs3211880	Intron 3		0.034	1.000
13388	A>G	rs3211881	Intron 3		0.056	1.000
13479	C>G		Intron 3		0.027	1.000
13528_13530	del3	rs3211882	Intron 3		0.109	0.628
13548	C>T		Intron 3		0.005	1.000
13577	A>T	rs3211883	Intron 3		0.363	0.284
13758	T>C	rs3211884	Intron 3		0.060	0.544
13936	C>T	rs3211885	Intron 3		0.311	0.082
13973	G>A	rs3211886	Intron 3		0.079	1.000
14075	C>T		Intron 3		0.005	1.000
14076	G>A	rs3211887	Intron 3		0.011	1.000
14081	C>A	rs3211888	Intron 3		0.058	0.528
14190	T>C	rs3211889	Intron 3		0.016	1.000
14299	A>G	rs3173799	Intron 3		0.189	0.204

Table 9 Continued

14455	A>T	rs3173800	Intron 3		0.105	0.655
14482	T>G	rs3211890	Intron 3		0.080	1.000
14500	T>C		Intron 3		0.005	1.000
14701_15292	del592		**Intron 3-Intron 4**	**Del Exon 4**	0.005	1.000
14882	T>C	rs3211891	Intron 3		0.063	0.619
14903	G>A	rs3211892	Intron 3		0.321	0.582
15062	T>C	rs3211893	Intron 4		0.043	1.000
15405	A>G		Intron 4		0.005	1.000
15425	G>A		Intron 4		0.026	1.000
15554_15557	del3		Intron 4		0.063	0.619
15676	A>G		Intron 4		0.016	1.000
15932	G>A	rs1924	Intron 4		0.237	1.000
15975	A>G	rs3211895	Intron 4		0.068	0.712
16040	G>T	rs3211896	Intron 4		0.068	1.000
16051	T>C		Intron 4		0.005	1.000
16163	G>A	rs3211897	Intron 4		0.090	1.000
16295	C>T		Intron 4		0.016	1.000
16565	T>G	rs3211898	Intron 4		0.011	1.000
16568	A>G	rs3211899	Intron 4		0.059	1.000
16741	A>G		Intron 4		0.021	1.000
16986	C>A		**Exon 5 (Nonsense)**	**Tyrosine> Stop**	0.011	1.000
17093	T>A	rs3173801	Intron 5		0.042	1.000
17282	T>A		Intron 5		0.042	1.000
17292	G>A		Intron 5		0.005	1.000
17421_17433	del13	rs3211903	Intron 5		0.047	1.000
17437	T>G		Intron 5		0.005	1.000
17462	delT	rs3211904	Intron 5		0.042	1.000
17743	A>G	rs3211905	Intron 5		0.080	1.000
17807_17813	del7	rs3211906	Intron 5		0.280	0.391
18119	G>A		Intron 5		0.009	1.000
18137	delA		Intron 5		0.352	0.740
18364_18367	del4	rs3211907	Intron 6		0.042	1.000
18463	C>T	rs3211908	Intron 6		0.012	1.000
18596_18597	insT		Intron 6		0.006	1.000
18650_18651	ins3		Intron 6		0.006	1.000
18659	T>C		Intron 6		0.005	1.000
18662	T>C	rs3211909	Intron 6		0.335	0.343
18784	C>T		Intron 6		0.005	1.000
18785	G>A	rs3211910	Intron 6		0.043	1.000
18821_18833	del13		Intron 6		0.005	1.000
18825	G>A		Intron 6		0.005	1.000
18911	T>G	rs3211911	Intron 6		0.084	1.000
18966	G>A	rs3211912	Intron 6		0.011	1.000
19151	G>A	rs3211913	Intron 6		0.494	1.000
19260_19261	ins4		Intron 6		0.016	1.000
19386	T>C	rs3211916	Intron 6		0.289	0.502
19426	C>T		Intron 6		0.005	1.000

Table 9 Continued

19673	A>T		Intron 6		0.016	1.000
19678	T>G		Intron 6		0.333	0.422
19742	A>G		Intron 6		0.005	1.000
19810	G>A		Intron 6		0.389	0.925
19877	T>C	rs3211917	Intron 6		0.053	1.000
20108	T>G		Intron 6		0.011	1.000
20119_20120	ins23		Intron 6		0.016	1.000
20141	T>A		Intron 6		0.005	1.000
20440	A>G	rs3211919	Intron 7		0.032	1.000
20584	A>T		Intron 7		0.026	1.000
20644	T>G	rs3211920	Intron 7		0.089	1.000
20843	T>C		Intron 7		0.005	1.000
21034	G>A		Intron 7		0.011	1.000
21111	G>A	rs3211923	Intron 7		0.011	1.000
21138	G>A		Intron 7		0.021	1.000
21174	T>C		Intron 7		0.011	1.000
21320_21321	ins4		Intron 7		0.005	1.000
21728	C>T		Intron 7		0.007	1.000
21882	G>A		Intron 7		0.027	1.000
21891	G>A	rs3211926	Intron 7		0.011	1.000
21927	C>T	rs3211927	Intron 7		0.011	1.000
21984	T>C		Intron 7		0.026	1.000
22296	C>G	rs3211928	Intron 7		0.174	0.875
22338	A>G	rs3211929	Intron 7		0.011	1.000
22351	A>C	rs3211930	Intron 7		0.079	0.189
22614	T>C		Intron 7		0.011	1.000
22749	C>T	rs3211931	Intron 7		0.170	0.956
23008	G>C		Intron 7		0.005	1.000
23219	G>A		Intron 7		0.005	1.000
23345	G>C		Intron 7		0.005	1.000
23459_23488	del30		Intron 7		0.005	1.000
23463	T>C	rs3211932	Intron 7		0.165	1.000
23533	G>A	rs3173802	Intron 7		0.069	1.000
23689	G>A	rs3173803	Intron 7		0.168	0.973
23780	G>A		Intron 7		0.026	1.000
24071	A>T		Intron 8		0.011	1.000
24080	T>C		Intron 8		0.059	1.000
24165	C>T	rs3211933	Intron 8		0.033	0.163
24166	G>A	rs3211934	Intron 8		0.054	1.000
24186	C>G	rs3211935	Intron 8		0.011	1.000
24426	T>A	rs3173804	Intron 8		0.168	0.973
24427	A>T	rs3211936	Intron 8		0.037	1.000
24496	A>G		Intron 8		0.005	1.000
24545	C>T		Intron 8		0.021	1.000
24657	T>C		Intron 8		0.026	1.000
25025	T>G	rs3211938	**Exon 9 (Nonsense)**	**Tyrosine> Stop**	0.226	0.503

Table 9 Continued

25048	G>T		**Exon 9 (Non-synonymous)**	**Cysteine> Phenylalanine**	0.005	1.000
25276	T>C	rs3211939	Intron 9		0.037	1.000
25305	C>T		Intron 9		0.005	1.000
25550	G>A	rs3211940	Intron 9		0.037	0.219
25580	T>C	rs3173805	Intron 9		0.060	1.000
25695	T>C	rs3211941	Intron 9		0.038	1.000
25849	T>G		**Exon 10 (Non-synonymous)**	**Tyrosine> Aspartic acid**	0.005	1.000
25850	A>T		**Exon 10 (Non-synonymous)**	**Tyrosine> Phenylalanine**	0.016	1.000
25945	C>A	rs3211942	Intron 10		0.063	1.000
26016_26019	del4		Intron 10		0.011	1.000
26076	G>A	rs1527483	Intron 10		0.011	1.000
26089	G>T	rs3211944	Intron 10		0.047	0.360
26158	T>C		Intron 10		0.016	1.000
26160	G>T	rs3211945	Intron 10		0.011	1.000
26262	T>C		Intron 10		0.016	1.000
26305	G>A	rs3211946	Intron 10		0.011	1.000
26346	C>T	rs3211947	Intron 10		0.016	1.000
26446_26449	del4		Intron 10		0.005	1.000
26484	G>A		Intron 10		0.016	1.000
26512	C>A	rs1405747	Intron 10		0.207	0.651
26692	C>T		**Exon 11**	**Arginine> Tryptophan**	0.005	1.000
26740	G>C		Intron 11		0.005	1.000
26776	T>C	rs3211949	Intron 11		0.011	1.000
26927	A>G	rs3211951	Intron 11		0.016	1.000
26948_26984	del37	rs3211950	Intron 11		0.011	1.000
27003	C>T	rs3211952	Intron 11		0.111	0.614
27041	A>G	rs3211953	Intron 11		0.011	1.000
27163	A>T		Intron 11		0.021	1.000
27234_27239	del6		Intron 11		0.005	1.000
27494	A>G		Intron 12		0.005	1.000
28080_28083	del4		**Exon- 3' UTR**		0.016	1.000
28278	G>T		**Exon- 3' UTR**		0.016	1.000
28302	A>G	rs8956	**Exon- 3' UTR**		0.037	1.000
28314_28329	del16	rs3212018	**Exon- 3' UTR**		0.037	1.000
28375	T>G	rs3211956	3' flanking		0.011	1.000
28507_28508	ins4		3' flanking		0.016	1.000
28509	A>T		3' flanking		0.016	1.000
28685	A>G	rs3211958	3' flanking		0.082	1.000
28747	C>G		3' flanking		0.011	1.000
28947	G>A	rs3211959	3' flanking		0.021	1.000
28969_28974	del6		3' flanking		0.016	1.000
28971	A>G		3' flanking		0.005	1.000
29082	G>A		3' flanking		0.043	1.000
29186_29187	ins4		3' flanking		0.021	1.000

Table 9 Continued

29225	A>G	rs3211960	3' flanking		0.059	1.000
29281_29945	del665		3' flanking		0.006	1.000
29344_29347	del4		3' flanking		0.005	1.000
29706	C>A		3' flanking		0.011	1.000
29815	G>A		3' flanking		0.011	1.000
29894_29902	del9	rs3044712	3' flanking		0.159	1.000
29963	G>A		3' flanking		0.016	1.000
30040	C>G		3' flanking		0.017	1.000

3.1.3 *CD36* Annotated Sequence

Figure 5 depicts the variants identified in *CD36* within a color FASTA representation of the annotated reference sequence adapted from the SeattleSNPs database (http://pga.mbt.washington.edu/) and modified by including additional variants identified in our study. The variants identified in this study as well as in the SeattleSNPs database are shown in **black font** with the nucleotide change, refSNP ID number, and SeattleSNP database location on the right side of the sequence. Variants reported in the SeattleSNPs database but not identified in our study are depicted in **green font**, and those variants identified only in this study are shown in **red font,** of which there are 183. One possible reason for these variants that are unique to either the SeattleSNPs database or our study could be that the SeattleSNPs sample was unselected with regards to HDL-C levels. The insertions identified only in this study are indicated by the flanking bases being highlighted in yellow. Large deletions found only in our study were highlighted in light blue. The suspicious variants with low sequence quality are marked with an *. The color code used in SeattleSNPs for the reference sequence is as follows: light grey for flanking regions and introns, green for UTR, dark blue for exons, purple for repeat regions, and light blue for regions not covered for SeattleSNPs database. Among the two large gaps present in the SeattleSNPs database, only one was also present in our study (21350-21672). It is important to note that our intron/exon numbering differs from the SeattleSNPs numbering after exon 3 due to the fact that the SeattleSNPs database separated exon 3 into two separate exons (exon 3 and exon 4). It also seems that SeattleSNPs did not highlight an alternative exon, and if we took this alternative exon into consideration then all of our exons and introns would be numbered one higher (exon 3 would be exon 4, etc.). In general, the literature has inconsistent numbering of introns and exons for *CD36*, and the numbering scheme used would be based on the reference

sequence used. The RNA form that seems to be must representative of the exon structure in SeattleSNPs is NM_001127443.1. This form includes all 12 coding exons and the non-coding first exon highlighted in SeattleSNPs.

```
ATGTCTTGCT GTTGATTTGT GAATAAGGTA TCGTAAATAA AACATCTGTT 50
ACCATACTTG CTTATCATTT AATGGAAAAC ACATCAGTCA ACCCACATTC 100
TGTTCGCAGG AGAGCTCCAG AAGGGGTGTG GAAGGTTGTG TTGGGTGGAG 150 | [G/T]_rs3211805 (106)
AAACCAGATA GTGAGGATGC AACTAAGTTG CTGAGACAAG GGAAGAGAGA 200 | Exon 1 | UTR
TGAGGGTGAG AGTTCTCCTT AGATAAGATT TCAATATGTT AATCATGTGT 250
AGAAAGAAAA TTAAAAAGGA GGAATATGAA GAAATTCAGA TATGACATTA 300 [GGA/-] (271_273)
TTAGTTCTGC CACTGGTAGG CATTAGAAGC AAGAAAAGGG AGACGGACCG 350
AGGAAGCCAC TTTGGTGAAA CAAAAAGAAA AGCATTTGTT TATTTAGAAC 400 [C/T] (400)
GGGCAAAATG ATACGTTTCA GTGGGTGTTT TCTTTGTACT TTGATCTTTT 450
TGTACTGATA TTTAAGCTTC TGTTTTATGA TCTCTTTCTA ATGATAGAAC 500 | Exon 2 | UTR
CAGAGCTTGT AGAAACCACT TTAATCATAT CCAGGAGTTT GCAAGAAACA 550
GGTGCTTAAC ACTAATTCAC CTCCTGAACA AGAAAAATGG GCTGTGACCG 600
                                      M   G  C  D  R 5
GAACTGTGGG CTCATCGCTG GGGCTGTCAT TGGTGCTGTC CTGGCTGTGT 650
  N  C  G   L  I  A   G  A  V  I   G  A  V   L  A  V   21
TTGGAGGTAT TCTAATGCCA GTTGGAGACC TGCTTATCCA GAAGACAATT 700
F  G  G  I   L  M  P   V  G  D   L  L  I  Q   K  T  I  38
AAAAAGGTAC AAGTAGTCCA AAGAATATGC CTTCTCATTT TGATTGATTC 750
 K  K                                                  40
TAACTTCTCT TTTTTTGCTT TGTATTTACC TGCTTTATAT TTCATGGTAA 800 | [C/A]_rs3211806 (757)
CTGCTAATTT TGTATCTTTG ACATAAAGGT AATTATGAAC CACTGCAACT 850
CTATATGATG TGACTTTATG TGAAATGTTA TAAGTATAAT GTATATTTAA 900 | [T/C]_rs1527463 (861)
CATGACTCCA TTGCTGTCTT AAATATAAAT ACCAAATTCT ATTAAAAGCT 950 | [AG/-] (947_948) | [-/A] (949_950)
GTCTACAGGT ATGCATGTTA GTAGAAATAA TTGTTTTAAG TTATGTCCAA 1000
AGAGCATGTT GGCATGCTTT TGAATAGGAA ATAAGTGAGT ATATTTTGTA 1050 | [T/C]_rs3211808(1021)|
[A/G]_rs3211809 (1024)
AAAGCACATT TATAAAAGAA GTTGCACTTT AGTTAATACT GAGAAAAGTA 1100
AAACTGTGTG TGTGTGTGTG TGTGTGTGTA ATGTGTTTAA TATTGAAACA 1150 |
TAAATCCTTA TTAAATTGTA GGTAAACTTG TTTGGTAATA CACTGTTTAG 1200
TAATCCACTA TTTTTATATA TGTGTAATAA TCTCATCTCA TAAATATTTT 1250
CTATTTGTGA AGCTTCATAT TGGAATCTTA GAAAATACTT TCAGAAATAT 1300 | [G/A]_rs3211999 (1272)
GCAGAACATG TCTTAGTATA AAACAAATTG ACTGTAGTGT GAAAAAACAG 1350
AATGATTGAA TAGATGGGCT TTGCACAACA ACCTAGAATT CATCTCCCAC 1400
CCTAGCTTAT TCGAATTAGC TACACACTCA CTCATCAGAA ACGTTGATTG 1450
ATAGGGGGAA GAAGAGAATA AAGGGAAGCA GTTCTGCTGA AGTTCTAAAT 1500 | [G/T] 1466
CAGGGATGGC AAATTCAAAT GGCTGCAGGA GTTTGGATAC AGGGGTTAAA 1550 | [G/A] 1529 | [T/G]_rs3211810 (1547)
CAAAATTTAA GGTGCCAAGA ACAAGATAAA ATGAAGAACA GGAACTGCTG 1600
TGAACTAGAC TGTGTGTTTA CTGCCTGAAG GCATAAATGT TCATTTTATT 1650
AAACATAATA CTGGCCAAAT AAAACGAGTT CTGCCTTCAA CTCTCTACCT 1700 | [C/A] (1675)
GGTTAGGATG GCAATGACCT AGACAAAGGG TTAGTAAGCA TAGTGCACGA 1750
TGGAGAACAG GAGGAATTGA ATTTTTATTA AATTTTTATT GCTATATTGT 1800
TAGTATTTTT AATATTTTTG ATCCACAGTT GGTTGAATCT GCAGATGTGG 1850 | REPEAT | [T/C]_rs3211811 (1848)
AACCCATGAA TACAGAGGGC TGACTGTATT GTGTTTATTG CTAACATATT 1900
GCAGAACTTC AGCCTCATGC ATTTAACTGA AAAAACATAT GTAAATTAGT 1950
ATCCCCTGCC TTGGAATTCT TAGTTTTCTT GATAGTATTT TTAATACAGG 2000
CCAATGGAAA CAGACAGGTA ATAGTGAGGT GTGGGGCTTA TCTAATAGTG 2050
GGATGGAGTG GCCATTGATC TGACATCCTT CCTACTCATA AACTATGTTC 2100 | [A/G] (2092)
TCACCCAGAT CTTATGCAGA GAAAGTACAA GATCAGTGTC TGTGTTAACG 2150
TACAGACTAC AACATCATTT GGAAAAGTTT TCCAAATTCA AAAATCACAA 2200 | [A/G]_rs3211812 (2162)
TAATCTTCCA ATGCACAGCG ATTTAGACTA CTTTTTTTTG CTGACACATA 2250
ATTATTGGTC AATAACTGAG TATTGATGAT TTAATTTTTT TCTTTGTCAT 2300 | [T/G]_rs3211813 (2273) | [C/A]
(2298)
GCAACAAAAC TGGTACATGT AGATTCTTTT GAATAGCATG TGAGGTGCTA 2350 | [A/G]_rs3211814 (2306)
GCATTTATTT TAATTCTTTA ATATTTATCT TTATGCTTTA TAATAGTTAA 2400 | [T/-]_rs3211815 (2389)
CCATGAGAAA GTAAGTTTTC TGACATCAAA ATGTGCTTTT GTATAATGAC 2450
TAAGAGAATA ATATAATCTC ATCTATTGTA AGCTATAACC AGGGGAAGAT 2500
ATATTATAAT AAAAAATACC GAGACCTATG AGACTCTAGA ATTGAATTGG 2550 | [G/A]_rs3211816 (2521)
AAAAGTAAAT GCTGTAATAC TTTGAAAGAG AAATCTCTCA GAGTATTCAA 2600
GAAACTTCAG GAAAAAGGTA GGACTTGATT CAGATTTTAA AGAAGTGTAG 2650 | [T/G]_rs3211817 (2638)
AATTTTGAAA GTCTCAAATA ACTGCAGTAA CCATATTAAA GGACTGATTG 2700 | [A/C]_rs3211818 (2652) | [A/G]
(2688)
ATAGTATGAA AAACCCTGTG AAAATGCTAC CAAATGCAGT GAAGATGGAA 2750 | [T/C] (2702)
GGAAGAAGGT AGAGAATGTG CCCATAATAC TGGATTAACA GCAGAAAAGA 2800
TTAGAAGTTA ACATGGGGTT AGAGTTAGGA AGCAAGCTTA GAACATCAGG 2850
CAGAAAAATA CTTAATTGAT GATACAATAC ACAATAGTCG TCGGTATGTT 2900 | [A/G] (2856) | [G/A]_rs3211819
(2890)
TTTCAGCAAG TGTAAGATAT GATAAAACCT GAATTTAGAG ATTAGATGGC 2950
TAACATTTAA TTCTTACTAA GAGTTTCCTA TATACTAGTT CACTTCATCT 3000 | REPEAT | [C/A]_rs3211820 (2996)
```

```
TCAAAAATAT CTGTGAGAAG TTGCTGTTAT TTTACTCATC TTACAGGCGC 3050 | [G/A] (3049)
AAGTCCTTAA ATCAACTTGC CAACTCACCT AGCCAATAGG GAAAGAAGTC 3100 | [C/G] (3079) | [A/G]_rs3211821
(3094)
ATTCTCCAAA TGTAAGCAAT CTGTAGGAAG AGCCTAGATG AAAAACCTGG 3150
CCTTGTGTAC TACTCTACAG CAGAAAGGTT GCAAAGTGAG GATACCAGAG 3200 | [G/A]_rs3211822 (3157)
ATAGGAAAAC AAACAAACAA AAAACTGTAG GATATGACCA CTTTCAAGGA 3250
GAATATAAGA TTCCCTTTTA GAAAGATATT TGCATAATCT GCACATTATA 3300
TAACAAAAGG CACACTGGTC TGGCATGGTT GAGTATATAG AAAAAAAAAA 3350 | [C/G]_rs3212000(3304) | [-
/A]_rs3211823 (3350)
CTGATAGATA ATAGGATGAA ATGGAAGCTT GGAGAGACAG ATTAAATTAG 3400
AGGAGTAAAA GCATCAAACT CAAGATGTCC AGTGAGTTAT TGATAGAAGG 3450 | [C/T]_rs3211824 (3412)
TATGTATATA CCTTACAAAG TAAAACAGAG GTTTTATAAG TGATGTTTTA 3500
TGAACATCAG TCTGTGTTTC AAAGGTCAGA GTTCAAATAG ATTTGATCAT 3550 | [C/G] (3505)
TGCTATAAAG ATTAAAAATA AAACTAAACA GTAGATTTAC TTTTAGAGAG 3600
TTTTAATGGA GAGAGGAAAA AGCAGAAGAA TTATTACAGA ATTTGTAATA 3650
ATTCAGAGGC AGGAAGACAA AAGTAGAAAA GGAAAAGAAT GGATCTTATAT 3700 | [G/A]_rs3211825 (3691)
GAAGAAACAA ACAAAAGAGTA CAATCCTAGA AGTTAATAGG AGGCCAGAAC 3750 | [AA/-]_rs3211826 (3714)
TCAAGGAAAG AAAGCATAAA ATGTTCTCAG GCAGATGAAA ATTAAGAAAA 3800
TAATGATTAG AACATGGTCA CCAATAATTT TGAAAGTATA CTTTGAGTAG 3850 | [A/-] (3835)
GGTAATAAAG AAGGGGGAAT TAAATCAATG GAGTGCTTTT CCTTCTTGTT 3900 | [G/A] (3852)
TTTGAAATAA GAATTATCCG TGCAATCTTT TACAAGGAAA AGAGAAGATT 3950
CAGACTTCCT AGTGTAGATT CAAACTCAAG GAGGTGGTAC AAACAGAGACT 4000 | [A/C]_rs3211827 (3991)
CATAAAAACT GTCAAATGAC AGTGCACATT GTAAGCCCCC TACATGTCCA 4050 | [C/T] (4039) | [G/T] (3046)
TAGACTAAGA CTCATCATCA GAATCCTTAT ATTGGTCAAG ATGTATGGTA 4100
CATTTAAGGC TAAGCTTCTT TTTGCTCTCA GATTTCCTACA CAACCATGTT 4150 | [G/A] (4108) | [T/G] (4133) |
[T/G]_rs3211828 (4134)
GACCTTGTTA ATTTAGTCAT ATGAAGTAGG TGGAATATTG TTTGAATAGA 4200 | [G/A]_rs3211829 (4180)
CTGTTTCTTT TTTCTTCAGT TTACAGACTA TTTTTTAAAT GATATTGGCG 4250 | [C/T]_rs3211830 (4249)
TATGTATATC TTTGGCATAG AGTAGGAGTC AAATATTTGA ATTTTGTGTA 4300 | [T/C]_rs3211831 (4259) | [C/T]
(4266)
AACCTAGAGA GCCAAGTGTC TTACTGCTGC CTCAGAAATA CTTAGAGTAA 4350 | [A/G] (4308)
TGCTATTTCA CAAACATATG GGTTTTGGTA CTGTTGACTT CTTATTTGTG 4400 | [A/T]_rs997906 (4366)
TACCATTAAA ACTCATTTAA TATTACTTTG TTTAGCTGAC TAATAGCAAA 4450 | [T/A]_rs3211833 (4423)
TTAAGAAAGC TATTGTATAC AAGTGATATT TGGAAAAAAA TAATTAAACT 4500
TCTTAATTGA GAATGATGAG AAAATCATTT CATATTTAGT GTAGAGTTGC 4550
AGTGCTGTGA TACCTTCCAA TTCTGAAAGA TGTCATTACC ATTTGAAGTA 4600 | [G/A] (4595)
GTCTAAAACA GTTTAGCAAG CTTATTTGTA CTTTATTTTT GCTCTTTCTT 4650 | [C/A]_rs3211834 (4648)
TTCAAGGGAT GTCTCTGGTA TCCTCATGCT TTTTATGGAT TATAGCTGCA 4700
ATCTTTCTTA CCAGTATTTT TGACCGTGAT TCTTAATAGA GTTGTGTGCA 4750
GCCAGCAAGT TTATGCAGTC TTTCAAAAAA AAAAAAAAAA CTAAAAATCA 4800 | [C/A]_rs3211835 (4752)
ATAGCAAGAG CTGCCATGAA AGTAGAAACT ACCAAAGCAC TGTGGAAAAG 4850
GGTCAAAATT CTACCACCAT TTACAAAAGA TAATTTATCA CTTGACATTA 4900 | [G/C]_rs3211836 (4879)
TTAATGTCAG TTGATAAAAA GTTTAGTGAA TATTGTATAA CATAACATAG 4950 | [A/G]_rs3211837 (4949)
GTAGTTGTAT GCTAAGCAGA TTAATGCAGG AACAGAAAAC CAAATACCAC 5000 | REPEAT | [C/T]_rs3211838 (4990) |
[A/G]_rs3211839 (4993)
ATGTTCTCAC TTGTAAGTAG GAGCTAAATA TTGAATACAC GTGGATATAA 5050
AGAAGGGAAC AATAGACACC TGGGGCTAAT CGAGGGTGGG AAGGGGGAGG 5100 | [C/T] (5081)
AGAGCGGAGG ATCAAAAAAC TACCTATCAG GTACTATGCT TATTACCTGC 5150
GTGAAGAAAT AATCCGTACA CCAAACCTCC ATGACACACA ATTTACCTAT 5200
ATAACAAACT TGCACATGTA CCTCTGACCC TAAAATAAAA GTTGGAAGAA 5250 | [G/C]_rs3211840 (5241)
AAAAATAAAA TTTAAATAAG TTCTATAGTT CTAACTATAG AACACCAAAA 5300 | [G/A] (5290)
AAGGAATTAT AGGTTTCTTT GTTCCTGTTA GAGAACTAAT ATTAAATATC 5350
ATCATTCAAT TATACCATTT TTCCTTCTCC CATTCTTGAA AACAATGATC 5400
CTTCTGTAGC TTGTAAGAAA TGCCTTCAGA CAAAATAAAT TGAAATCATG 5450
ATTCATTAAA ATGGTTATAC ATGGCTGGGC ATGGTGGCTC ATGCCTATAA 5500 | REPEAT [A/G] (5497)
TCCTAGTACT TCGGGAGGCT GAGACGGGTG GATCACTTGA GCTCAGGAGT 5550 | [T/A] (5504)
TTGAGACCAG CCTGGGCAAC ATGGCAAAAC CCCATCTCTA CGAATAATAC 5600
AAAAAAATTA GCTGGATATG GTTACACATG CCTGTAGTCC TAGCTTTTGT 5650
GGGGCTGAAG TGGGAGGATC CCTTGAGCCC AGGAGGTTGA GGCTGCAGTG 5700
AGCTAAGGTC GCGCCTCTGT ACTTCAGCCT GGGTGACAAA GTGAGACCCT 5750
GTCTCCCCCC GCAAAAAATT ATGTTTGAAT TAATAGTCCA TATTTTAAAC 5800
ATATAATTTG TTTATATTAC CTTTATAATG TATGCTTCTA GTATTTCCAG 5850
ATATAATAAA ATAATTTGTA CTTTTTCATA GTATCTAAAA GCTCAAAAAT 5900
AAATGACTAA AAATATTAAA GATGGTAAAG TTTGTTTTAT ATTTAATAAC 5950 | [A/-] (5913)* | [C/T] (5950)
GTAAGAACAA CCCAAAATAT TTAAATTAAA ATAGCAGTAT AAAAACGAAT 6000
TAAACACTAT GTATATGTGC ATATAGCTGT AATAACAAAT GTCACTTAAA 6050 | [C/T]_rs3211841 (6007)
GACAATGTTT CAAGAAATAG ACAAGCATAT ATGTTTCTTC TACTAGTCCA 6100
ATTTAATATC AACATAGTTT ACTTTTCATG CCAAACACAA CATAGTTTAC 6150 | [T/C] (6146)
TTTTCATACC ATTTTAGGTA CTTACATAAA TTCTTGGAGG CTAATTTCTA 6200
GGTTTCAGAA ACATATTAAA GCTTACTACT CCAATACATT TCACCTAATT 6250 | [A/G] (6238)
TTATTCTGGG TTTTAGTTGA TGATGATGTT CTGAAACTAA AGCCATTTTC 6300
```

```
AGATTCTCTG CATAACTGTC ACTACAATTC TAATTCATTC TCAACTTACA 6350
CACACTTGCA TAATGCAAAA ACATTAGAAT TTCTTTTAAC ATGGGGAAAT 6400
GATTCGTTTT AGGCCCAATT ACAAGCAAAT GGTAGCAGCT AGTGGGAAGG 6450
ATGGAAGTAT GCTTCTGGTT TTTAAGAAAG TTTTTACTTT ACTAGAAAGA 6500
GAAATATTAA AATGATTAAA TAACAGAAAA ATAATTGTCA CAGGATATTA 6550
TATAAGTTAT TGGCACATGA CTGGCTTAGA CAGTAAATGC TATGAACCAA 6600 |
GCAATTAAAG CAGAATGGAG AAGTAAATGG TATTAGAGTT CTTCCTCTAA 6650 | [G/T] (6638)
AGGATGGATA GAATTTGTAA AATACTAATA GCTGACATTT ATTGAATGTC 6700 | REPEAT [G/T] (6652)
TGGTTTGTAC TAACACTGTT CTAAGTGTTC ACGTTTATTG ATCCCCATAA 6750
CAACCTATGA GGATAAATAT TTAAGCCACA TGAAAATGCC AAAATTCAAC 6800 | REPEAT
TTTTTTTGAT TTTCAAAAAT ACTAAGTCTA TATGGTTCAA TCTAACATTA 6850 | REPEAT
CTGTTATAAT TTCACAGATA AGGAAAGTGA GGTACAAAGA AATTAAATAA 6900 | [G/A] (6877)
TCTGCTCAAA CTTCAATGAT GATGATTATT ATTAAGTGAT CAAACTCATA 6950
GCATCTAATT TCATTGTTGG CCACTTCAGT CTATGCTTTA GGGAAGAAGC 7000
CACTATACCT TGTTAGGGTG GGGATAGGGT AAAAAAATCT AACTCGTAGA 7050 | [-/A] (7037_7038) | [G/C] (7046)
TTAAAATCTG CTAAACATAC TTCTGTCTCT TGGTAGAAAT GTGTAAAGTC 7100
CTTGAAAAGT GAAACAAAAT ATTTTCTACA TTTGTACTCT ACTCCATGCG 7150
TAGCAAAGTA TGAAGGGCAT TTTAATAATG AAAGTTGTCA GTTATGTAGA 7200 | [G/A]_rs3211842 (7167)|
[A/G]_rs3211843 (7175)
TGATTATTGC TGTTGCATGA AAACAGTATT CAAGTGAGTG CATGACTACA 7250
TATAGGATAT TAAGGTCCCA ACAAGCATCT GTTTTCAAAT TAATGTAGCA 7300 | [G/T]_rs3211844 (7265)
GAATAGAAGT TTCAAGGTCA GTGCTTTCAA CCCAGCTATA GAACATGCTA 7350 | [G/A]_rs3212001 (7306)
TTGTGAGAGTC ACCAGGTTGA AATCCAGTTG TTGTTGGGGC TTGGTAGTCT 7400 | [T/A] (7351)
GCATGCTTAG AGTCAGGGAC AAATACACAC ATACACATTA GCCAAGTGAG 7450 | [C/T] (7414)| [C/A]_rs3211845
(7430)
AAAATCTTAT ATCACTTTTT CTGTGTACAT ACTTACCAGC TTTGTACTTC 7500 | [C/A]_rs3211846 (7486)
AAATTTAGTA AAATATTTGC AAAATAGTCT AGACAGTCCA AATTCAATTT 7550
AAAATCAAGT TCTTTTAAAT TCACAAAGGA CATAACATCT TTTGTTTCTT 7600
GTGACATATT TAGTCTTGGT ATTCAAGGAG TGACTACAAT TTAATACATA 7650
TAGTAGTGGT AGCGCTCAGAA GCATATTATT ACTACTTCGA CTACTTCAAT 7700 | [G/C]_rs3211847 (7662) |
[C/T]_rs3211848 (7663) | [G/A]_rs3212002 (7664)
ATTTTTCTCT ATGTCTCTAA ACAAATCATG ATGCTATGGA AACTATATAA 7750
AATTGAAGTG AAAGGACCTG ATTTTTCTAT TATCTGCTAG TAAATTTGTC 7800 | REPEAT
ACATAACATT GAGTCAGACA TCTCTTTTCT CTAGAACCCA TTTTCCACCT 7850
CTAGAAAAAT GGACGGGGTT AGATAGTCCT GATTCTAACA GTTGATAATT 7900 | [G/A]_rs3211849 (7854)
CAACACATTC ATTTCAACCC TGTGTTTCAC GTTTGCCTTA TTTGTCGCTA 7950 | [G/A]_rs3211850 (7947)
GAGATATTTT TTAAGGGAAA AGTGATTTGT GGCCTTTAGA AAAAAAGGTC 8000 | [A/C]_rs3211851 (7988)
TCATACACGT ACAATGTCTT TTGGTGTTTA GATGGACACT CATTGTAGAG 8050
TTAGGAAGGC ATTCAATCCT GTCAAGCACA TTCTGTCACA GACTCTCTTC 8100
AATTCTATAA GAGTTTGTTG AGTTTCTTTC ATTAAGAGAT ATTAGGTCGG 8150 | REPEAT
TGCAAAAGTA ATTGTGATTT TTGTCATTAA AATCGCAAAA ACCGTAATTA 8200 | [G/A]_rs3211852 (8152)
CTTGTGTACA AACCTAATAT GTACCTTGAG AGGCACTTGA TGAATAACAG 8250 | [G/A] (8228)
GTATTGTTTT TGAGCATTAA GATTCTAATA GTGTGAGGAG AATAATGTTT 8300
GTGTTTGTGT GTTCTCCCTT GTTAAGTCTC TTGAAATGGA GAAATCAACC 8350
GTTTTCAGCT TCTATTTAGC TTTGAGCTTC AACATCCTCA CCAAACATAT 8400
TTACATTTAG AAGAATATCC ATAATGCTTT AAGCAATTTT GTCTTAATAT 8450 | [A/T] (8414)* | [G/T]_rs3211853
(8426) | [C/A] (8427)*
TAGACTTAGT ATGGAAATAA GCCTTTTTGG GAGCAAATAA AAAGTTTTTC 8500
AAGGACAATT TGTTGTAAAA AGTAATAAAG ATTCCATCAC TCTAGCAGGT 8550
CCAATGTACC CCGTATAAGT GGATTACAGC AGCGTAGGAT TTTGTAATTC 8600 | [C/T]_rs3211854 (8583) |
[T/A]_rs3211855 (8595)
CATTTTAACT TGGTGCCCAG TTCTTCTCTA AGGAATTCGC TTATTCCCTA 8650 | [G/A] (8639)*
CATTAAGCAA TAAGTAAAAT ATCATGAAGA TATGAAACTT AATTTGACTT 8700
CTACCAACTA AGTGCAGAGG GTGACGGAAA AAAAAAAAAA AAAGACAGTC 8750 |
ATAGCATGGT GTAAAGGAGG ACTAGTCTAA GGGTCATGGA TGAGGATTTT 8800 | REPEAT | [A/G]_rs3211856 (8796)
TGTCATGAAT TCACTACTGT TTGCCTAGGC TGTCCTTGGG CCAGCCACCT 8850
GGTTTCTCAG AGCCTCAGTG TGATCCACTG GAAAATGTGA TCATGTTCAA 8900 | [A/G] (8873)
TGTCCTTTCC TAATCTGTCA TTCTGTGAAG TGGTGCTTCC AGGTCAAGGA 8950
AACAAATAGG TCAGAATGTT CCTGTGGTCA TTGTTCTCTT TAGGCTCATT 9000
TTCCATGAAA TAGTTGTATT CCTCTTCCTG TAGCACACTC AGGTTAAGTT 9050 | [T/A] (9045)
GGAGCTCTTT AACCTGGGCA TTCTTCCCAA TACTTCATTA GAAATACCGC 9100
CTGCTTCAAA ATCACCTTTG TTTGCACTAT CTACTACAGC GTAATAGACT 9150 | [T/G] (9117)| [A/T] (9136)
TCTGGCTCAG AGAGTGAAAG CTGTGGAAGA AAACTAGAAG AATATTTATG 9200 | [C/-]_rs3211857 (9152) | REPEAT
TTAACTCCCT TATTTTACAG CTAATGAAGT TGACCCTAGG GAGGTTTGAT 9250
GACTCTCCAC CCAGTCCCAA AGCTCCTAAA GCAAGGAATC GAGTTCCTAA 9300 | [G/A] (9291)*
TGTAGGATTT TTAATTCTCT TATTGGTGTA TTTTCCACTT GTCTATCTTT 9350 | [C/G] (9347)
ACCAAAGGAG CATCAGTGAT TTTTAGTGGA TTTCACAAAG GAGTAATAGG 9400
ACAGCTTCCT TGTCCTGTTT TTTTCTTTAA AAACTACAAC CCAATAATGA 9450
TGAGTGAGGC TGTTGTAATT TTTGGAAACA TGATAATGGG TTGTTGCAAA 9500 | [T/C]_rs1054516 (9473) |
[G/A]_rs3211858 (9474)
TATATTGAAA GTTAGGGGCT GTGTGCAGTG GCTCACACCT GTAATTCCAG 9550 | [T/A]_rs3211859 (9505) | REPEAT |
```

```
[C/T]_rs1054517 (9534)
CCTTTTGGGA GTCCAGGTCA GGTGGATTGC TTCAGCTCAG GAGTTCAAGA 9600 | [A/G]_rs1133344 (9600)
TCAGCCTGAC CAATATTGGTG AAATCCTGTT TCTACTAAAA CTACAAAAAC 9650 | [T/C]_rs3212003 (9616)
TAGGTGGGCA TGGTGGCACA GGTGAGAGGA TCACTTGAGC CTGGGAGGTC 9700 | [G/A]_rs3212004 (9663) | [T/C]
(9699)
AAGGCTGCAA GGAGCCAAGA TTGTACCACT GTACTCCAGC CTGGGTGACA 9750
GAGCAATATC TTGTCTCAAA AAAAAAAAAA AGAAAAGAAA AAGAAAAAAG 9800 | [A/-] (9786) | [AAAAAG/-]
(9794_9799)
TTAGGGAAGC CAGATTGAGC AGGATTTTGA AGACACGGTG TCAAATTTGT 9850
GTTCTTCATG TAGATATTTT TATTTTCTTA ATGTAGTGAG TGTTAAGTAC 9900
CATCCTCTGA AGTGTCCAAC TCTTGAAGGA AAAGAAAAAG CTTCCAATAC 9950
CATTAGATTT TTCACTCAGT TTTTGTGTTT TGTTTCATCA GTCCAACTTA 10000
GGGGCAGAGA AGAAGTTCAA CTCAGTAAGA ACTTTTTTGA ACTTCGGTAA 10050 | [C/T] (10045)
AATTTGCTTA ATAAATTATG TTGACTGTTG CAATATTTTC AAAGCGTCAC 10100
TCTAAAGCTT GCCGAAGGGT CACTTTAAAG TTTGCCTTAA AATCAACAGT 10150 | [T/C] (10103)
CGTGTCTTCA GTATTACACA CTGATTCTCT TTGTAAAAGG CTAAAAAGAC 10200 |
TGCTGTAATA ATAATTTGTT GAAAACATTT CTGCTGCAAG TGTATGGTAA 10250
GGTTGCAAAG GTTCTCATGA ATGAGGTACT TGGGCTTGGT CCTTTTATTC 10300 | [C/T] (10279)
TGGCTGACTC AAGGCTGCAA ACAATCTTCC AGAAGTGCCT GTACTTACTA 10350 | [T/C] (10342)
CAAAGACATA ACCCAAACTT ATTTTCTTTT TCATAGCAAG TTGTCCTCGA 10400 | [T/C]_rs3173798 (10381) | Exon 3
                                      Q   V  V  L  E 45
AGAAGGTACA ATTGCTTTTA AAAAGTTGGG TTAAAACAGG CACAGAAGTT 10450 | [AA/-]_rs3211861(10423_10424) |
[G/-]_rs3211862 (10425) |
  E  G  T   I  A  F   K  N                              W      V     K     T     G
     T     E     V   61
TACAGACAGT TTTGGATCTT TGATGTGCAA AATCCACAGG AAGTGATGAT 10500 | [G/A] (10465)
 Y  R  Q   F  W  I  F   D  V  Q   N  P  Q   E  V  M  M 78
GAACAGCAGC AACATTCAAG TTAAGCAAAG AGGTCCTTAT ACGTACAGGT 10550
  N  S  S   N  I  Q   V  K  Q  R   G  P  Y   T  Y  R   94
GAGTGAGTCC CCACAAATAT GAGACACTCT TACCTTGACC ATGTATTTCT 10600
GAGAAGTCTT CTACTTGGCA AATGTCATTG TATTGAAATG TACTTATTAT 10650 | [A/G]_rs3211863 (10613)
TTTCTTGCCA AAAATATACT TTAAAATATT TTTCCTGTCT GTATAGAATC 10700 | [C/A] (10654)
CTAATCTAAG AATTTAATGA TTATAAGGTA TTTATTTTGA AAAAAGTGGA 10750
AGATATATAC ATTATCCAAA TATTTATTAG ACAATATATA GAAAAGATAA 10800
ACAGATTTTA TTATAAACTT ACCAGTATCA AATAACTTCC CTATGTTTAT 10850
GAAGTTAGTG TTTTGTTATT GTGGCATATA TTCTCAACAA CTCTGAAATA 10900 | [A/G]_rs3211864 (10876)
TTCCTGCTGA GGAAAAAAAA ATCAGTTTTC ACATCTTAAA ATTTAAAGCA 10950
TATTTTAACA GTGGTTCTCA AAGTGTAGCT GGGCCACAGC AAAGTGCTGG 11000 | REPEAT | [G/A] (10975)
GCCAGCTGCA TCAGCATTAA CAGGGAATTG GGAAACATGC AAATCCCTCC 11050
ACCCCATCCC AGACCAAATG AATGAAAGAG TTGAGGCCCA CTAAAGAGTT 11100
TTACAAATCC TGGGTGATTT GAATACACAA TTTTGAGAAC CACCGTGCTA 11150 | [G/-]_rs3211865 (11137)
AATTCATTTT TCCTAACTGC GGTACCTACT CAAGATTTAT ATTCAAATTA 11200 | [C/T]_rs3211866 (11155)
ACTTCTGGGT CATATAGATG TTTACATTGA TCTATTTTAG TTTCACTCTT 11250 | [T/G] (11249)
ATATTTCCTC TAGTGTTGAG ACAATATTTT ATATCTTTCC TCTGCCTTCC 11300
CTGATAGTTT AACACACTTT TGAACAGAAT CTTAGAGTAT TAGAACTAGA 11350 | REPEAT
GGAAATGATC AGCCAGAATC TGTCATTTTG TGAGTTAAAA ATGCCTGATC 11400
CAGTGGATGT GAGGTCACTT GTCCATGTTC ACACAGATGT AGGTACTGGT 11450 | [T/C] (11440)
AGCTCCTAGA TTTTTGTATA CCTTTTCTTC CATTTCAATT ATAGACACTC 11500 | [C/A]_rs3211867 (11472)
TGTCTTTTTC ACTAAAAAGG GAAAAAAGGG GGAGATTGAG GTGCCAAAAA 11550
CACTACCTTA ACCAAGGATT AGGAATATAT TTATGGTATC CATATGGAAA 11600 | [T/C]_rs3211868 (11554)
CATAGTTATG TTCTGCTATA GAATTTGTTA CCTAAATGTG TATTTTCCCT 11650 | [A/G] (11618) | [G/A] (11640)
CATGGATAAT CACAACTAGT TATACACCAT GATTAGTTCT CATCACAAAG 11700 | [T/A]_rs3211869 (11684)
TTTTGTATTT CTAATGATTT TTTATAATTT GTAGTTTTAG CATCTTATCT 11750 | [C/T]_rs3211870 (11741)
GTTTTGGTTA TCGGATGATT GTATAATTTT CCTTCTTCCC CAAAGCATTA 11800
CTATGCTTCC AGATTGAAAA TTCACACCCA AAATAAAATG AAATAACAGG 11850
CATGGTGAGG GTGGTGCCTA CTACCATAGC ACATGACATA TAAGCCCCAG 11900 | [A/G]_rs3211871 (11890)
AGGGCCTTGG TGCATGTCCT GTCTTAGCAG TTTTACCCGA ATCGGGTGTC 11950 | [G/C] (11904) | [C/T] (11943)
TCTGTGCGCA TCTGTGACAA TTCATAATTT TGTCAATGAA GACTTAGTTT 12000
TTATGGGTAC AAAAACATCT TACCTTAGAA TTTTCTTCAT GGTCTTAAAG 12050
TTTTTATATT ATTAGCAAGA GTCTATACAA GTATCTTGTT GCAAAGAATG 12100
TTGTTAGGTG TACCATATAT ATATACACAC ATTATATATAT ATATGTATTC 12150 | [T/C]_rs3211872 (12124) | [-/TA]
(12125_12126) |[T/C] (12132) | [AT/-] (12143_12144)* | [G/A]_rs3211873 (12145)
ATAGGTGTAT ATATACTTAT GAAAATCTAG AATTCACATA TGACCTAAGG 12200
AGAGTATCAG CATAATTGTT TAAGAAATAG AACCTAAGTG AACTGAGTAA 12250
GTGTGAGGAA AGAAAATAGT TGTTAGAATT TGAGAACTCA TCATGATTTT 12300 | [GTT/-]_rs3211874 (12272_12274)
TTTTGTTTGT TTGGGGGAAG TTATATTTTT CACTTCTGTT TTTAATTTTG 12350 | REPEAT
TTTATTTTTT GAGGCAAGGT CTTGTTCTGT CACCCACTCT AGAGTGCAGT 12400
GGCACGATCA CAGTTCACTG CATCCTCAAC CTCCTAGGCT CAAGTGATCC 12450 | [C/G]_rs3211875 (12403)
TCCCACCTCA GCCTCCTGAG TAGCTGGGAC CACAGGCGCA TGCCACTATA 12500
CCCAGCTAAT TTTTATGGGT TTTGTTTTGT AGACAAGGTT TCGCCATGTT 12550
GCTCATGCTG GCCTCAAACT CCTGGGCTCA AGCAATCTAT GTGCCTTGGC 12600
```


CTCGCAAAGT GTTGGGATTA CAGGTGTGAC CCACCATGCC CGACCATATT 12650 | [G/C] (12609) | [G/A] (12642)
TTTCACTTTT AAATATATTA TTTGGGATAT GAGAAAACAT ATTATTGAAA 12700 | [A/T] (12691)
TAGTTCTCCT CTGGTAGTTG ACAGCTGATC TATATTTTAT TCTTAAAGCC 12750 | [A/C] (12745)
TTTAACCTAT CTCTCTAAGG GACAAAAAGG GGATGTGTTA GATATATTTG 12800 | [A/G]_rs3211876 (12775)
GTTTGCTTAG AGAGGCAATG GAGCAAAGAC TGGAGAAGAA CAGTTGCATT 12850
TACAGCTAAA TTTTAGAATC CATAAATAGC TCTTTCTGTA ACAATTTTAA 12900
AGGGATTCCT GATTATCTAT AGGTCATCTT GTCTCAGCAT GTCACCAAAA 12950 | [A/G]_rs1358337 (12919) | [A/T] (12936)
TAGTCTTTTA TTGTTTGCCT GAGTGCCTTA AATAATGGAA AAACAACCAG 13000 | [C/A]_rs3211878 (12998)
TACCTTTTAG AAAAAAAATT AACACTTTGA TAGTGCATGT GTTGAGCTAA 13050
ACATGCTTTT TCATAACTAA TTATACCCTA AATCCATCTG ACATTGGAAG 13100 | [T/A]_rs3211879 (13094)
TATGTGAGAA TGTCCCTCCT CAAACAGAAC CACAGGCTGT ATTTGGCCAT 13150 | [T/G](13140)*
TGTCTGCTAA AGTAAGCTTG ATTACACTTT GACAAGATAT GACCTGAATC 13200
AAAGCACGAA ATTGCTTGGG TTGAGATCTT ATGTGTGTTT CTACCTACAC 13250
CATCTCTATG AAGATGCTGT AATTCCTTCA GAATATTCTT AGATCTGTAA 13300 | [C/T] (13279)
TTCATTTTTT GAGTTACTTT GCTTTGGAAA GAATATAAAA AGTAATCTTG 13350 | [T/C]_rs3211880 (13346)
AGAAAAGCAT AAATTCCTTC CCTTGACCTT ATAAACAAAC ACGTTCCAAA 13400 | [A/G]_rs3211881 (13388)
ACTTCCATAC AGTCTAACTT TTAAACAAAA AATTCAGTAA TTATGTTCCC 13450
TTTTTAAAAC AACATTCTAA AGGCAGGGCT GTTTTAAGGA TTTTTCTTCT 13500 | [C/G] (13479)
AGTGACTGCC ATCTACTGGT ATAATGTTGT CAGTACACTT TAGGTAACGT 13550 | [TGT/-]_rs3211882 (13528) | [C/T] (13548)
GACAATTATG TTGCTATTGT TTTAATTTAT GGGCATATTT CTGAAAATTG 13600 | [T/A]_rs3211883 (13577)
AAGTACTTTG ATTTGAAATA TTAGTAATAT GCACACTAAT TTAAAATACA 13650
ATTGTCCTGA GTGCTGTTTC AGATTTTTAT ATTCTTTGTG TGTCTGCAGT 13700
ACTTCAAGAA TAGTAGTTTT AATAAAATTA TATTAAAAAT ATATTTCTAG 13750
CCAACTTTGA ATCCTCTTTT TAAAGAATCT GAGACATTTG TTTAGTAAAT 13800 | [T/C]_rs3211884 (13758) | REPEAT | [A/G] (13799)
ACATTTACTG AGAACTCATT ATAATACCAGG CATGATGCTT AGAGATACAA 13850 | [A/G]_rs3212005 (13823)
AACACTCATA AACCCCTACC CTGTAGGAAT TTACATTTTT ATTATAATGG 13900
AATAAAATGG GTTTTTAAAA AATTCAGTCC AATTGCTTAA GTCTTTACCT 13950 | [C/T]_rs3211885 (13936)
TGGCTATCTA CTCACTTCAC AGGCATTCAA ATATGACAGT GCAATTATGA 14000 | [G/A]_rs3211886 (13973)
TATATATCTT ATTTCTTCTC AGTGAATATC TTCCACTTGT GAAATGTATT 14050 | [T/C]_rs3212006 (14003)
TTTCCTGTAA TAATCACCCT TCACCGCGTT CCTAAAATGT AGTTGGTACT 14100 | [C/T] (14075) | [G/A]_rs3211887 (14076) | [C/A]_rs3211888 (14081)
TTGCAGCCTA GATATGTTAG AGTCAGGAGA TTAGAACTAG CAAAAGTATT 14150
CAATATACTT GAGGGGCTGG GGAAGTTAGT CTATAACCCT TAAAAATGGA 14200 | [T/C]_rs3211889 (14190)
AATTATATTC TTATCTCTAA TTCTCATTAA AGAAGTCCTG AAAAAGGGAA 14250
AAAGCAGTGC TGGATAGGCC TTGTATTTTG ATCATTCTTT CTAATATAAG 14300 | [A/G]_rs3173799 (14299)
AAAATGTGCT CTAAATATTC CCCTGCTGAA GCCCCTGTAC TTTTTCCTCC 14350
TTATCCTCAG GATAAAGTCT CGTTTGAGTA TTAGATGCAA CTCTTTTATG 14400
TTTGAGATCT TACTTGTTTC AATAACTTTA TTTCCACTAC TCTCCTCACC 14450
TCTTACTCTG TAATCCAACC ACACTCAAAT ATCTGCATTT CCCAAAACAT 14500 | [A/T]_rs3173800 (14455) | [T/G]_rs3211890 (14482) | [T/C] (14500)
ATTATGTTCG TTCTTATGCA TGCTACCATC TGCCGTACTT TACCTAACTA 14550 | [G/A] (14510)
AACTTGTATT TGTTACATAG ACCATTTATG TAGACTTGTA CTAAAAAGTG 14600
TTTTAACCCT CACCCCTACC CTGACCATTC TAAATCTGGA TCTGGATCTG 14650
TTTTGTGTTT TCCCATGGAG TCTCCATTAT AACAATATTC CTTCTCTCTC 14700
ACTATTTTAA TTTCCAGTTT GGCAGTTTAA GACCCGTTCT CGTTAGTTTG 14750 | [592bp/-] 14071_15292
CTAGAGACCC TGGCTGATTT CTCATTTTAA CACCATTGCT TACATACTAT 14800 | REPEAT
CTATCTGGCA TATTCTGTGT GTTCAAAAAT CATTTGTTGA ATGAATGACA 14850
TTTGAGATGT AATGTTGACA TATGACAAAT GTTTTGAATT TTGTTTACTG 14900 | [T/C]_rs3211891 (14882)
CTGTTCTTTT AGAGTTCGTT TTCTAGCCAA GGAAAATGTA ACCCAGGACG 14950 | [G/A]_rs3211892 (14903) | Exon 4
V R F L A K E N V T Q D 106
CTGAGGACAA CACAGTCTCT TTCCTGCAGC CCAATGGTGC CATCTTCGAA 15000
A E D N T V S F L Q P N G A I F E 123
CCTTCACTAT CAGTTGGAAC AGAGGCTGAC AACTTCACAG TTCTCAATCT 15050
P S L S V G T E A D N F T V L N L 140
GGCTGTGGCA CTGAGTAGAC AAAGAAGAAA GTTATCTATT TTAAAATAGT 15100 | [T/C]_rs3211893 (15062) | [A/G] (15075)
A V A 143
CTAGAACTCA TGTAATTAAT CCTATCATTA GAATTATTAG GTTTTACTTG 15150
TTTTTCCATT TTTATCAAAA CGTATTCCTA TATCATTATT TTGAATGGAG 15200
GCATGGTCAT TTGGAAAGTG ACAAAATCAT AAGAATGGAC TGCATTAGAA 15250
GACGACTTTA TTTTTTTCAT TATTTTTTGT CGAAACATTC TGTATATTTA 15300
ATTTCAATTT TATAACAGAT TACAGGAAGA TGCTTAAGAA ACAAGTACAA 15350
CATTTGTTTC AGTATGTCTT TAAATGAAAG ACTTTTAAGT ATGTAAGCAA 15400
CTATATAATA AAAGGTTTCC AAACGCAGCC TGTAAGAAAT CAGGCAAATT 15450 | [A/G] (15405) | [G/A] (15425)
TTACTATAAG CAATAAACCA TTCCGAGCTT TCCAGACAGT GTACCAGTAG 15500
CTGTACCAGT GGGCAATAAC CTTTAACCAA ACAAAAACAA ACAAACAAAC 15550 | [C/T]_rs3212007 (15521)
AAACAAAAGC ACTTTGCAAT TTGTTGCTGC AAAATGGGGA GAAAAAAAGA 15600 | [CCA/-] (15554_15557)
GTATATAAAC TTGATGGAAT CACAACTGTC AATATAATTT AAGGGAAAAT 15650

```
AAAGTCGATA AGGTTGATGG TGTCTATTGT TTGGAAAGTC GAATTCGGCT 15700 | [A/G] (15676)
ATTTGCTTGG GGCTCTAGAG ACCACACCAC TGAATAAACA AAACTCTGCA 15750
GAGTCTAGCT ATCCGCCAAC AGGGGGTGCC GTCCAAATTC ATGGCAAATA 15800
AAGGGCATTT GGTGCTCACC ATTCACAACA GGCGGGCATT TATGTGGATG 15850 | [C/T]_rs3212008 (15833)
AAGTACAATT TCTTCAGCAA GCTCAGCAAA AACTCTTAAG GGGCAAATAT 15900
GAACTCTGCA TTTTAAGAAA AATAGAAAAC GGAAACACAA AATCCTAAAA 15950 | [G/A]_rs1924 (15932)
AGTATATGAA GGCTCTGCAT TAGCACACTG GCATCAAACC ACACATCCAC 16000 | [A/G]_rs3211895 (15975)
AGCACATCCT AATTCTATGG GAAACAATCC TTGCTTATCG GAGGTGCATA 16050 | [G/T]_rs3211896 (16040)
TCTAAATTCA ATAGCTTACC AATATATTCT TGGGAGTAGG CCAAAAGGAA 16100 | [T/C] (16051)
ACAGAAAACC CACATTAAAA AAAGAAGTTT CTTTTCCTAA ACATTTTCCT 16150
GAAGCTGAAA TTGAAGTGGA GAGGAAGTCA GTTGTCCTCG TCGAAATCGT 16200 | [G/A]_rs3211897 (16163)
AGTCCTCCTC ATCCTCCCCA ACCTGGGACA CCGGGGTCTT CACCCTGGAG 16250
ATGCTGTACT GAGACCTGTT GGAGCTTGTG GCCAGCATTT CATCCGCACC 16300 | [C/T] (16295)
ATTGGTCAGG TCACTGGCAG AGAGCCTCGT GCCGTTAGAC GTGGAACCTG 16350
CCGTTGTGAT GAACACGCCT GCAACATTG TCTGCGCCAT TTCTGTCACG 16400 | | [A/G] (16377) [C/T]_rs3212009
(16385)
TGTGGCTCCA GCGCCTTTGG GACCAGACTT ATGGCTTTTT TTTTTTTAA 16450 | REPEAT
GTTCTGGGAT ACATGTGCTG AATGTGCAGG ATTGTTACAT AGGTATACAT 16500
GTGCCATGGT GCTTTGCTGC ACCTATCAAA CCATCATCTA GGTTTTAAGC 16550
CCCGTATGCC TTAATGCATT AGATATTTGA GAAGACCACT TTAAGTGTAT 16600 | [T/G]_rs3211898 (16565)|
[A/G]_rs3211899 (16568)
CGTTAAATAA ATAGAGCTTA ACTTGGAATG TCGTCTTCTT GTGGCTGGCA 16650
CTGAGGCAAA GAAATGTAAT CATCTAGGAA TTAGACGAAT TGCATTTTGA 16700
GTTTTGGCAG GATCTGGCAG TAATTTTAAA GATAAGCTTT AAAAAGTTTT 16750 | [A/G] (16741)
GTATTAAGCT CAATATTAGC ATTTAATCCA TTTATTTGTT AAAATCTAAT 16800
ATTGTATTCT TGTCTTAAAC AGTGACTTTG TTTTTGTAGG CTGCATCCCA 16850 | [C/T]_rs3211900 (16809) | [G/T]
(16824)| [A/G]_rs3211901 (16838) | Exon 5
                                                A  A  S  H 147
TATCTATCAA AATCAATTTG TTCAAATGAT CCTCAATTCA CTTATTAACA 16900
  I  Y  Q   N  Q  F   V  Q  M  I   L  N  S   L  I  N   163
AGTCAAAATC TTCTATGTTC CAAGTCAGAA CTTTGAGAGA ACTGTTATGG 16950
K  S  K  S   S  M  F   Q  V  R   T  L  R  E   L  L  W  180
GGCTATAGGG ATCCATTTTT GAGTTTGGTT CCGTACCCTG TTACTACCAC 17000 | [G/A]_rs5956 (16983) | [C/T]
(16986)
 G  Y  R   D  P  F  L   S  L  V   P  Y  P   V  T  T  T 197
AGTTGGTCTG TTTTATCCTG TAAGTACCAA ATATGAATGG CAATATTATT 17050
  V  G  L   F  Y  P                                   203
ACATTTTAAT TTAATTAATT CAATGGCATT GGCAAGGCAT AATTTTATAA 17100 | [T/A]_rs3173801 (17093)
TTTAGCTCAT TAGTCTTATT GCTGATCTGG AGACATATAT CCTAACTTTT 17150
TAAAAAGTCC ACTTCTCATT ATAGCTTCAG CTTTCCTAGT TGGGAAATTC 17200
ATCTGAATTT AACAATTAAA TTTAAACCTG AAGAATAGAT TTAATAAGGT 17250
TTCTACTCAT TTATAAATAC ACAATTTTTT TTAAATTAGC CGGTAAGCTA 17300 | [A/T] (17274) | [-/T]
(17282_17283)| [T/A]_rs3211902 (17282) | [G/A] (17292)
GTCTGTAAAT CTTTGAGCAC TGTTTTTGGC TTTTATTTCC CTATTCACAT 17350
AATCAAGTTT AATACCATAT TTTTATTTGT TTTAAATATA CTCCTCATTC 17400
CTCCTTTTCT AGACCTCTTA CAATTTTAAT TTATATTTAA AAGTTAGCCT 17450 | [TTACAATTTTAAT/-]_rs3211903
(17421_17433) | [T/C] (17437)
AATGTTCACA TTCTCAATAC TGATAAGGTA ATAGACTTCA TTTTAATGGG 17500 | [T/-]_rs3211904 (17462)
ATTTGTAAAT AAGAATTTTT AGTAGTCCAT AATGTCATGA AATGGCAGCT 17550
TGAAGATTAA GGAAAATGTG AACTTGATGG TGTACTTGAT TACCGCTTAA 17600
TGTTTTGAAT TACAAATAGG ATAAGCTAAC TACTGAATTG GAAGTTGGAC 17650 | [G/A] (17641)
TATGACTTCA TTTGGCACTA TATGTGAATA TTATGTTCTC TAGTTCATTG 17700
TTTTACTTTT AGATACTGTT AGGATTACAA GGTTATATAT CAATTATAAA 17750 | [A/G]_rs3211905 (17743)
TGAATGTAGA AAGCCATAAT GAATCAAATT CATTCTGATT TTAACTCAAT 17800 | [TACTCAT/-]_rs3211906
(17807_17813)
ACTCATTACT CATAGCTCTT TCTTTGGCTA ATGCTTTAAC TTTTGGATGT 17850
CTAATTTTTA TCATTTTAGT AACCACTTAT TATCATTTTA GTAACCACTT 17900
ATTAATGACT GTAACATTTA GAATACCCCT AGAAATCATT GTTCTTATAG 17950
GTTTGTTCAA TGTCCCTCAC CTCAACATAG TAAGAATAGT GATCAAAATG 18000
CCCTCATTGG CTTAATTATG ACAGAGTGCT AGAGTTCACA TCATGTCAGC 18050
TTCTGATATG TATCTTCTTT GTCACAGCAT CTAGCACTTA TTTCTAGGCA 18100
CCTTTCACAA TTTTTAAGGC CAATAATTTA AAAAAAATGT ATTGCAGATG 18150 | [G/A] (18119)| [A/-] (18137)
TATTTCAAGT CATTTGAGTA ACCAGTGATT GAGAAATGTG AAAGTGAGTT 18200
ATGTATTGTA CAACTTTGAA AAAATGACTT GTAGAAGTAA CATTTTCCCA 18250
TACATATATT TCAGTACAAC AATACTGCAG ATGGAGTTTA TAAAGTTTTC 18300 | Exon 6
               Y  N   N  T  A   D  G  V  Y   K  V  F  215
AATGGAAAAG ATAACATAAG TAAAGTTGCC ATAATCGACA CATATAAAGG 18350
 N  G  K   D  N  I  S   K  V  A   I  I  D   T  Y  K  G 232
TAAAAGGTAA GTAAGTATTC TGGTAAAATG TGCATGTATG TTACTAGGGT 18400 | [AGTA/-]_rs3211907 (18364_18367)
  K  R                                              234
ACTCTTAAGC AGGAATAGTA TTCATTTAAC ATCTCATAAG ACATAGGCAT 18450
```

```
CAACCTATAG AACAGACCTG GTTATAATTC AGCTCTGGAA ACTCCTGTTC 18500 | [A/G]_rs3212011 (18457)|
[C/T]_rs3211908 (18463)
TGCTAGGTAT TAACTCTTTA GTTGTGGTAA CTGGTGAGTT CACACCAGTG 18550
CATAGCTGCT GACTATCAGC TCCACTTTAA GGTTTGGTTC ACCTTTCTGC 18600 | [-/T] (18596)
ACAGGTTATG GTTGTGTTAC ATAAATCCCC AAAGGGACTA TTTTTTCATC 18650 | [-/ATC] (18650_18651)
TCTGCTACTT ATTCCAGCATT ACAGTATAAT TATTCTTACA ATTAGATAAC 18700 | [T/C] (18659)| [T/C]_rs3211909
(18662)
CATAAATGAA AAGGTAAAAA AAAAAAAAAC AACACATCAA CTGATTGTGT 18750 | [A/C] (18724) | [A/G] (18726)
AGTAGATGGA AACTTTTTTT TTACTTTTTA AATCGAGCAT ATCGAATTCC 18800 | [C/T] (18784)| [G/A]_rs3211910
(18785)
ATATTCCAGT GGCATGACCT AAATGTGTCT ATAAAGATGG AAGCTTAATG 18850 | [AAATGTGTCTATA/-] (18821_18833)|
[G/A] (18825)
AATCCAGGCA ACTGCTTTCA TGACCTTCCC CCTGCAAATA GTCTTTAATA 18900
ATTTTCCATA TTTGATAACTC AGCTTTTTTA ACTTTATCAA TGCAAAAATA 18950 | [T/G]_rs3211911 (18911)
GAATGAATAT TTCAAGTGCA GTTCTACAAT GTAAATACAA AATGTGAAAA 19000 | [G/A]_rs3211912 (18966)
TGAAGACTTT GCCAACTTTA AAGTGGTAAA ATAACAAATC AGCTTCCTAA 19050
GCCATTATTT CCTTTTTTTT TTTCTAGCTC CAGCCTATTC ACCTAAAGAA 19100
TTTATAATTT ATCATATATG TAAACTAGGA AGAACCCTAA TAAATATCAG 19150
AGGAACAAGC TTTCTTCTCC ATAAAATTAA CAATTGTGTG TTTGACTAGT 19200 | [A/G]_rs3211913 (19151)
TTTTTTCTGA AAAAGAATAA AGTGATCAAA ACCAAAAGAG ATGAAATGTT 19250 | [A/C]_rs3211914 (19228)
TTTATTAATA TAAAGTAACC CCTACCTCAA ACTGAGTCTA TCTAATGATT 19300 | [-/AATA] (19260_19261)|
[TCAAACTGAGTCTATCTAATGATTGCTTT/-]_rs3211915 (19307_19335)
GCTTTTCAAA CTGAGTCTAT CTAATGATTG CTTTTAACAG CTAAATATTA 19350
CTAGTGGAGT ACTTTTTCTT CTAAAGAGTC CACAGTTACA TATTTTTATA 19400 | [T/C]_rs3211916 (19386)
GAAAAAGTCA GTAGAGGGAA AAAAACACTT CTAAGTATTC ACTTAAAAGG 19450 | [C/T] (19426)
AAATCACAGC AATTTTTTAT ATTGAGAAAT AACGAGCATT TCACTCTAAT 19500
ATTACAGAGA GATGTGGAGG GGAGTTGCAA AGCACTCCTA GTTAGAGTAA 19550
GAATTTCACA TCATTTTAAG ATTGTAAGGT TGATTTAACT CATGGCAAGA 19600 | NOT SCANNED
CTGAACATGA TTAACCACTT ATTTTGTTTA AGCAATGGCT TGTCTTGACA 19650
TGTCACTGTC AGGAGTGGAA AAAAAGTTTT TCCATCACCA AAAAGTGAGG 19700 | [A/T] (19673)| [T/G] (19678)
ATGCAGGGAA ACTATTACTA TTGTTTTTAT CACTTCCAGA TATATATAGC 19750 | [A/G] (19742)
TTATTTAGAC AGCAAAGTAT AAGTTAGTAA ACTCTTTCCT ACTTTAACGC 19800
CAGTCCATGG GAGAATTAAA GGGAGGAAGG GGCAAGAATA TAAATTAATC 19850 | [G/A] (19810)
ATCTACATGT TTTGAATGTT TGCTGATCCA AACATCTGCT TCTTTCCTCT 19900 | [T/C]_rs3211917 (19877)
TCCCTGCTTC TCTCTTCTGC CTGCCTTGTA CTTAGCTTAT ATCAGTTACC 19950
CACTTAACAC TTTCCTTCTA TTCATTCTGA ACACTTATAC TACTGAGTAA 20000
TTCATGTATC CCATTGGAAA AAAAAAAATT TCCCCTAAGT AGAGATCTAA 20050 | [A/-]_rs3211918 (20028)
GTTTCTGATA AATGTTTTTA ATTTTTTTCT TATTCCAAGA TGTTTCTAAT 20100
TAACACCTGG CAGTTTTTAT TTTATGATCT GGCTACCTAA TGGCATCAGG 20150 | [T/G] (20108)| [-
/TAATTAACACCTGGCAGTTTTTA] (20119_20120) | [T/A] (20141)
TACATTGCAA TAAGATAAAA GGTTCAAACA AAACATAAAC AGAATTGAAC 20200
ATTTCTTAAA CTTAGTACTT GTCACATTTA AATGCATCAT ATTAACAGAA 20250
GTATTGAATT ATAATAGAAA AAGTAATGTA AGAAAGGTAT TCTTTAAATA 20300 | [T/C]_rs3212012 (20277)
AGAATGTTTA TTCATTGTCT TTTTCTATTC CTAGGAATCT GTCCTATTGG 20350 | Exon 7
                                    N  L   S  Y  W  239
GAAAGTCACT GCGACATGAT TAATGGTACA GGTAAGAATA TTTGTTTTGT 20400
 E  S  H   C  D  M  I   N  G  T                        249
GGTCATCACA GTTAATCCAC CTCCCTTTCC CACAAATCCA CCGTTGTACT 20450 | [A/G]_rs3211919 (20440)
GACAGTGTTC TGAAAGTTGA GGGTGTGTGT TTACTTGCCT TTATATCCCC 20500
ACAACAAAAT TCAGAGTCAC TATTTGTATA TAGGCTAAAG GTCTGTTCAA 20550
TGTGATGGGA TTTGTAAAGA TTAATCATAA AAAATTAGCC CTTACTGCAT 20600 | [A/T] (20584)
GATTCAGAGA AATGTTTGCT TTGTAAAAAC TTTACTGCCA TCCTGGCAAC 20650 | [C/T]_rs3212013 (20630) |
[T/G]_rs3211920 (20644)
AGAAGTAAAT GGTAAAAAAC AAACAACAAC ACAAAACACT TTAGTTACAT 20700
TTCAAATATT TTATAAATAG TATACAGATA AGTTAGAATT GAAAGAATTA 20750
AAAAAAGCTA ATTACAAAAT AGAAAACATA TCTAAGCAAG GCTTATAGCT 20800 | [C/G] (20758) | [T/C] (20759)
GAGATAGAAT AAAGTTGCAA GACAGAACCA TCCATTTTTG GATAAGGTGG 20850 | [T/C] (20843)
TCAGAATGAG AGAGAAAACA ATAACAAAGT CAGATGTCTT CTTAACTGTG 20900
AAAAATATAT GTAAAAAATG AGCATTCATA AACTTAGATT TAAGGTATGG 20950
TTAGATCAGG CTTATTCTCC TTCAGGAAAA GCAAAGTACA GATATGCCTA 21000
TTAGTTTAAG AATTAAGATT TTTAAAAAGC AATGGGAGGC TGGTCACAGT 21050 | [T/A]_rs3211921 (21005) | [G/A]
(21034)| REPEAT
GGCTCATGTC TATAATCCCA GCACTTTGGG AGACCAAGGC AGGTAGACTA 21100 | [C/T] (21084)
CTTGTGCCCC GGAGGTTGAG ACCAGCCTGG GCAACATGGC AAAATCCCAT 21150 | [C/T]_rs3211922 (21110) |
[G/A]_rs3211923 (21111) | [G/A] (21138)
CTCTACAAAA AAAATACAAA TATTAGCTGG GTGTGCTGGC ACACGCTTTT 21200 | [T/C] (21174)
AGTTCCGGCT ACTTGGGAGG CTGAGGCAGG AGGATCACTT GATTCCAGGA 21250
GGTGGAGGCT GCAGTGAGCT ATAATTGTGC CTCTGCACTC TAGCCTGGGC 21300
AATAGTGTGA GACCCTGTCT CAAAAAAAAA AAAAAAAAAA AAAGTGGTGG 21350 | [-/GTCT] (21320_21321)
GAATCGAATA CCATGAAGTA ATTAATATCC AGAGGCAAGC TTTCAACTAC 21400 | NOT SCANNED
TAGGAAACAA AAGCATGAAG TTTAGAAAAT TAGGGAATAT CCCATAGGAA 21450
```

```
AATGATATAA CAAAGATAAA GACACTTAAG TAAAAATGAA GTACAGAGAA 21500
TGTTTACTTT GCCAAAGGTA TGATTATTGA CTTCGTATGC GTACTTGATT 21550
CTATCATTGC CTTGTAGGAA ATGATAAACA AAAAAAGCAA ACAAAAAAAA 21600
CAGCAAAAAC ACTGTTGGAT TTGCAGGGGT TTTCTTTTGT TTTTTTTTTT 21650 | REPEAT
TTTTTTTTTT GAGACATAGT CTGGCTCTGT TGCAGGCTGG AGTGCAGTGG 21700
TGCGATCTCG GCTCACTGCA ACCTCTGCCT GCCTGGGGTT CAAGTGATTC 21750 | [C/T]_rs3211924 (21703) | [C/T]
(21728)| [GCCT/-]_rs3211925 (21731)
TCCTACCTCA ATCTCATGAG TAGTTGGGAC TACAGGCATG TGCCACCACA 21800
CCCAGCTAAT TTTTGTATTT TTAGTAGAGA CGGTTTCACC ATGTTGGCCA 21850
GGATGGTCTT GATCTCTTGA CCTTGTGATC CGCCCCCTC GGCCTCCCAA 21900 | [G/A] (21882)| [G/A]_rs3211926
(21891)
AGTGCTGGGA TTACAGGCTT GAGCCACCAC ACCTGGCCAA GGCTTTGCAG 21950 | [C/T]_rs3211927 (21927)
GGTTTTGAGC CATAGGAGTG GGCAAATAAG CTATTTTCTA AGTAAAGCAT 22000 | [T/C] (21984)
TTCTGGAAAT AAATTTTCAA GTCCTCAATA CTACCAATGC TAGTAGTTGT 22050
ATAAGCGGAA TACTTAGTCC TTAGATGCAG AGATAAATAT ACTGTGTTAA 22100
TCCTAGTAAA GGATTCTTTA AGGAGAAAAA TCTTTAAAAA GAAGATTCAT 22150
TGACAATAAT AGAAATGCAT GAGCATCAGG CATGATGTTA AAAATATCTA 22200
TGGGGAAAAT AGGGCTGATA CAAACTAGGA AAAGAGGGAA TTTGGACAGA 22250
GCAGGGGATG ATCCCTGTGA GAGAGATTCT TGCTTATGGC ACTAGCGAAA 22300 | [C/G]_rs3211928 (22296)
AGAAAGTAAT AAAACAGTAA GGTTCTAATG TGTTTTCATC GTAATTTAGA 22350 | [A/G]_rs3211929 (22338)
AGTGCATCCT CTTTACCTTC TTCAGTAAGT TTGAAAAGAC TTAAAATGAA 22400 | [A/C]_rs3211930 (22351)
AGTATTGGTA TACATTTAGA ATGTTTCTTT TGGGAGGACA GAAAAAAAGC 22450 | [G/A]_rs3212014 (22434)
TCTGCAAATA TCAACTATGT TGCAGTCATT GTCCATGGTA GTTGCTGGGT 22500
ATAATTAGGA CATAATCTTC AATGATATTT TAGATAATGA TCTTAAAAGC 22550
TTTTCAATGC ATTTTTCTTT AATTATTAAG GAAATGGAAA AGCAGTTGTT 22600
GTCAATGTCC ATATGGGGGC TTATATTTTA CTTGTCACAA GAGTGTAAAC 22650 | [T/C] (22614)
TAATGAAAAC AAGGTCACAA GGTTGACACA CTTGCCAGAG AAAGAACCTA 22700
CAATGCATAA TCAAAGATGA GACTTACCAT TTAAAACTCA TTCAATGACG 22750 | [C/T]_rs3211931 (22749)
GCTCGTATTA GTGATATTCA TGTGAGAAGT ACATGGAGGA GTTAACTATT 22800
TCTTCATATA TCCAATGTTT GCCATGGATC ATAGCTGTAG CAAAAAGTAA 22850 | [C/T]_rs3212015 (22830)
ATTCAATAAA TGACTTTTTT ATACACTTTC AAAGCATTTG AAGTTCTGTT 22900
TAGTGAGAGA ACTTTAGTAA ATATTTTTAG AAGTAGTAAT CAGCCATTAG 22950 | [G/C] (22907)
GACAAATGAG AAAAAAAATC ACTACAAATA AATGTGGACA TGGCAGGAGA 23000
TCCAAATGAA CTTCACTGGA AGAAAAGTGC CACTCTACTG GTGGGGTAGG 23050 | [G/C] (23008)
GCATTTCAAA AAACAAACAC AATGTTAGCC TTAACATTTC ATGTTTAAGT 23100
TTCTTTTATT TTGTACCATT AAATATGTAT AGTATGTAGA TTTGTTGTTG 23150
ACAATAGCAG CCGCCAGCCA TATGTAACTG TTAAGGACTC AAAATCTGGC 23200 | REPEAT
TAGTATGATT TGAAATGTGC TGTAAATATA AAATGCACAG TAGATTTTGA 23250 | [G/A] (23219)
GACTTTAAGA ATTTAAAATA TTTTATAATG GTTACATGTT AAAATATTTT 23300
GGATATTAAT AAGTTAAATA TAGTTTTGTT TTGTTTGAGA CAAAGTTTTC 23350 | [G/C] (23345)| REPEAT
CACTTGTCGC CCAGGCTGGA GTGCAATGGC ATGATCTTGG CTCACTGCAG 23400
TATCCGCCTC CTGGGTGCAA GCAATTATCC TGCCTCAGGA TAGCTGGGAT 23450
TACGAGCGCC TATGTCAGGA TAGGTCAGGG AGGCCTATCC TGCCTCAGGA 23500 | [30bp/-] (23459_23488)| [A/G]
(23463)| [T/C]_rs3211932 (23463) | REPEAT
TAGCTGCCTC CCAAGTAGCT GGGATTACAA GCGCCTACCA CCACACCCAG 23550 | [G/A]_rs3173802 (23533)
CTAATTTTGT GTATTTTTAG TAGAGATGGG GTTTCACCAT GTAGGCCAGG 23600
CTGGTCTCAA ACCCCTGACC TCAGGTGATC CACCCGCCTT GGCCTTCCAA 23650
AGTGCTGGGC TTGCAGGCAT GAGCCAGCAT GCCTGGCCGG TTATTTCTTT 23700 | [G/A]_rs3173803 (23689) | REPEAT
TTACCTTTTA AAAATGTGGC TGCTAGATAA ATTACTTAGG CAGTTTGCAT 23750
TACATTTCTA TGGACTACAC TGGAGGAGAG ATTTCTAGGT TTTTTTCTAG 23800 | [G/A] (23780)
AACACACATT ACATCTAATC ATTTGCCACT CGATTTTTAA ACAGATGCAG 23850 | Exon 8
                                                  D  A   251
CCTCATTTCC ACCTTTTGTT GAGAAAAGCC AGGTATTGCA GTTCTTTTCT 23900
A  S  F  P   P  F  V   E  K  S   Q  V  L  H            265
TCTGATATTT GCAGGTAAGA CAGATACTGA AGTATAAGTA TGTCTGAGTC 23950
AGACCCCAGG TGACAAAATG CAGACCAAGA AACTTAAACA CAGCATAGGA 24000
AATTCATCAT GTTTATTAAC TAACTCTTTG CAAAATGTTC TTCTGCATCT 24050
TCCAATTTTT AATAGTACAA ATTTTTTTTT TTTTTGCCAT TTCTATCTAA 24100 | [A/T] (24071) | [T/C] (24080)
GCAAGAACCA TTTTGCCTTT TAAAAACTAA ACTAGTAGTC TATTAACGAT 24150 | [A/G] (24104)*
CAAGTCCAGA AGGGCGTGCC CAATCTTTCT AAAGACATAC AGAGAAGATG 24200 | [C/T]_rs3211933 (24165) |
[G/A]_rs3211934 (24166)| [C/G]_rs3211935 (24186)
ATGCAAATGT AACAGAACTG TGTTAATGTG CTCTGCTCAG ATTTGTGGGG 24250 | REPEAT
CCCAGTAGAT AACACACAAG CACTGGAGCC TTTACCACTA CCCTTGAGCC 24300
CAAGCTCTGT CCCTAAAGAC TGTATGCTTG TGGGTAGGCA TTTAACATCT 24350 | [T/C] (24331)
CTGAATTTTT TTCTTATCTG TACAATTTAA GACAGCAGTA TTTCTTCATA 24400
TACGTGTATC TGGGGATAAG AAAAATATGT GTATTGAACC CTATGATAGA 24450 | [T/A]_rs3173804 (24426) |
[A/T]_rs3211936 (24427)
CACTTGGTAA ACGATGGAAA TGTACCAGGT ATATATCCAG CATGTATATG 24500 | [A/G] (24496)
CATACACTCC AGTGAGTGGT CTTTCTTTCC AGGATTAATT AAGCCGTGAA 24550 | [C/T] (24545)
AGAAACTATT TCATTTAAAC TGATCACAAA TAAAGTATTT GAAGGAAGTC 24600 |
CTATAAATAT TTACTCTATT GGATAAATTG CCTGTGAGAA GTAACTTGAG 24650
```

```
TATAAATAAA CATGGTACTT CACAAACAAG AATAGTTCAT GCTTGGCTAT 24700 | [T/C] (24657)| [G/A]_rs3211937
(24685)
TGAGTTTTAG TATGTGTTAA AATTTCCCAA TCACTTTTTT TCTAAGAATG 24750
AAACAAGAAT TTAAAAGAGT ATATGATGTT TCTAAGTTAA AACAAGAATA 24800
AGAAAAAATG AATCTCCAGA ATGTAAGTTC AGGTTCCTGG AATGCAGCTC 24850 | Exon 9
                                                    S 266
TTTTTTCTCT GTATTTAGGT CAATCTATGC TGTATTTGAA TCCGACGTTA 24900
  F  F  S   V  F  R   S  I  Y  A   V  F  E   S  D  V   282
ATCTGAAAGG AATCCCTGTG TATAGATTTG TTCTTCCATC CAAGGCCTTT 24950
N  L  K  G   I  P  V   Y  R  F   V  L  P  S   K  A  F  299
GCCTCTCCAG TTGAAAACCC AGACAACTAT TGTTTCTGCA CAGAAAAAAT 25000
 A  S  P   V  E  N  P   D  N  Y   C  F  C   T  E  K  I 316
TATCTCAAAA AATTGTACAT CATATGGTGT GCTAGACATC AGCAAATGCA 25050 | [T/G]_rs3211938 (25025) | [G/T]
(25048)
  I  S  K   N  C  T   S  Y  G  V   L  D  I   S  K  C   332
AAGAAGGTGA GTAAATAACC TCAGTAGCAC AGTCCATACC ATAATTTGTG 25100 | [AAG/-]_rs3212016 (25054)
K  E                                                   334
ATATTCTTTA AGATGAGAAC TTTACCATAA TCCTTTAGCA ACCAAAATTT 25150
AAAATATATC ATAATTTGTG ATATTCTTTA AAATGAGAAC TTTACCATAA 25200
TCCTTTAGCA ACCAAAATTT AAAATTAAAG TAAGAAAGTA ATTAGGGCAG 25250
AAGAAAGAAT GGTGGCAGAA AATTTTAGTG CTGATTTTGT ATTTTGGGAA 25300 | [T/C]_rs3211939 (25276)
GATCCCACTT GTGTTTCAGT ATTACAAAAT TTAGTTAAAA CCACACCAGT 25350 | [C/T] (25305)
ATTTCCTTGT GGCTGCTTTT AGATTTAGGG TGAAATGAAA ATAATTCCGA 25400
GAACACATTA AACATCCTGT TATTCATCTG TCCTAACTTT TTTCACTAGA 25450
AAATGGTACA GGTAAATGTA TTTTCAGTAT GTATCTAAAG CTAGAGTTAA 25500
ACATAAAATT TGGAGACTAG CTTATCCTGT ACATATTTAT CATACTAACG 25550 | [G/A]_rs3211940 (25550)
TGGGTGTGGA AGAAGAAAGA AAAAACTAGT GTTAAATAAA TTCTTAGTCC 25600 | [-/A] (25575_25576)|
[T/C]_rs3173805 (25580)
ATAGACATAT TACTGCCTGA AAGCTTTACA TATTGAAAAT TAATACTGAA 25650
GGAGTTTATA GTAGAAATCA ACTGACATAA TTCTTCCCCA CCCATGTTAA 25700 | [T/C]_rs3211941 (25695)
AAACCATGTA TTTTTTAATG CAAGAAGCTT TAGTTTTGTG GAAATATTTT 25750
TTGAGTTATA TGTGAAATGA AGGAAGTTAT TAATTCCAAT TGACTCTTAA 25800 | [C/G]_rs3212017 (25794)
AACTTGTCTT CAGGGAGACC TGTGTACATT TCACTTCCTC ATTTTCTGTA 25850 | Exon 10 | [T/G] (25849) | [A/T]
(25850)
             G  R  P   V  Y  I   S  L  P   H  F  L  Y 347
TGCAAGTCCT GATGTTTCAG AACCTATTGA TGGATTAAAC CCAAATGAAG 25900
  A  S  P   D  V  S   E  P  I  D   G  L  N   P  N  E   363
AAGAACATAG GACATACTTG GATATTGAAC CTGTAAGAAA ACACCTTATT 25950 | [C/A]_rs3211942 (25945)
E  E  H  R   T  Y  L   D  I  E   P                     374
GATCTGATTT GGTTGATATT TTTAAAAATA CAATTGAAAT AAAAATAATC 26000
TTGTCGATGA TTATTTATTC AATAAATAAT CATATTTATT GAATCACATT 26050 | [TATT/-] (26016_26019)
CTTGAAAGTT ACTGAAACTT AGGTCGATTT CTTCCTATGG TTTATGAAGT 26100 | [G/A]_rs1527483 (26076) |
[G/T]_rs3211944 (26089)
GATTCTAATT GGCTTAAAAT AATTTTATAT AAATATTTAT GTTTAGATGA 26150
GGAGTTATTG TATATTATGC AAAAGTCAAA GAGTATATGT GAGTTTGACT 26200 | [T/C] (26158)| [G/T]_rs3211945
(26160)
ATACTTATTG TCAAAATCTA TAAAATTGTT GTAGCGCAAC AGTTTTAATC 26250
ATCTTTATTT TTGTATTTTC CTTTTCAAAA AGTAAATGAA ATTCTAGGAT 26300 | [T/C] (26262)
TTTAGTAGTT ATGTTTTAGT TTAAACAATG ACACATGGAT TCTAACTGAA 26350 | [G/A]_rs3211946 (26305) |
[C/T]_rs3211947 (26346)
TATATATTTG ACCAGGAATT ATCTGAGATC TTATATTTTG TTGCTGATTT 26400
TTGATTTTTT AAAAACCATT TCAAGTAACT CACAAATCTA ACTAACTAAA 26450 | [CTAA/-] (26446_26449)
ACCTTGACAT TCGATTGGGC AAATAAATTG TGTGTATCTA TATGGATGCA 26500 | [G/A] (26484)
TGTGTATATA ACTATAACTA TATATGCAGT TTTAAAAGTT TCAATTAGTC 26550 | [C/A]_rs1405747 (26512)
CTGTTTAACC TTAAGTTACT ACCTTCTCTT CTGCTGTAAG AAAAATAAGT 26600
TTTGAATAGT ATAAAATAAT GTTTTTAAAA GTTGGTAATT ATTTAGTTGT 26650
TCTCTTTTTA GATAACTGGA TTCACTTTAC AATTTGCAAA ACGGCTGCAG 26700 | Exon 11 | [G/T] (26669) | [C/T]
(26692)
             I  T  G   F  T  L   Q  F  A  K   R  L  Q  387
GTCAACCTAT TGGTCAAGCC ATCAGAAAAA ATTCAGTGAG TCTCTTGAAA 26750 | [G/C] (26740)
 V  N  L   L  V  K  P   S  E  K   I  Q                399
ATGGTTATTT TGATATGATC TGTAGTATCG TAGTATCTTC TTGTAAGAAC 26800 | [T/C]_rs3211949 (26776)
ATGAGTAAAT CTATGTAAGT AAGTGGAAAT AACATCTGAT ATCAACTTAT 26850 |[A/C] (26822)
CTTTAGCTTA ATGTCACCAA TCATTATTAA ATGCTTATGA CTAATTTCAC 26900
AGATTTTGGA ATGGTTTTAT GGTTTTATTT GAGCATTTGA TAGCATCATG 26950 | [A/G]_rs3211951 (26927) |
[ATGGTTTTATGGTTTTATTTGAGCATTTGATAGCATC/-]_rs3211950 (26948_26984)
GTTTTATGGT TTTATTTGAG CATTTGATAG CATCTCTTAT TTTGTTAGCT 27000
GCCCAAATAT TTCTATGACA ATAATTAATT TTTGGAATTC ATATTTCAGT 27050 | [C/T]_rs3211952 (27003) |
[A/G]_rs3211953 (27041)
TCCCCGAGAA TTTATTGAAA GGAAAAATCC ACACTTGTGA AAAAAAATCA 27100
ATGTGATTAG AAGACATATA AGAGCAAAGG AAGTCAAAAA CAACTATATT 27150
```

```
AAAATTTAAA TGAGTCATTA CAGGAACAAA ATCAAATTAG CAACAGCAAC 27200 | [A/T] (27163)
TAATTTATGA ACATTTATTT TAAAGTTTGT TATATATAAA TATTAGTTTA 27250 | [ATATAA/-] (27234_27239)
TATGTTCATA ATTATTTTCA ACGTATATTA CAGAGTATTA AAGAATCTGA 27300 | Exon 12
                                    V  L   K  N  L   404
AGAGGAACTA TATTGTGCCT ATTCTTTGGC TTAATGAGGT TTGTATTTGC 27350 |[T/C] (27309)
K  R  N  Y   I  V  P   I  L  W   L  N  E              417
AGCTGTTAGT CATTAAAAAC AACCTTCTTT GTATATAAAC AAGCTCTTGA 27400
TGTTTCAAAA GAATGTATAG TATTTAAAGC TATATGTATT TCCATTACCC 27450 |[G/T] (27411)| [T/C]_rs3211954
(27418)
ATATGGATGA GTATACATTT ATTTAACCTA TTTGAGATGA TCCAATTGAA 27500 | [A/G] (27494)
CAAAAACATT TCCTATCATT TAAGATTTTC TTCAAAAATG CATCTATTAA 27550
ACACATTTTC TTGTTGTAAC ATTTGTCTTC TATTGCCTGA CAAGGTATTT 27600
TTACTATAAA TCCATGCATT GATAGCTATA AAAATAGGAA AAACATTGAA 27650
TAAGTCTTTG GAGCAAATGA AACTGTTGAC CCTTTGATAG TTCTGAAGAG 27700
CAAATGAATC CTAGTACATT GAAGAGTACC GTACTCTATC TGGCACTTAA 27750
TTGCCTTTCT TGACTTGCAA AAGGAATTCC ATTAACTTGC CTTATAGATA 27800
CTGATGACTA ACACCAATAG AGGTGTTAGA AAAAAGGGTG ATAGGCAATT 27850
GAAGGGTTTA TTTTGTTTTA CTAACGTACC CAAATAATGT TGATTATTAA 27900
CTTGATTACA GACTGGGACC ATTGGTGATG AGAAGGCAAA CATGTTCAGA 27950 | Exon 13
            T  G  T   I  G  D   E  K  A  N   M  F  R  430
AGTCAAGTAA CTGGAAAAAT AAACCTCCTT GGCCTGATAG AAATGATCTT 28000
 S  Q  V   T  G  K  I   N  L  L   G  L  I   E  M  I  L 447
ACTCAGTGTT GGTGTGGTGA TGTTTGTTGC TTTTATGATT TCATATTGTG 28050
  L  S  V   G  V  V   M  F  V  A   F  M  I   S  Y  C   463
CATGCAGATC GAAAACAATA AAATAAGTAA GTATGTACCA AAAAATATTG 28100 | UTR |[AGTA/-] (28080_28083)
A  C  R  S   K  T  I   K                              471
CTTCAATAAT ATTAGCTTAT ATATTACTTG TTTTCACTTT ATCAAAGAGA 28150
AGTTACATAT TAGGCCATAT ATATTTCTAG ACATGTCTAG CCACTGATCA 28200
TTTTTAAATA TAGGTAAATA AACCTATAAA TATTATCACG CAGATCACTA 28250
AAGTATATCT TTAATTCTGG GAGAAATGAG ATAAAAGATG TACTTGTGAC 28300 | [G/T] (28278)
CATTGTAACA ATAGCACAAA TAAAGCACTT GTGCCAAAGT TGTCCAAAAT 28350 | [A/G]_rs8956 (28302) |
[GCACAAATAAAGCACT/-]_rs3212018 (28314_28329)
TGACTGGTTC ATTTCTCAAT TATATAGCTA GTTATATATT ATCTGATACT 28400 | [T/G]_rs3211956 (28375)
TAAAAATAAT TGACTAGGAA ATGGTTTCAT AAGACCAGGA TTGCTGCATG 28450 | [G/A] (28412)| [G/A]_rs3211957
(28423)
TAGACATGCT GGCCGTGCAT TTTCCAAATC CAGAAAAGTC CTGAACAAAA 28500
ATTTGAATAT AATAGTATCA GGAAATAGGG GAAAATAGTG TTCAACATAG 28550 | [-/AAAA] (28507_28508) | [A/T]
(28509)
TAGAACCAAG TACTCAGAGT GGTGTACAAG AGTTTAGGCC TCTATGCTTT 28600
AGATATTAGT GTCCATGCAC TCTTAAAGAT GAATGAATGC CTGACCTTTC 28650
CTAAAGGAAA ACCTTTTTTT AGTTATCACA GGTAACATTG GTGTTGCCTG 28700 | [A/G]_rs3211958 (28685)
TGGGGGAGAA TTTTTTACTT TCTCCCTATT TTTCTTCAGC CCACCTCAAG 28750 | [C/G] (28747)
GCAATGTAAA TAAAAGCAGT ATGTGTGTTA TAATAACTTA TTTAGTTGTT 28800
TACTGTGTGC CAGTACCATG CTAAAGAAAG ATTCTTTATC TCACATGATC 28850
CCTAATACAA TTTCATGATT TTCCAATAGC CTGGATTCAT CAGCATCCAT 28900
TCTATCTTCT AGAGATGTTC CATGAAATCA TACATTTTAA TTGTTTGTGA 28950 | [G/A]_rs3211959 (28947)
CCAAAGCATA AATCAGTGAA AGTGGAATTT GGGGACCAAT AATTTCCTTT 29000 | [AAAGTG/-] (28969_28974) | [A/G]
(28971)
TGTGAGATGG AGAGCTTGTG TTTGAAAAGG CAACCTAATT TTTGGTTCTA 29050
ATCACTTCTA CCACTATTTA GTCACCAAAA AGACAATAAT TCTGCATCCA 29100 | [G/A] (29082)
AACTATTTGG ACAGAATGGC TTCAAAATGC TAGCCTAAAA TGTTCACATT 29150 |[C/G] (29112)
ATAAAAAGTT AAATATTACC TTCAATACCT GTCAGTAGCC TACTGACAAA 29200 | [-/CAGT] (29186_29187)
TTATGACTAA ACAAAGGTAT TTGTATGACT ATGTAATAGA TCATCCGCTG 29250 | [A/G]_rs3211960 (29225)
AAAAGTAAAA CAAAATAACA AAAAAACTTG TCCTAATGGG AAAGCATGCT 29300 | [665bp/-] (29281_29945)
TAATAAAAGG AAATGCAGAA GTTATAAACA TGTTTTGTAA GTAAGTATTC 29350 | [AGTA/-] (29344_29347)
AGAATTAAAA TTATGTGATA CATTTTTATG ATTGCTTAAT GATCCTTGGA 29400
TGTCAGATTC CTTGGGTCTA TTTATAGCTA AATTATAATG AAAAATTCAA 29450
GGCTTGCTGG AGCAACTTTG TCAACAAATA TATTAGTTTT GCTTATATAT 29500
TTGATTTTTA TGTGGAAAAA TTACTACCCT TTTTTACAAG CAGAGAATAA 29550
ACTGTTGATT ACTTGATTTA CTAGATTTAG AAGAATCACA AAAGATATGT 29600
AGATTTTCTT AAGCAAAATT CAGCTCTTAA TATCATAAAA ATTATATCTT 29650
TGGGCAGATT TGTAAACAAT AGGAACAAGT AAGAAGACAG GTATGTAAGA 29700
AGTAGCAAAG TGTAAGACGG TGGAGCTTTA ATGTGTGGTT TTACTCAGGG 29750 | [C/A] (29706)
GGTCACAAGA AAATAACCCA ATGGTCCTTT TGAAAGTAGT ACATACACCT 29800
TTAAATGGAA CTTGGTCTGA AGGGTGTGGA AATGCTGGTC CAGGGCAGAT 29850 | [G/A] (29815)
GCACTTCAGC TACCGTTCCT TGCCTGGTGT GTGCTTGTCT TGAGGGTTGA 29900 | [GGGTTGAGA/-]_rs3044712
(29894_29902)
```

```
GAAGCACCCT TCTTATGTGC CTTTCTGGAA TGGGAATGTG TTTTGCCCTA 29950
TTGGTGAAGT CCGTTAGAGA ACTTCTCTCT GCAGCTAGAG AAACTGAACT 30000 | [G/A] (29963)
GATTAATAAA AAGTATGATA TCTAAGATAT ACTGCCACTC TACAGGTAGT 30050 | [C/G] (30040)
AAATGATTCT TCACATTTGT GCGCCAGCTA CATCATGACC CTTGGTAATA 30100
CCGACCAACA GGATGAAAAT TGGATATGCT TTACTTTTAC AAAGCACACC 30150
TTATCTTTTA GTTGAAGAAA TGGTGGCACT
```

Figure 5. *CD36* annotated sequence.

3.2 DISTRIBUTION OF *CD36* VARIANTS IN HIGH AND LOW HDL GROUPS

3.2.1 Non-Hispanic Whites

Of the 131 variants identified in NHWs, 44 had a MAF ≥5%, 52 had a MAF 1-5%, and 35 had a MAF <1%. Of the variants identified in our study but not previously identified in the SeattleSNPs database the MAF range was 0.005-0.396%. Among the common variants (MAF ≥5%), three variants had a statistically significant difference when comparing the allele frequencies between the high HDL-C and low HDL-C groups: 4249C>T(p=0.047), 13094T>A(p=0.040), and 14299A>G(p=0.047). However, 13 of the common variants have a p-value of between 5-10%, which may be statistically significant due to the small sample size.

Of the relatively uncommon or rare variants (MAF<5%), 21 were present only in the low HDL-C group versus 25 present only in the high HDL-C group. For NHWs: 14 out of 47 (29.8%) individuals with high HDL-C had more than two rare variants versus 16 out of 48 (33.3%) individuals with low HDL-C; and 11 out of 47 (23.4%) individuals with high HDL-C had more than three rare variants versus 6 out of 48 (12.5%) individuals with low HDL-C:, and 8 out of 47 (17.0%) individuals with high HDL-C had more than four rare variants versus 5 out of 48 (10.4%) individuals with low HDL-C. Of five exonic variants identified, one was found only in

the low HDL-C group (27309T>C, p=0.321), one was found only in the high HDL-C group (26669G>T, p=0.311), and two of them were found in both the high and low HDL-C groups. Table 10 summarizes the distribution of common *CD36* variants, and Table 11 summarized the distribution of relatively uncommon or rare *CD36*. Those variants highlighted in yellow are only found in the low HDL-C group and those highlighted in blue are only found in the high HDL-C groups.

Table 10. Distribution of common *CD36* variants in high and low HDL-C groups in NHWs.

CD36 Variant	Alleles	refSNP ID	Location	All MAF (n=95)	High HDL-C MAF (n=47)	Low HDL-C MAF (n=48)	*p*-value
2521	G>A	rs3211816	Intron 2	0.389	0.340	0.438	0.170
2996	C>A	rs3211820	Intron 2	0.405	0.351	0.458	0.132
3094	G>A	rs3211821	Intron 2	0.453	0.426	0.479	0.458
3157	G>A	rs3211822	Intron 2	0.404	0.348	0.458	0.123
3991	A>C	rs3211827	Intron 2	0.389	0.340	0.438	0.170
4249	C>T	rs3211830	Intron 2	0.053	0.085	0.021	**0.047**
4366	A>T	rs997906	Intron 2	0.395	0.340	0.448	0.130
4648	C>A	rs3211834	Intron 2	0.405	0.351	0.458	0.132
4993	G>A	rs3211839	Intron 2	0.441	0.413	0.468	0.450
7167	G>A	rs3211842	Intron 2	0.437	0.415	0.458	0.546
7854	A>G	rs3211849	Intron 2	0.442	0.415	0.469	0.455
9473	C>T	rs1054516	Intron 2	0.452	0.426	0.479	0.464
9534	C>T	rs1054517	Intron 2	0.436	0.415	0.457	0.556
9600	A>G	rs1133344	Intron 2	0.065	0.085	0.043	0.248
9794_9799	del6		Intron 2	0.058	0.097	0.026	0.078
10876	A>G	rs3211864	Intron 3	0.063	0.096	0.031	0.068
11741	C>T	rs3211870	Intron 3	0.437	0.415	0.458	0.546
11890	A>G	rs3211871	Intron 3	0.388	0.337	0.438	0.157
12132	T>C		Intron 3	0.388	0.337	0.438	0.157
12143_12144	del2		Intron 3	0.389	0.330	0.446	0.110
12272_12274	del3	rs3211874	Intron 3	0.379	0.321	0.433	0.129
12919	G>A	rs1358337	Intron 3	0.437	0.415	0.458	0.546
13094	T>A	rs3211879	Intron 3	0.068	0.106	0.031	**0.040**
13388	A>G	rs3211881	Intron 3	0.062	0.091	0.034	0.120
13528_13530	del3	rs3211882	Intron 3	0.396	0.356	0.435	0.275
13577	T>A	rs3211883	Intron 3	0.06	0.089	0.032	0.103
13936	C>T	rs3211885	Intron 3	0.437	0.415	0.458	0.546
13973	G>A	rs3211886	Intron 3	0.389	0.340	0.438	0.170
14299	A>G	rs3173799	Intron 3	0.053	0.085	0.021	**0.047**
14455	A>T	rs3173800	Intron 3	0.389	0.340	0.438	0.170
19307_19335	del29	rs3211915	Intron 6	0.468	0.447	0.490	0.555
22296	C>G	rs3211928	Intron 7	0.437	0.500	0.375	0.097
22749	C>T	rs3211931	Intron 7	0.426	0.468	0.385	0.249
23463	T>C	rs3211932	Intron 7	0.426	0.468	0.385	0.249

Table 10 Continued

23533	G>A	rs3173802	Intron 7	0.426	0.468	0.385	0.249
23689	G>A	rs3173803	Intron 7	0.426	0.468	0.385	0.249
24426	T>A	rs3173804	Intron 8	0.426	0.468	0.385	0.249
25580	T>C	rs3173805	Intron 9	0.426	0.468	0.383	0.238
26512	C>A	rs1405747	Intron 10	0.426	0.468	0.385	0.249
27003	C>T	rs3211952	Intron 11	0.426	0.468	0.385	0.249
28314_28329	del16	rs3212018	**Exon-3'UTR**	0.163	0.181	0.146	0.514
28685	A>G	rs3211958	3' flanking	0.426	0.468	0.385	0.249
29225	A>G	rs3211960	3' flanking	0.426	0.468	0.385	0.249
29894_29902	del9	rs3044712	3' flanking	0.426	0.467	0.385	0.256

Table 11. Distribution of relatively uncommon or rare *CD36* variants in high and low HDL-C groups in NHWs.

CD36 Variant	Alleles	refSNP ID	Location	All MAF (n=95)	High HDL-C MAF (n=47)	Low HDL-C MAF (n=48)
271_273	del3		Intron 1	0.037	0.011	0.062
861	T>C	rs1527463	Intron 2	0.016	0.011	0.021
1024	A>G	rs3211809	Intron 2	0.016	0.011	0.021
2389	delT	rs3211815	Intron 2	0.047	0.074	0.021
2638	T>G	rs3211817	Intron 2	0.048	0.076	0.021
2688	A>G		Intron 2	0.005	0.011	0.000
3049	G>A		Intron 2	0.011	0.000	0.021
3304	C>G	rs3212000	Intron 2	0.043	0.043	0.042
3349_3350	insA	rs3211823	Intron 2	0.011	0.000	0.021
3691	G>A	rs3211825	Intron 2	0.016	0.032	0.000
4108	G>A		Intron 2	0.005	0.011	0.000
4134	T>G	rs3211828	Intron 2	0.047	0.074	0.021
4595	G>A		Intron 2	0.005	0.000	0.010
4990	C>T	rs3211838	Intron 2	0.043	0.067	0.021
5497	A>G		Intron 2	0.005	0.011	0.000
6146	T>C		Intron 2	0.005	0.011	0.000
6652	G>T		Intron 2	0.005	0.000	0.010
7306	G>A	rs3212001	Intron 2	0.026	0.011	0.042
7664	G>A	rs3212002	Intron 2	0.005	0.000	0.010
7988	A>C	rs3211851	Intron 2	0.048	0.076	0.021
8595	T>A	rs3211855	Intron 2	0.016	0.011	0.021
8639	G>A		Intron 2	0.005	0.000	0.011
9136	A>T		Intron 2	0.005	0.011	0.000
9347	C>G		Intron 2	0.005	0.011	0.000
9616	T>C	rs3212003	Intron 2	0.018	0.017	0.019
9663	G>A	rs3212004	Intron 2	0.005	0.011	0.000
9786	delA		Intron 2	0.036	0.065	0.013
10279	C>T		Intron 2	0.005	0.000	0.010
10381	T>C	rs3173798	Intron 2	0.047	0.074	0.021
11249	T>G		Intron 3	0.005	0.011	0.000
11440	T>C		Intron 3	0.016	0.021	0.010
11472	C>A	rs3211867	Intron 3	0.043	0.067	0.021
11554	T>C	rs3211868	Intron 3	0.048	0.076	0.021
11618	A>G		Intron 3	0.005	0.011	0.000

Table 11 Continued

11684	T>A	rs3211869	Intron 3	0.043	0.065	0.021
11904	G>C		Intron 3	0.005	0.000	0.010
12145	G>A	rs3211873	Intron 3	0.039	0.067	0.000
12403	C>G	rs3211875	Intron 3	0.039	0.068	0.011
12642	G>A		Intron 3	0.016	0.011	0.021
12775	A>G	rs3211876	Intron 3	0.026	0.000	0.052
12936	A>T		Intron 3	0.005	0.000	0.010
13279	C>T		Intron 3	0.006	0.000	0.011
13456	A>G		Intron 3	0.011	0.011	0.011
14510	G>A		Intron 3	0.005	0.000	0.010
14903	G>A	rs3211892	Intron 3	0.016	0.011	0.021
15075	A>G		Intron 4	0.005	0.011	0.000
15833	C>T	rs3212008	Intron 4	0.005	0.011	0.000
15932	G>A	rs1924	Intron 4	0.048	0.076	0.021
16377	A>G		Intron 4	0.016	0.011	0.021
16385	C>T	rs3212009	Intron 4	0.011	0.011	0.010
16824	G>T		Intron 4	0.005	0.000	0.010
16983	G>A	rs5956	**Exon 5 (Synonymous)**	0.042	0.043	0.042
17274	A>T		Intron 5	0.005	0.011	0.000
17282_17823	insT		Intron 5	0.005	0.011	0.000
17641	G>A		Intron 5	0.016	0.000	0.031
17743	A>G	rs3211905	Intron 5	0.016	0.011	0.021
18137	delA		Intron 5	0.016	0.011	0.021
18463	C>T	rs3211908	Intron 6	0.032	0.043	0.021
18662	T>C	rs3211909	Intron 6	0.016	0.011	0.021
18724	A>C		Intron 6	0.016	0.011	0.021
18726	A>C		Intron 6	0.022	0.000	0.043
18966	G>A	rs3211912	Intron 6	0.016	0.011	0.021
19151	A>G	rs3211913	Intron 6	0.01	0.000	0.024
19228	A>C	rs3211914	Intron 6	0.042	0.074	0.000
19678	T>G		Intron 6	0.016	0.011	0.021
19810	G>A		Intron 6	0.016	0.011	0.021
20630	C>T	rs3212013	Intron 7	0.021	0.032	0.010
20758	C>G		Intron 7	0.011	0.000	0.021
20759	T>C		Intron 7	0.005	0.011	0.000
21084	C>T		Intron 7	0.005	0.011	0.000
21110	C>T	rs3211922	Intron 7	0.016	0.000	0.031
22830	C>T	rs3212015	Intron 7	0.005	0.011	0.000
22907	G>C		Intron 7	0.005	0.011	0.000
24071	A>T		Intron 8	0.006	0.000	0.011

Table 11 Continued

24104	A>G		Intron 8	0.006	0.012	0.000
24331	T>C		Intron 8	0.005	0.011	0.000
25575_25576	insA		Intron 9	0.016	0.011	0.021
26076	G>A	rs1527483	Intron 10	0.037	0.054	0.021
26669	G>T		**Exon 11 (Non-synonymous)**	0.005	0.011	0.000
26822	A>C		Intron 11	0.005	0.011	0.000
27309	T>C		**Exon 12 (Non-synonymous)**	0.005	0.000	0.010
27411	G>T		Intron 12	0.026	0.032	0.021
28080_28083	del4		**Exon- 3'UTR**	0.021	0.032	0.010
28375	T>G	rs3211956	3' flanking	0.037	0.054	0.021
28412	G>A		3' flanking	0.005	0.011	0.000
28572	G>T		3' flanking	0.011	0.011	0.010
29112	C>G		3' flanking	0.005	0.000	0.010

3.2.2 Blacks

Of the 281 variants identified in Blacks, 103 had a MAF ≥5%, 110 had a MAF 1-5%, and 68 had a MAF <1%. Of the variants identified in our study but not previously identified in the SeattleSNPs database the MAF range was 0.005-0.289%. Among the common variants (MAF ≥5%), one variant had a statistically significant difference when comparing the allele frequencies between the high HDL-C and low HDL-C groups: 16568A>G (p=0.022). Of the relatively uncommon or rare variants (MAF<5%), 59 were present only in the low HDL-C group versus 32 present only in the high HDL-C group. For Blacks: 32 out of 48 (66.7%) individuals with high HDL-C had more than two rare variants versus 34 out of 47 (72.3%) individuals with low HDL-C; 28 out of 48 (58.3%) individuals with high HDL-C had more than three rare variants versus 30 out of 47 (63.8%) individuals with low HDL-C; 22 out of 48(45.8%) individuals with high HDL-C had more than four rare variants versus 24 out of 47 (51.1%) individuals with low HDL-C. Of 13 exonic variants identified, four were found only in the low HDL-C group (25048G>T, p=0.311; 25849T>G, p=0.311; 26692C>T, p=0.311; 10465G>A, p=0.311), one was found only in the high HDL-C group (25850A>T, p=0.087), and eight of them were found in both the high and low HDL-C groups (one of these variants was missense). Table 12 summarizes the distribution of common *CD36* variants, and Table 13 summarized the distribution of relatively uncommon or rare *CD36*. Those variants highlighted in yellow are only found in the low HDL-C group and those highlighted in blue are only found in the high HDL-C group.

We particularly looked at variants causing nonsense and frameshift changes in the Black population. Except for two (10465G>A and 14701_15292del592), the remaining were not uniquely present in low HDL-C groups. 10465 was in exon 3 causing a nonsense change and was present in an individual in the low HDL-C group. 14701_15292 results in the partial removal of intron 3, total removal of exon 4, and partial removal of intron 4 was present in the individual with the lowest level of HDL-C in our sequence sample.

Table 12. Distribution of common CD36 variants in high HDL-C and low HDL-C groups in Blacks

CD36 Variant	Alleles	refSNP ID	Location	All MAF (n=95)	High HDL-C MAF (n=47)	Low HDL-C (n=48)	P value
106	G>T	rs3211805	5' flanking	0.063	0.042	0.085	0.218
947_948	del2		Intron 2	0.084	0.104	0.064	0.317
949_950	insA		Intron 2	0.084	0.104	0.064	0.317
1021	T>C	rs3211808	Intron 2	0.068	0.038	0.095	0.151
1024	A>G	rs3211809	Intron 2	0.244	0.256	0.232	0.716
1547	T>G	rs3211810	Intron 2	0.147	0.156	0.138	0.727
1848	T>C	rs3211811	Intron 2	0.058	0.031	0.085	0.112
2162	A>G	rs3211812	Intron 2	0.058	0.031	0.085	0.112
2389	delT	rs3211815	Intron 2	0.205	0.219	0.191	0.642
2521	G>A	rs3211816	Intron 2	0.100	0.115	0.085	0.498
2638	T>G	rs3211817	Intron 2	0.105	0.094	0.117	0.601
2652	A>C	rs3211818	Intron 2	0.084	0.104	0.064	0.317
2996	C>A	rs3211820	Intron 2	0.416	0.438	0.394	0.540
3094	A>G	rs3211821	Intron 2	0.295	0.292	0.298	0.925
3157	G>A	rs3211822	Intron 2	0.400	0.427	0.372	0.441
3350	delA	rs3211823	Intron 2	0.196	0.208	0.182	0.651
3412	C>T	rs3211824	Intron 2	0.054	0.031	0.080	0.149
3714_3715	del2	rs3211826	Intron 2	0.168	0.177	0.160	0.747
3991	A>C	rs3211827	Intron 2	0.074	0.106	0.043	**0.096**
4134	T>G	rs3211828	Intron 2	0.221	0.229	0.213	0.785
4249	C>T	rs3211830	Intron 2	0.116	0.104	0.128	0.613
4259	T>C	rs3211831	Intron 2	0.089	0.104	0.074	0.473
4366	A>T	rs997906	Intron 2	0.095	0.115	0.074	0.345
4648	C>A	rs3211834	Intron 2	0.305	0.333	0.277	0.396
4879	G>C	rs3211836	Intron 2	0.094	0.117	0.070	0.279
4990	C>T	rs3211838	Intron 2	0.200	0.223	0.174	0.412
4993	A>G	rs3211839	Intron 2	0.489	0.511	0.465	0.542
7167	G>A	rs3211842	Intron 2	0.289	0.323	0.255	0.304
7175	A>G	rs3211843	Intron 2	0.106	0.106	0.106	1.000
7430	C>A	rs3211845	Intron 2	0.054	0.031	0.078	0.160
7854	A>G	rs3211849	Intron 2	0.489	0.489	0.489	1.000
7947	G>A	rs3211850	Intron 2	0.063	0.062	0.064	0.970
7988	A>C	rs3211851	Intron 2	0.191	0.213	0.170	0.458
8595	T>A	rs3211855	Intron 2	0.210	0.217	0.202	0.798
8796	A>G	rs3211856	Intron 2	0.055	0.054	0.056	0.972
9152	delC	rs3211857	Intron 2	0.065	0.096	0.033	**0.087**
9473	T>C	rs1054516	Intron 2	0.277	0.250	0.307	0.390
9474	G>A	rs3211858	Intron 2	0.056	0.054	0.057	0.942

Table 12 Continued

9505	T>A	rs3211859	Intron 2	0.054	0.052	0.056	0.916
9534	C>T	rs1054517	Intron 2	0.261	0.312	0.205	**0.096**
9600	G>A	rs1133344	Intron 2	0.394	0.402	0.385	0.818
9786	delA		Intron 2	0.193	0.217	0.167	0.395
9794_9799	del6		Intron 2	0.051	0.033	0.071	0.243
10381	T>C	rs3173798	Intron 2	0.189	0.208	0.170	0.503
11137	delG	rs3211865	Intron 3	0.069	0.096	0.043	0.151
11472	C>A	rs3211867	Intron 3	0.300	0.302	0.298	0.950
11554	T>C	rs3211868	Intron 3	0.189	0.208	0.170	0.503
11684	T>A	rs3211869	Intron 3	0.188	0.207	0.170	0.526
11741	C>T	rs3211870	Intron 3	0.371	0.402	0.340	0.383
12132	T>C		Intron 3	0.072	0.098	0.045	0.175
12143_12144	del2		Intron 3	0.088	0.120	0.056	0.127
12145	G>A	rs3211873	Intron 3	0.162	0.203	0.125	0.191
12272_12274	del3	rs3211874	Intron 3	0.090	0.120	0.058	0.152
12403	C>G	rs3211875	Intron 3	0.101	0.139	0.066	0.141
12919	G>A	rs1358337	Intron 3	0.463	0.447	0.479	0.661
12998	C>A	rs3211878	Intron 3	0.058	0.031	0.085	0.112
13094	T>A	rs3211879	Intron 3	0.058	0.042	0.074	0.333
13388	A>G	rs3211881	Intron 3	0.056	0.034	0.076	0.219
13528_13530	del3	rs3211882	Intron 3	0.109	0.130	0.087	0.343
13577	A>T	rs3211883	Intron 3	0.363	0.359	0.367	0.911
13758	T>C	rs3211884	Intron 3	0.060	0.033	0.087	0.120
13936	C>T	rs3211885	Intron 3	0.311	0.344	0.277	0.317
13973	G>A	rs3211886	Intron 3	0.079	0.115	0.043	**0.066**
14081	C>A	rs3211888	Intron 3	0.058	0.031	0.085	0.112
14299	A>G	rs3173799	Intron 3	0.189	0.208	0.170	0.503
14455	A>T	rs3173800	Intron 3	0.105	0.125	0.085	0.370
14482	T>G	rs3211890	Intron 3	0.080	0.096	0.064	0.419
14882	T>C	rs3211891	Intron 3	0.063	0.031	0.096	**0.068**
14903	G>A	rs3211892	Intron 3	0.321	0.333	0.309	0.714
15554_15557	del3		Intron 4	0.063	0.031	0.096	**0.068**
15932	G>A	rs1924	Intron 4	0.237	0.250	0.223	0.666
15975	A>G	rs3211895	Intron 4	0.068	0.042	0.096	0.140
16040	G>T	rs3211896	Intron 4	0.068	0.052	0.085	0.367
16163	G>A	rs3211897	Intron 4	0.090	0.106	0.074	0.446
16568	A>G	rs3211899	Intron 4	0.059	0.021	0.100	**0.022**
17743	A>G	rs3211905	Intron 5	0.080	0.062	0.098	0.372
17807_17813	del7	rs3211906	Intron 5	0.280	0.255	0.304	0.456
18137	delA		Intron 5	0.352	0.302	0.412	0.127
18662	T>C	rs3211909	Intron 6	0.335	0.267	0.405	**0.058**
18911	T>G	rs3211911	Intron 6	0.084	0.104	0.064	0.317
19151	G>A	rs3211913	Intron 6	0.494	0.545	0.446	0.181
19386	T>C	rs3211916	Intron 6	0.289	0.260	0.319	0.372
19678	T>G		Intron 6	0.333	0.277	0.391	**0.097**
19810	G>A		Intron 6	0.389	0.333	0.447	0.109
19877	T>C	rs3211917	Intron 6	0.053	0.052	0.053	0.973

Table 12 Continued

20644	T>G	rs3211920	Intron 7	0.089	0.083	0.096	0.764
22296	C>G	rs3211928	Intron 7	0.174	0.125	0.223	**0.073**
22351	A>C	rs3211930	Intron 7	0.079	0.062	0.096	0.396
22749	C>T	rs3211931	Intron 7	0.170	0.125	0.217	**0.092**
23463	T>C	rs3211932	Intron 7	0.165	0.117	0.213	**0.077**
23533	G>A	rs3173802	Intron 7	0.069	0.052	0.087	0.346
23689	G>A	rs3173803	Intron 7	0.168	0.125	0.213	0.106
24080	T>C		Intron 8	0.059	0.031	0.087	0.104
24166	G>A	rs3211933	Intron 8	0.054	0.062	0.045	0.610
24426	T>A	rs3173804	Intron 8	0.168	0.125	0.213	0.106
25025	T>G	rs3211938	**Exon 9 (Nonsense)**	0.226	0.234	0.217	0.786
25580	T>C	rs3173805	Intron 9	0.060	0.043	0.078	0.314
25945	C>A	rs3211942	Intron 10	0.063	0.031	0.096	**0.068**
26512	C>A	rs1405747	Intron 10	0.207	0.160	0.256	0.108
27003	C>T	rs3211952	Intron 11	0.111	0.094	0.128	0.456
28685	A>G	rs3211958	3' flanking	0.082	0.053	0.111	0.151
29225	A>G	rs3211960	3' flanking	0.059	0.042	0.078	0.297
29894_29902	del9	rs3044712	3' flanking	0.159	0.109	0.211	**0.059**

Table 13. Distribution of relatively uncommon or rare variants in high HDL-C and low HDL-C groups in Blacks

CD36 Variant	Alleles	refSNP ID	Location	All MAF (n=95)	High HDL-C MAF (n=47)	Low HDL-C MAF (n=48)
400	C>T		Intron 1	0.005	0.010	0.000
861	T>C	rs1527463	Intron 2	0.037	0.052	0.021
1466	G>T		Intron 2	0.005	0.000	0.011
1529	G>A		Intron 2	0.005	0.000	0.011
1675	C>A		Intron 2	0.005	0.000	0.011
2092	A>G		Intron 2	0.005	0.000	0.011
2273	T>G	rs3211813	Intron 2	0.037	0.042	0.032
2298	C>A		Intron 2	0.026	0.031	0.021
2306	A>G	rs3211814	Intron 2	0.037	0.000	0.074
2702	T>C		Intron 2	0.005	0.000	0.011
2856	A>G		Intron 2	0.005	0.010	0.000
3079	C>G		Intron 2	0.032	0.031	0.032
3505	C>G		Intron 2	0.005	0.000	0.011
3835	delA		Intron 2	0.026	0.010	0.043
3852	G>A		Intron 2	0.005	0.010	0.000
4039	C>T		Intron 2	0.005	0.010	0.000
4046	G>T		Intron 2	0.005	0.010	0.000
4133	T>G		Intron 2	0.005	0.000	0.011
4266	C>T		Intron 2	0.005	0.000	0.011
4308	A>G		Intron 2	0.005	0.000	0.011
4752	C>A	rs3211835	Intron 2	0.016	0.014	0.019
5081	C>T		Intron 2	0.005	0.000	0.011
5241	G>C	rs3211840	Intron 2	0.022	0.022	0.023
5290	G>A		Intron 2	0.006	0.000	0.011
5504	T>A		Intron 2	0.005	0.000	0.011
5913	delA		Intron 2	0.016	0.032	0.000
5950	C>T		Intron 2	0.005	0.000	0.011
6007	C>T	rs3211841	Intron 2	0.016	0.010	0.021
6238	A>G		Intron 2	0.005	0.000	0.011
6638	G>T		Intron 2	0.032	0.031	0.032
6877	G>A		Intron 2	0.005	0.010	0.000
7037_7038	insA		Intron 2	0.032	0.031	0.032
7046	G>C		Intron 2	0.005	0.011	0.000
7265	G>T	rs3211844	Intron 2	0.021	0.010	0.032
7351	T>A		Intron 2	0.005	0.000	0.011
7414	C>T		Intron 2	0.005	0.000	0.011
7486	C>A	rs3211846	Intron 2	0.021	0.010	0.033
7663	C>T	rs3211848	Intron 2	0.038	0.052	0.022

Table 13 Continued

8152	G>A	rs3211852	Intron 2	0.021	0.031	0.011
8228	G>A		Intron 2	0.005	0.011	0.000
8414	A>T		Intron 2	0.011	0.011	0.011
8427	C>A		Intron 2	0.016	0.011	0.021
8583	C>T	rs3211854	Intron 2	0.043	0.054	0.032
8873	A>G		Intron 2	0.011	0.021	0.000
9045	T>A		Intron 2	0.032	0.031	0.033
9117	T>G		Intron 2	0.033	0.022	0.044
9291	G>A		Intron 2	0.005	0.000	0.011
9699	T>C		Intron 2	0.021	0.021	0.021
10045	C>T		Intron 2	0.011	0.000	0.022
10103	T>C		Intron 2	0.005	0.000	0.011
10342	T>C		Intron 2	0.005	0.000	0.011
10423_10424	del2	rs3211861	**Exon 3 (Frameshift)**	0.016	0.010	0.021
10465	G>A		**Exon 3 (Nonsense)**	0.005	0.000	0.011
10654	C>A		Intron 3	0.005	0.010	0.000
10876	A>G	rs3211864	Intron 3	0.037	0.042	0.032
10975	G>A		Intron 3	0.005	0.010	0.000
11155	C>T	rs3211866	Intron 3	0.021	0.011	0.032
11640	G>A		Intron 3	0.011	0.010	0.011
11890	A>G	rs3211871	Intron 3	0.026	0.042	0.011
11943	C>T		Intron 3	0.021	0.000	0.043
12124	T>C	rs3211872	Intron 3	0.021	0.010	0.032
12125_12126	ins2		Intron 3	0.033	0.021	0.044
12609	G>C		Intron 3	0.011	0.010	0.011
12691	A>T		Intron 3	0.016	0.021	0.011
12745	A>C		Intron 3	0.005	0.000	0.011
13346	T>C	rs3211880	Intron 3	0.034	0.034	0.033
13479	C>G		Intron 3	0.027	0.022	0.033
13548	C>T		Intron 3	0.005	0.011	0.000
14075	C>T		Intron 3	0.005	0.010	0.000
14076	G>A		Intron 3	0.011	0.000	0.021
14190	T>C	rs3211889	Intron 3	0.016	0.000	0.032
14500	T>C		Intron 3	0.005	0.010	0.000
14701_15292	del592		**Intron 3-Intron 4**	0.005	0.000	0.011
15062	T>C	rs3211893	Intron 4	0.043	0.043	0.043
15405	A>G		Intron 4	0.005	0.000	0.011
15425	G>A		Intron 4	0.026	0.021	0.032
15676	A>G		Intron 4	0.016	0.010	0.021
16051	T>C		Intron 4	0.005	0.000	0.011
16295	C>T		Intron 4	0.016	0.021	0.011
16565	T>G		Intron 4	0.011	0.021	0.000
16741	A>G		Intron 4	0.021	0.010	0.033
16986	C>A		**Exon 5 (Nonsense)**	0.011	0.010	0.011

Table 13 Continued

17093	T>A	rs3173801	Intron 5	0.042	0.042	0.043
17282	T>A		Intron 5	0.042	0.042	0.043
17292	G>A		Intron 5	0.005	0.000	0.011
17421_17433	del13	rs3211903	Intron 5	0.047	0.052	0.043
17437	T>G		Intron 5	0.005	0.000	0.011
17462	delT	rs3211904	Intron 5	0.042	0.042	0.043
18119	G>A		Intron 5	0.009	0.000	0.018
18364_18367	del4	rs3211907	Intron 6	0.042	0.042	0.043
18463	C>T	rs3211908	Intron 6	0.012	0.000	0.024
18596_18597	insT		Intron 6	0.006	0.011	0.000
18650_18651	ins3		Intron 6	0.006	0.011	0.000
18659	T>C		Intron 6	0.005	0.010	0.000
18784	C>T		Intron 6	0.005	0.000	0.011
18785	G>A	rs3211910	Intron 6	0.043	0.042	0.044
18821_18833	del13		Intron 6	0.005	0.000	0.011
18825	G>A		Intron 6	0.005	0.010	0.000
18966	G>A	rs3211912	Intron 6	0.011	0.000	0.021
19260_19261	ins4		Intron 6	0.016	0.031	0.000
19426	C>T		Intron 6	0.005	0.010	0.000
19673	A>T		Intron 6	0.016	0.021	0.011
19742	A>G		Intron 6	0.005	0.010	0.000
20108	T>G		Intron 6	0.011	0.000	0.021
20119_20120	ins23		Intron 6	0.016	0.021	0.011
20141	T>A		Intron 6	0.005	0.010	0.000
20440	A>G	rs3211919	Intron 7	0.032	0.031	0.032
20584	A>T		Intron 7	0.026	0.021	0.032
20843	T>C		Intron 7	0.005	0.010	0.000
21034	G>A		Intron 7	0.011	0.010	0.011
21111	G>A	rs3211923	Intron 7	0.011	0.010	0.011
21138	G>A		Intron 7	0.021	0.021	0.021
21174	T>C		Intron 7	0.011	0.010	0.011
21320_21321	ins4		Intron 7	0.005	0.010	0.000
21728	C>T		Intron 7	0.007	0.000	0.015
21882	G>A		Intron 7	0.027	0.031	0.022
21891	G>A	rs3211926	Intron 7	0.011	0.010	0.011
21927	C>T	rs3211927	Intron 7	0.011	0.010	0.011
21984	T>C		Intron 7	0.026	0.031	0.021
22338	A>G	rs3211929	Intron 7	0.011	0.010	0.011
22614	T>C		Intron 7	0.011	0.010	0.011
23008	G>C		Intron 7	0.005	0.000	0.011
23219	G>A		Intron 7	0.005	0.000	0.011
23345	G>C		Intron 7	0.005	0.010	0.000
23459_23488	del30		Intron 7	0.005	0.010	0.000
23780	G>A		Intron 7	0.026	0.031	0.021
24071	A>T		Intron 8	0.011	0.000	0.022
24165	C>T	rs3211933	Intron 8	0.033	0.042	0.023
24186	C>G	rs3211935	Intron 8	0.011	0.010	0.012

Table 13 Continued

24427	A>T	rs3211936	Intron 8	0.037	0.021	0.053
24496	A>G		Intron 8	0.005	0.000	0.011
24545	C>T		Intron 8	0.021	0.042	0.000
24657	T>C		Intron 8	0.026	0.031	0.021
25048	G>T		**Exon 9 (Non-synonymous)**	0.005	0.000	0.011
25276	T>C	rs3211939	Intron 9	0.037	0.021	0.053
25305	C>T		Intron 9	0.005	0.000	0.011
25550	G>A	rs3211940	Intron 9	0.037	0.043	0.032
25695	T>C	rs3211941	Intron 9	0.038	0.022	0.053
25849	T>G		**Exon 10 (Non-synonymous)**	0.005	0.000	0.011
25850	A>T		**Exon 10 (Non-synonymous)**	0.016	0.031	0.000
26016_26019	del4		Intron 10	0.011	0.000	0.021
26076	G>A	rs1527483	Intron 10	0.011	0.000	0.021
26089	G>T	rs3211944	Intron 10	0.047	0.042	0.053
26158	T>C		Intron 10	0.016	0.010	0.021
26160	G>T	rs3211945	Intron 10	0.011	0.010	0.011
26262	T>C		Intron 10	0.016	0.010	0.021
26305	G>A	rs3211946	Intron 10	0.011	0.010	0.011
26346	C>T	rs3211947	Intron 10	0.016	0.021	0.011
26446_26449	del4		Intron 10	0.005	0.000	0.011
26484	G>A		Intron 10	0.016	0.010	0.022
26692	C>T		**Exon 11 (Non-synonymous)**	0.005	0.000	0.011
26740	G>C		Intron 11	0.005	0.010	0.000
26776	T>C	rs3211949	Intron 11	0.011	0.010	0.011
26927	A>G	rs3211951	Intron 11	0.016	0.021	0.011
26948_26984	del37	rs3211950	Intron 11	0.011	0.000	0.021
27041	A>G	rs3211953	Intron 11	0.011	0.010	0.011
27163	A>T		Intron 11	0.021	0.031	0.011
27234_27239	del6		Intron 11	0.005	0.000	0.011
27494	A>G		Intron 12	0.005	0.000	0.011
28080_28083	del4		**Exon- 3'UTR**	0.016	0.021	0.011
28278	G>T		**Exon- 3'UTR**	0.016	0.021	0.011
28302	A>G	rs8956	**Exon- 3'UTR**	0.037	0.021	0.053
28314_28329	del16	rs3212018	**Exon- 3'UTR**	0.037	0.042	0.032
28375	T>G	rs3211956	3' flanking	0.011	0.000	0.021
28507_28508	ins4		3' flanking	0.016	0.000	0.032
28509	A>T		3' flanking	0.016	0.000	0.032
28747	C>G		3' flanking	0.011	0.022	0.000
28947	G>A	rs3211959	3' flanking	0.021	0.010	0.032
28969_28974	del6		3' flanking	0.016	0.010	0.021
28971	A>G		3' flanking	0.005	0.000	0.011
29082	G>A		3' flanking	0.043	0.043	0.043
29186_29187	ins4		3' flanking	0.021	0.031	0.011
29281_29945	del665		3' flanking	0.006	0.000	0.011

Table 13 Continued

29344_29347	del4		3' flanking	0.005	0.000	0.011
29706	C>A		3' flanking	0.011	0.000	0.022
29815	G>A		3' flanking	0.011	0.010	0.011
29963	G>A		3' flanking	0.016	0.021	0.011
30040	C>G		3' flanking	0.017	0.023	0.011

3.3 LD AND TAGGER ANALYSES OF *CD36* VARIANTS

SNPs that are lie close to one another along the same chromosome are often inherited together, also known as linkage disequilibrium (LD). These SNPS that are transmitted together can be grouped into haplotypes, and uniquely identifying "tag" SNPs can be used to identify these haplotypes. Tagger analysis can be used to identify these tagSNPs in a group of SNPs, which helps to reduce the number of SNPs needed for genotype screening.

3.3.1 Non-Hispanic Whites

LD and Tagger analysis of the 44 common variants (MAF ≥5%) found in NHWs using an r^2 cutoff of 0.9 identified 8 tagSNP bins. Strong pair wise LD was detected for common variants in NHWs. Table 14 summarizes the Tagger results with variants genotyped in our entire NHW population underlined, and Figure 6 shows the LD plot obtained by analysis.

Table 14. Tagger results using Haploview of CD36 common variants in NHWs

Bin	Variants Included in Bin
1	2521, 2996, 3157, 3991, 4366, 4648, 11890, 12132, 12143_12144del2, 12272_12274del3, 13528_13530del3, 13973, 14455
2	22296, 22749, 23463, 23533, 23689, 24426, 25580, 26512, 27003, 28685, 29225, 29894_29902del9
3	3094, 4993, 7167, 7854, 9473, 9534, 11741, 12919, 13936
4	9794_9799del6, 10876, 13094, 13388
5	4249, 14299
6	9600, 13577
7	19307_19335del29
8	28314_28329del16

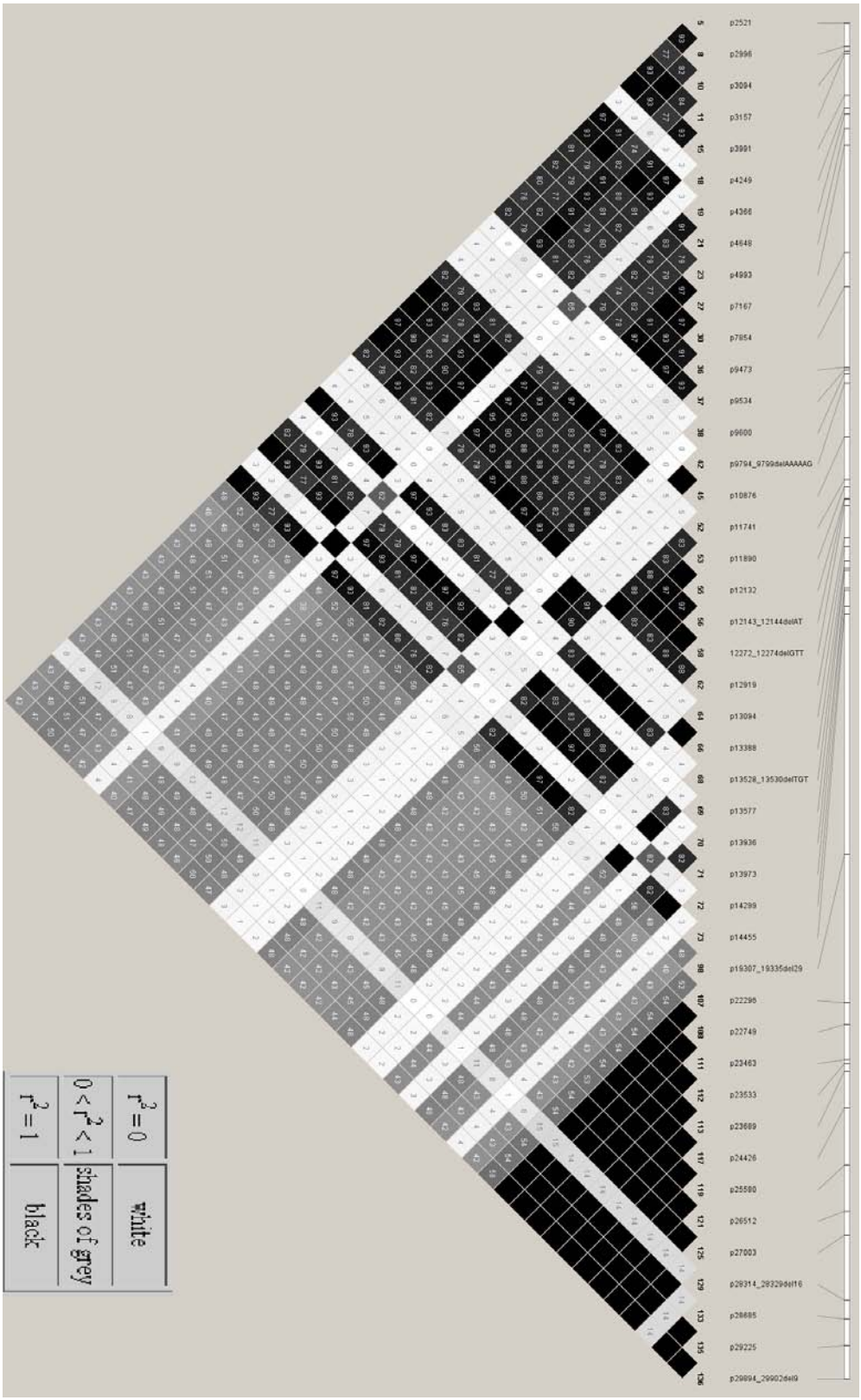


Figure 6. LD analysis for NHWs

3.3.2 Blacks

LD and Tagger analysis of the 103 common variants (MAF≥5%) found in Blacks using an r^2 cutoff of 0.9 identified 57 tagSNP bins. Table 15 summarizes the Tagger results with variants genotyped in our entire Black population underlined, and Figure 7 shows the LD plot obtained by analysis.

Table 15. Tagger results using Haploview of *CD*36 common variants in Blacks

Bin	Variants Included in Bin
1	106, 1021, 1848, 2162, 3412, 7430, 12998, 13758, 14081
2	2389delT, 4990, 7988, 9786delA, 10381, 11554, 11684, 14299
3	947_948del2, 2949_950insA , 2652, 4259, 4879, 14482
4	22296, 22749, 23463, 23689, 24426
5	3991, 12132, 12272_12274del3, 13973
6	9794_9799del6, 13094, 13388
7	18137delA, 18662, 19678
8	8796, 9505, 9474
9	1024, 3350delA, 8595
10	2996, 3157
11	2521, 4366
12	17807_17813del7, 19386
13	13528_13530del3, 14455
14	24080, 25945
15	25580, 29225
16	7167, 13936
17	9152delC, 11137delG
18	12145, 12403
19	3094
20	9473
21	20644
22	15932
23	4993
24	14882, 15554_15557del3, 15975,
25	19151
26	16163
27	7854
28	12919
29	7947
30	28685
31	12143_12144del2
32	25025
33	27003
34	16040
35	4648
36	4249
37	19810
38	3714_3715del2
39	7175
40	19877

Table 15 Continued	
41	16568
42	11741
43	29894_29902del9
44	23533
45	14903
46	26512
47	4134
48	2638
49	22351
50	1547
51	17743
52	11472
53	18911
54	9534
55	9600
56	13577
57	24166

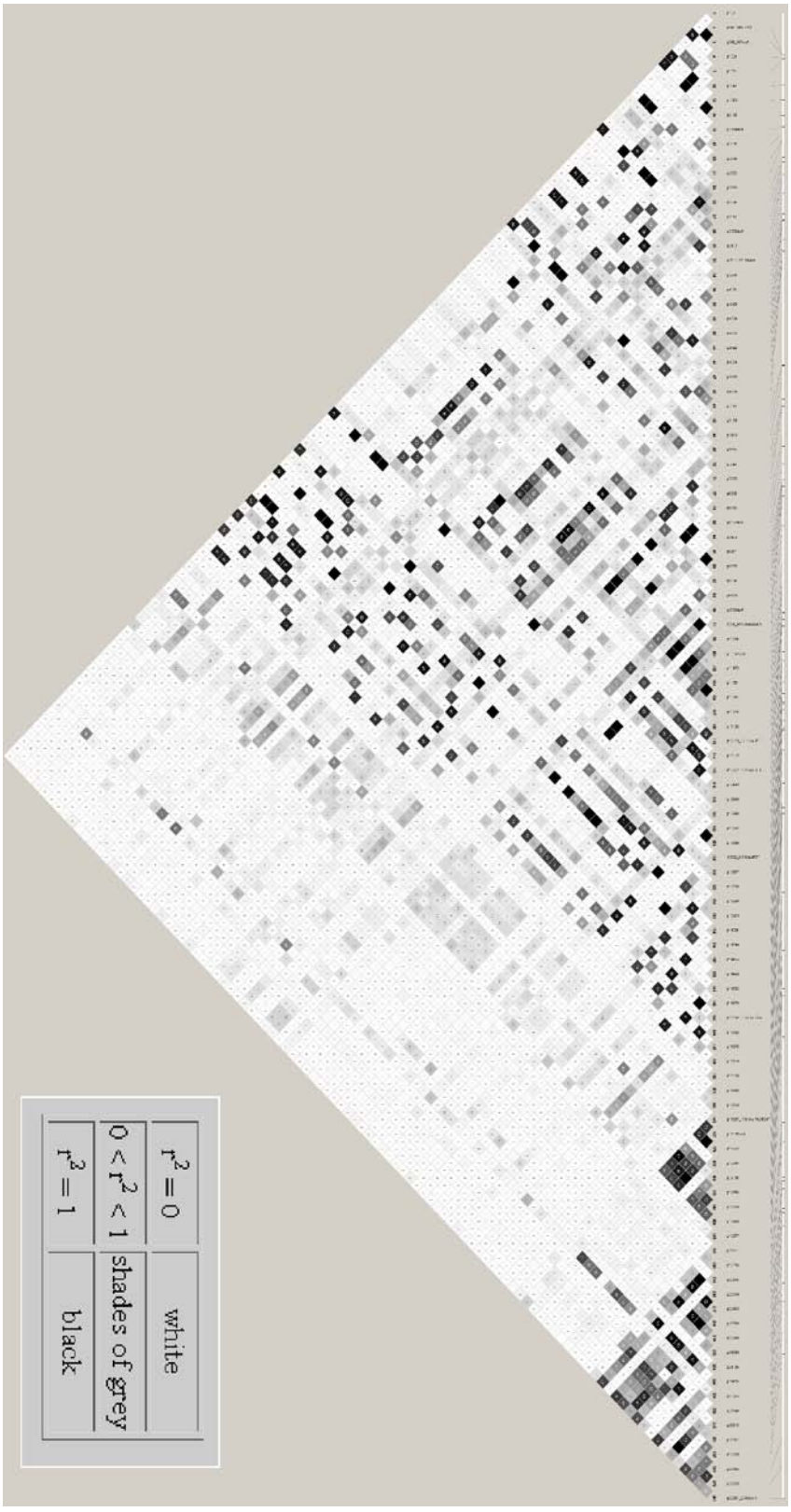


Figure 7. LD analysis for Blacks

3.4 GENOTYPING OF IDENTIFIED VARIANTS IN THE ENTIRE NHW AND BLACK POPULATIONS

3.4.1 Taqman SNP Genotyping

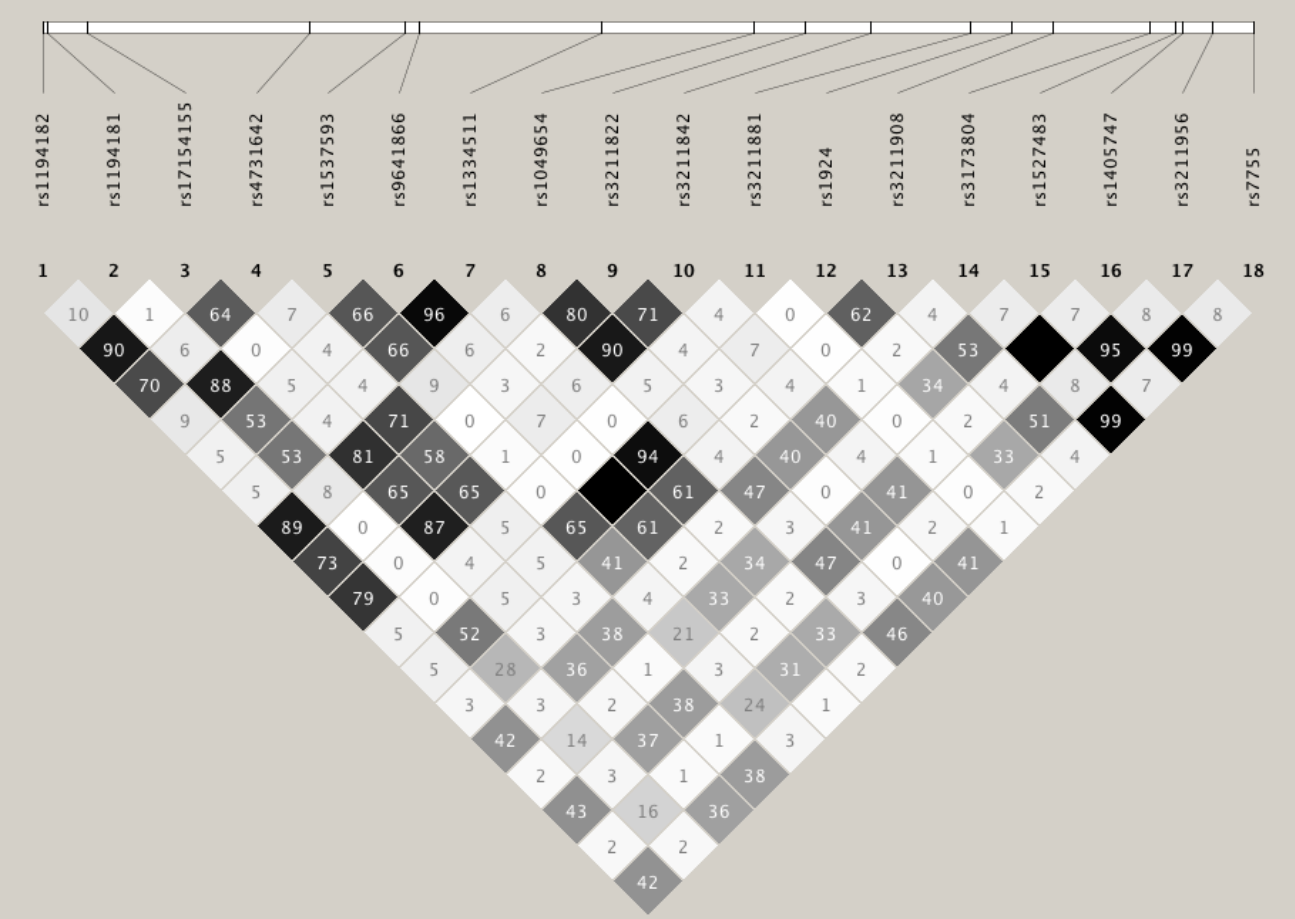
We genotyped 18 variants in Blacks and 18 variants in NHWs using pre-made Taqman SNP Genotyping Assays to date. Genotyping call rates for a total of 19 assays (17 common to both populations) is listed in Table 16**.**

Table 16. Genotyping call rates for TaqMan

***CD36* Reference SNP ID**	**Position**	**Location**	**NWH (%)**	**Black (%)**
rs3211822	3157	Intron	99.36	97.20
rs3211842	7167	Intron	98.24	96.70
rs3211881	13094	Intron	100.00	99.04
rs1924	15932	Intron	99.20	98.34
rs3211908	18463	Intron	99.84	96.57
rs3173804	24426	Intron	99.84	98.48
rs1527483	26076	Intron	98.72	98.73
rs1405747	26512	Intron	99.04	98.60
rs3211956	28375	Intron	98.88	99.11
rs1334511		Intron*	98.72	97.59
rs1537593		Intron*	99.04	97.46
rs9641866		Intron	98.88	96.44
rs1194182		Exon*	98.24	99.23
rs17154155		Intron*	99.04	98.60
rs10499858		Intron*	--	97.84
rs1049654		Exon*	99.36	96.82
rs1194181		Intron*	99.04	98.48
rs4731642		Intron*	99.36	--
rs7755		Exon*	99.04	97.07

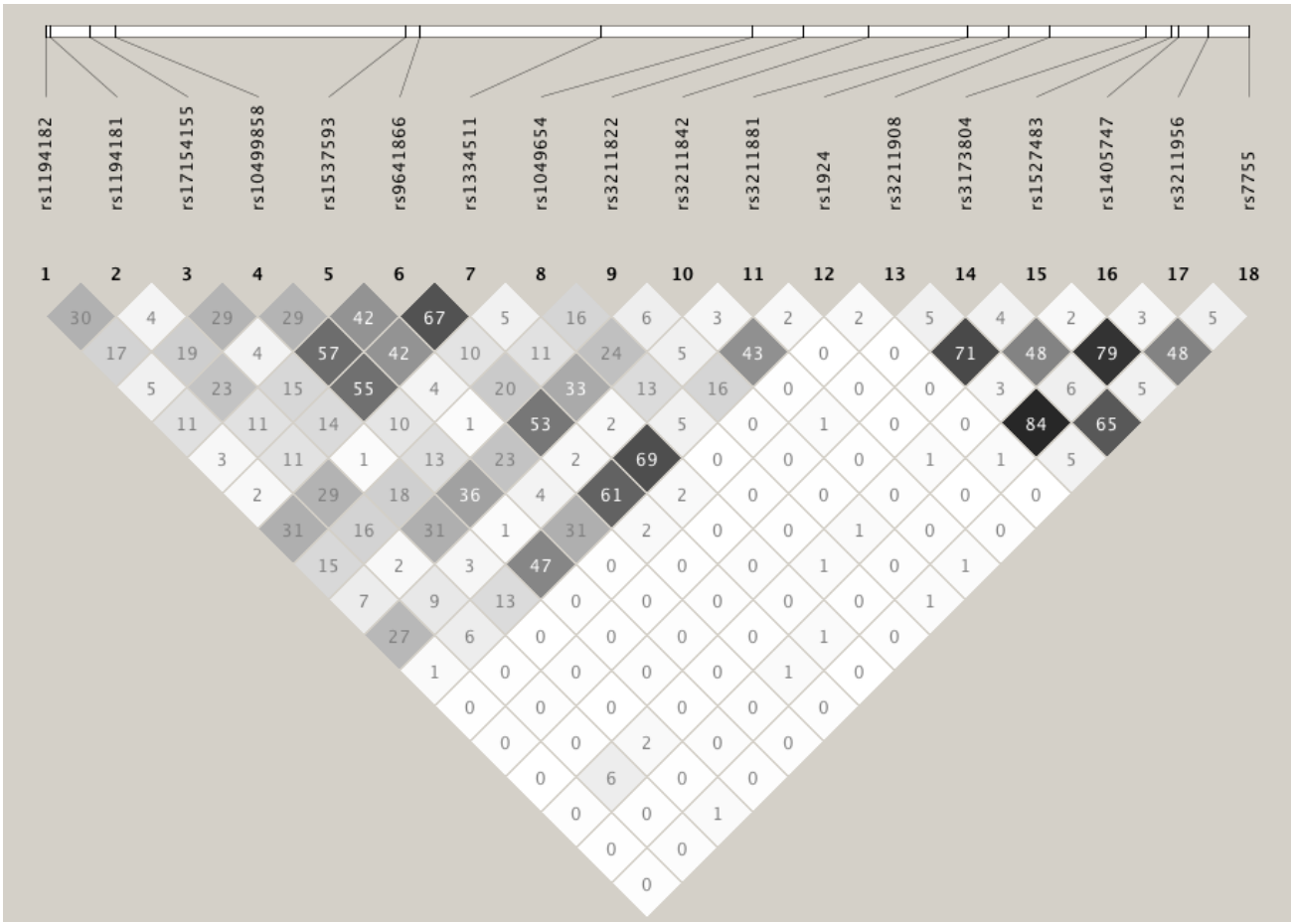
* SNPS falling outside of our sequenced region

None of the variants in the NWH group deviated from HWE (p>0.05). In the Black population, one variant (rs3173804) slightly deviated from HWE (p=0.0212). This SNP was also found to be out of HWE in the SeattleSNPs database for their Caucasian population. For the variants screened using TaqMan assays, the LD analysis was repeated (shown in Figures 8 and 9), and the LD patterns were similar to those observed in the high and low HDL populations used for sequencing.



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

Figure 8. LD analysis of the variants screened in the entire NHW population



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

Figure 9. LD analysis of the variants screened in the entire Black population

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

Tables 17 and 18 show the genotype counts, covariate adjusted mean HDL-C levels (for each genotype), and adjusted p-values for each variant screened in the entire NHW and Black sample populations, respectively. The model used for this analysis was either dominant or additive depending on the MAF. We did not find any p-values that would indicate association between any significant association with HDL-C levels for the variants screened to date in the entire NHW and Black samples. The genotyping of remaining variants is under way using either TaqMan SNP genotyping assays of the Sequenome® iPLEX genotyping platform.

Table 17. Genotype distribution, mean HDL-C levels, and adjusted p-values for 18 *CD36* variants in NHWs

Variant	Males*			Females**			All***		
rs1049654_Add	CC (94)	AC (146)	AA (51)	CC (102)	AC (152)	AA (72)	CC (196)	AC (298)	AA (123)
HDL-C Adjusted Mean ± SE	45.26±1.05	43.19±0.85	44.38±1.43	56.63±1.36	56.8±1.11	57.57±1.61	50.98±0.88	50.03±0.71	50.93±1.11
p-value	0.447			0.7			0.769		
rs1194181_Dom	GG(33)	GA/AA(2)		GG (276)	GA/AA (49)		GG (531)	GA/AA (84)	
HDL-C Adjusted Mean ± SE	43.94±0.64	44.54±1.72		56.41±0.82	59.13±1.95		50.20±0.53	51.94±1.34	
p-value	0.703			0.19			0.226		
rs1194182_Add	GG (89)	CG (150)	CC (48)	GG (100)	CG (155)	CC (68)	GG (189)	CG (305)	CC (116)
HDL-C Adjusted Mean ± SE	45.63±1.08	42.95±0.83	45.49±1.47	56.65±1.37	56.85±1.1	56.47±1.67	51.15±0.89	49.97±0.70	50.72±1.14
p-value	0.627			0.85			0.579		
rs1334511_Dom	AA (263)	AG (27)		AA (288)	AG (35)		AA (551)	AG (62)	
HDL-C Adjusted Mean ± SE	43.83±0.63	45.80±1.96		56.44±0.8	58.3±2.31		50.16±0.52	52.08±1.56	
p-value	0.319			0.41			0.201		
rs1405747_Add	CC (101)	CA (136)	AA (53)	CC (101)	CA (151)	AA (73)	CC (202)	CA (287)	AA (126)
HDL-C Adjusted Mean ± SE	44.09±1.02	43.48±0.88	45.40±1.40	56.45±1.37	56.6±1.12	57.53±1.61	50.23±0.86	50.06±0.73	51.58±1.10
p-value	0.493			0.63			0.365		
rs1527483_Dom	GG (263)	GA (25)		GG (284)	GA (41)		GG (547)	GA (66)	
HDL-C Adjusted Mean ± SE	44.09±0.63	42.34±2.04		56.94±0.81	56.26±2.14		50.50±0.52	49.71±1.51	
p-value	0.528			0.94			0.892		
rs1537593_Dom	CC (255)	CT/TT (36)		CC (273)	CT/TT (51)		CC (528)	CT/TT (87)	
HDL-C Adjusted Mean ± SE	43.93±0.64	44.31±1.70		56.75±0.83	58.1±1.92		50.36±0.53	51.30±1.32	
p-value	0.778			0.47			0.465		
rs17154155_Add	GG (98)	GT (147)	TT (45)	GG (113)	GT (150)	TT (62)	GG (211)	GT (297)	TT (107)
HDL-C Adjusted Mean ± SE	44.88±1.03	43.42±0.85	44.44±1.53	56.95±1.29	56.46±1.12	57.46±1.74	50.94±0.85	50.01±0.71	50.79±1.19
p-value	0.676			0.97			0.689		
rs1924_Dom	GG (263)	GA/AA (27)		GG (291)	GA/AA (35)		GG (554)	GA/AA (62)	
HDL-C Adjusted Mean ± SE	43.84±0.63	45.69±1.96		56.77±0.8	57.77±2.32		50.34±0.52	51.57±1.56	
p-value	0.359			0.67			0.401		

Table 17 Continued

rs3173804_Add	TT (101)	TA (138)	AA (54)	TT (103)	TA (151)	AA (73)	TT (204)	TA (289)	AA (127)
HDL-C Adjusted Mean ± SE	44.11±1.01	43.41±0.87	45.52±1.39	56.58±1.35	56.66±1.12	57.53±1.61	50.31±0.86	50.04±0.72	51.65±1.09
p-value	0.459			0.68			0.363		
rs3211822_Add	GG (111)	GA (140)	AA (42)	GG (118)	GA (153)	AA (53)	GG (229)	GA (293)	AA (95)
HDL-C Adjusted Mean ± SE	45.02±0.97	43.57±0.86	43.03±1.57	56.84±1.26	56.46±1.11	57.88±1.88	50.95±0.81	50.04±0.72	50.53±1.26
p-value	0.215			0.8			0.501		
rs3211842_Add	GG (100)	GA (138)	AA (48)	GG (113)	GA (146)	AA (65)	GG (213)	GA (284)	AA (113)
HDL-C Adjusted Mean ± SE	44.81±1.03	43.93±0.88	43.24±1.48	56.35±1.29	56.65±1.13	58.2±1.7	50.61±0.85	50.30±0.73	50.87±1.16
p-value	0.398			0.43			0.993		
rs3211881_Dom	AA (256)	AG/GG (37)		AA (293)	AG/GG (35)		AA (549)	AG/GG (72)	
HDL-C Adjusted Mean ± SE	43.65±0.63	46.75±1.67		57.27±0.79	53.19±2.31		50.52±0.52	50.04±1.45	
p-value	0.19			0.14			0.875		
rs3211908_Dom	CC (277)	CT (16)		CC (300)	CT (27)		CC (577)	CT (43)	
HDL-C Adjusted Mean ± SE	44.05±0.61	43.90±2.55		56.87±0.79	56.09±2.63		50.49±0.51	49.86±1.88	
p-value	0.92			0.89			0.933		
rs3211956_Dom	TT (264)	TG/GG (26)		TT (280)	TG/GG (44)		TT (544)	TG/GG (68)	
HDL-C Adjusted Mean ± SE	44.17±0.63	42.88±2.00		57.06±0.82	55.71±2.07		50.60±0.53	49.60±1.47	
p-value	0.677			0.82			0.804		
rs4731642_Add	AA (77)	AG (142)	GG (72)	AA (93)	AG (163)	GG (70)	AA (170)	AG (305)	GG (142)
HDL-C Adjusted Mean ± SE	44.65±1.16	43.44±0.86	44.89±1.20	56.69±1.42	56.37±1.07	57.82±1.64	50.62±0.94	49.90±0.70	51.48±1.03
p-value	0.809			0.65			0.559		
rs7755_Add	GG (98)	GA (138)	AA (54)	GG (102)	GA (151)	AA (72)	GG (200)	GA (289)	AA (126)
HDL-C Adjusted Mean ± SE	44.15±1.03	43.48±0.87	45.53±1.39	56.67±1.36	56.83±1.12	57.4±1.62	50.37±0.87	50.16±0.72	51.58±1.10
p-value	0.457			0.74			0.408		
rs9641866_Dom	TT (262)	TA/AA (27)		TT (291)	TA/AA (33)		TT (553)	TA/AA (60)	
HDL-C Adjusted Mean ± SE	43.81±0.63	44.90±1.96		56.74±0.8	58.27±2.39		50.31±0.52	51.56±1.59	
p-value	0.59			0.52			0.418		

* Mean and *p*-values adjusted for BMI
** Mean and *p*-values adjusted for BMI, age, and smoking
***Mean and *p*-values adjusted for BMI, age, smoking, and sex

Table 18. Genotype distribution, mean HDL-C levels, and adjusted p-values for 18 *CD36* variants in Blacks

Variant	Males*			Females**			All***		
rs1049654_Add	AA (208)	AC (202)	CC (51)	AA (96)	AC (140)	CC (41)	AA (304)	AC (341)	CC (92)
HDL-C Adjusted Mean ± SD	47.20±0.84	45.43±0.85	46.00±1.70	50.68±1.26	52.02±1.04	48.18±1.94	48.59±0.70	48.05±0.66	46.38±1.28
p-value	0.2			0.5176			0.127		
rs10499858_Add	AA (304)	AG (151)	GG (13)	AA (188)	AG (76)	GG (14)	AA (491)	AG (227)	GG (27)
HDL-C Adjusted Mean ± SD	46.34±0.69	45.38±0.99	45.18±3.37	51.41±0.91	50.79±1.44	47.60±3.34	48.26±0.55	47.26±0.82	46.52±2.37
p-value	0.46			0.4037			0.3066		
rs1194181_Add	AA (153)	AG (224)	GG (95)	AA (86)	AG (130)	GG (62)	AA (239)	AG (354)	GG (156)
HDL-C Adjusted Mean ± SD	46.50±0.99	45.56±0.81	47.00±1.25	50.32±1.33	52.20±1.08	50.46±1.57	48.00±0.80	48.02±0.65	48.20±0.98
p-value	0.99			0.8113			0.949		
rs1194182_Add	CC (313)	CG (149)	GG (13)	CC (174)	CG (94)	GG (13)	CC (487)	CG (242)	GG (26)
HDL-C Adjusted Mean ± SD	46.44±0.69	45.44±1.00	45.69±3.38	51.44±0.95	50.60±1.30	52.21±3.48	48.27±0.56	47.46±0.79	48.27±2.42
p-value	0.35			0.8519			0.462583		
rs1334511_Add	AA (262)	AG (182)	GG (22)	AA (168)	AG (92)	GG (17)	AA (429)	AG (274)	GG (39)
HDL-C Adjusted Mean ± SD	46.73±0.75	44.93±0.90	48.94±2.59	52.13±0.97	49.89±1.31	51.78±3.06	48.76±0.60	46.80±0.75	49.90±2.00
p-value	0.51			0.3506			0.277		
rs1405747_Add	CC (315)	CA (141)	AA (14)	CC (179)	CA (94)	AA (8)	CC (493)	CA (235)	AA (22)
HDL-C Adjusted Mean ± SD	46.70±0.68	44.71±1.02	49.31±3.23	51.24±0.94	51.06±1.30	49.70±4.45	48.39±0.55	47.17±0.80	49.67±2.62
p-value	0.33			0.7915			0.4023		
rs1527483_Dom	GG (464)	GA (9)		GG (272)	GA (7)		GG (735)	GA (16)	
HDL-C Adjusted Mean ± SD	46.19±0.56	44.56±4.05		51.34±0.74	48.61±4.63		48.11±0.45	46.14±3.05	
p-value	0.64			0.5346			0.4876		
rs1537593_Add	CC (170)	CT (241)	TT (53)	CC (104)	CT (133)	TT (42)	CC (273)	CT (374)	TT (95)
HDL-C Adjusted Mean ± SD	46.78±0.93	45.71±0.78	44.78±1.67	52.34±1.22	50.22±1.08	52.45±1.92	48.84±0.74	47.46±0.64	47.85±1.27
p-value	0.28			0.685			0.3076		
rs17154155_Add	GG (141)	GT (250)	TT (80)	GG (96)	GT (128)	TT (56)	GG (237)	GT (377)	TT (136)
HDL-C Adjusted Mean ± SD	47.09±1.02	45.13±0.77	47.87±1.35	50.77±1.29	51.80±1.11	49.70±1.68	48.59±0.80	47.46±0.63	48.47±1.06
p-value	0.97			0.7829			0.7825		

Table 18 Continued

rs1924_Add	GG (264)	GA (173)	AA (30)	GG (168)	GA (96)	AA (19)	GG (431)	GA (269)	AA (49)
HDL-C Adjusted Mean ± SD	46.26±0.75	45.57±0.93	47.21±2.22	51.95±0.97	49.87±1.28	50.43±2.90	48.41±0.59	47.27±0.75	48.35±1.77
p-value	0.96			0.243			0.4786		
rs3173804_Add	TT (394)	TA (68)	AA (10)	TT (206)	TA (65)	AA (7)	TT (599)	TA (133)	AA (17)
HDL-C Adjusted Mean ± SD	46.51±0.61	44.04±1.48	45.49±3.85	51.60±0.86	49.96±1.54	60.43±4.68	48.39±0.50	46.36±1.07	52.22±2.98
p-value	0.16			0.8192			0.4913		
rs3211822_Add	GG (141)	GA (233)	AA (88)	GG (102)	GA (130)	AA (46)	GG (243)	GA (362)	AA (134)
HDL-C Adjusted Mean ± SD	46.60±1.02	44.54±0.79	48.41±1.29	50.61±1.25	51.66±1.10	51.05±1.86	48.03±0.79	47.22±0.65	49.51±1.06
p-value	0.46			0.731			0.4044		
rs3211842_Add	GG (214)	GA (204)	AA (42)	GG (131)	GA (123)	AA (22)	GG (345)	GA (326)	AA (64)
HDL-C Adjusted Mean ± SD	45.93±0.82	45.90±0.84	46.92±1.86	51.08±1.10	51.44±1.13	48.38±2.67	47.85±0.66	48.00±0.68	47.70±1.53
p-value	0.59			0.6202			0.823199		
rs3211881_Dom	AA (411)	AG/GG (62)		AA (236)	AG/GG (47)		AA (647)	AG/GG (108)	
HDL-C Adjusted Mean ± SD	46.29±0.60	44.95±1.54		51.25±0.81	50.63±1.83		48.11±0.48	47.40±1.18	
p-value	0.36			0.8942			0.6092		
rs3211908_Dom	CC (453)	CT (6)		CC (270)	CT (7)		CC (722)	CT (13)	
HDL-C Adjusted Mean ± SD	46.15±0.57	41.23±4.98		51.38±0.76	47.73±4.74		48.13±0.46	43.75±3.42	
p-value	0.28			0.4576			0.188402		
rs3211956_Dom	TT (467)	TG (6)		TT (276)	TG (6)		TT (742)	TG (12)	
HDL-C Adjusted Mean ± SD	46.14±0.56	41.21±4.96		51.31±0.75	46.04±5.11		48.08±0.45	43.15±3.55	
p-value	0.27			0.2993			0.1484		
rs7755_Add	GG (374)	GA (87)	AA (6)	GG (192)	GA (75)	AA (6)	GG (565)	GA (162)	AA (12)
HDL-C Adjusted Mean ± SD	46.46±0.63	44.65±1.31	48.09±4.99	51.59±0.90	49.92±1.45	55.25±5.08	48.35±0.52	46.66±0.97	51.17±3.55
p-value	0.35			0.6729			0.3677		
rs9641866_Add	TT (263)	TA (178)	AA (23)	TT (154)	TA (101)	AA (16)	TT (416)	TA (279)	AA (39)
HDL-C Adjusted Mean ± SD	47.12±0.74	44.02±0.90	47.43±2.50	51.05±0.98	52.38±1.21	45.52±3.01	48.60±0.59	47.09±0.73	46.36±1.94
p-value	0.1			0.5947			0.0947		

* Mean and *p*-values adjusted for waist measurement
** Mean and *p*-values adjusted for waist measurement, smoking, and exercise
*** Mean and *p*-values adjusted for waist, smoking, exercise, and sex

4.0 DISCUSSION

CD36 is a class B scavenger receptor which is known to act as a receptor for FA, LDL-C, HDL-C, and VLDL. Mutations in the *CD36* gene have been found to be associated with platelet glycoprotein IV deficiency, which is an autosomal recessive condition. Patients with this deficiency have been found to have an altered lipid profile, with elevated LDL-C and TG with but reduced HDL-C levels (Miyaoka et al., 2001 and Yanai et al., 2000). A list of the seven mutations reported to date in *CD36* causing the protein deficiency is located in Online Mendelian Inheritance in Man (http://www.ncbi.nlm.nih.gov/omim), one of which (rs3211938, T>G) was identified in this study.

In addition to this association with a Mendelian disorder, further studies are needed to determine the associations of rare and common variants in *CD36* with complex genetic traits/diseases. Some studies have hypothesized atheroprotective roles for *CD36 (*Nagy et al., 1998; Tontonoz et al., 1998; Masuda and Ross, 1990; Endemenn et al., 1993*)*, however, other studies have implicated *CD36* in alterations in the lipid profile that may increase risk for CVD (Ma et al., 2004). To date, GWA studies have not identified *CD36* as a major gene associated with HDL levels. However, several genomewide linkage scans have linked a region of chromosome 7 (7q11.2–7q21.11) with components of the Metabolic Syndrome, including insulin resistance and dyslipidemia (Love-Gregory et al., 2008). This region harbors the *CD36* gene, which has prompted researchers to further investigate the relationship of common SNPs and

haplotypes with the lipid profile and other components of MetS (Love-Gregory et al., 2008; Madden et al., 2008; Goyenechea et al., 2008). The rare variant hypothesis has not been tested to our knowledge, and the only study that has sequenced the *CD36* gene did not investigate genotype-phenotype correlations (Gelhaus et al., 2001). In order to fully understand the correlation of *CD36* with the lipid profile, HDL-C in particular, deep sequencing of this gene is the next step to clarify its association with HDL-C levels and expand the catalogue of variants known to be responsible for variation in lipid levels.

In our study, we evaluated the role of *CD36* genetic variation by resequencing the entire gene (covering all coding exons and their flanking regions as well as introns) in a subset of samples from healthy individuals with HDL-C levels in the upper and lower 5^{th} percentiles. The portion of the CD36 gene sequenced in our study was ~30kb in length. There is an alternative splice-form that was not sequenced in this study, which includes additional non-coding exons and alternative introns. This alternative splice form spans ~79kb, which is ~49kb longer than the portion that we screened. The purpose was to catalogue the rare and common variation in this gene by using those individuals with extreme HDL-C levels, followed by comprehensive analysis of identified variants to test both the "common variant-common disease" and "rare variant-common disease" hypotheses.

4.1 COMPARISON OF OUR STUDY RESULTS WITH SEATTLESNPS DATABASE

We compared our sequencing data with one published study (Gelhaus et al., 2001) along with the publicly available SeattleSNPs data. Gelhaus et al. (2001) resequenced the promoter region up to -253bp relative to the transcriptional start site and all 15 exons in 12 individuals from Ghana,

who were selected by their spleen size and hypothesized susceptibility to malaria. Gelhaus et al. (2001) identified 24 variants, with 5 of those variants affecting the amino acid sequence. Among these 5 variants they found to affect amino acid sequence, only 1 (1264) was also detected in our study.

SeattleSNPs also completely sequenced the *CD36* gene in 24 and 23 randomly selected African-American and European adults, respectively that were unselected for HDL-C levels. A total of 95 African adults and 95 American NHWs, selected based on their extreme HDL-C levels (in the upper and lower 5^{th} percentiles), were sequenced in this study. We noted several differences between the data we obtained through resequencing and the information available in the SeattleSNPs database which are likely due to differences in sample selection criteria and sample sizes. There is a greater likelihood of admixture with other ethnic groups in African American samples (used by SeattleSNPs) when compared to African samples (used in our study), which may be responsible for some sequence differences. Any variants identified in our study but not reported in the SeattleSNPs database may be a result of our larger sample size and selection factors, or they may be a result of their association with HDL-C levels.

All variant and exon locations are given using SeattleSNPs database nomenclature. A total of 187 variants were reported in the SeattleSNPs database, with 115 have a MAF<5% and 72 having a MAF≥5%. A total of 343 variants were identified in this study, however we do not have combined population frequency data so we cannot break down the number of variants with a MAF<5% and a MAF≥5% for our combined populations. A total of 80 sequence variants were reported in the SeattleSNPs database for the European population (19 with MAF<5% and 61 with MAF≥5%). By comparison, a total of 131 sequence variants were identified in this study in NHWs (87 with MAF<5% and 44 with MAF≥5%) . A total of 167 sequence variants were

reported in the SeattleSNPs database for the African-American population (69 with MAF<5% and 98 with MAF≥5%) , and a total of 281 variants were identified in this study (178 with MAF<5% and 103 with MAF≥5%). Aside from the variants at position 20028 and 27418 (one of which we believe may be a sequencing artifact in the SeattleSNPs sequence), we found all common variants reported by the SeattleSNPs database for the African-American and European populations. However, several rare variants reported in their study were not identified in our study and vice versa, and we have several suspicious variants identified in our study but not present in the SeattleSNPs database that need to be confirmed with other methods. We also found some common variants that were not reported by SeattleSNPs, which may be due to a sequencing gap from 19590-19849 that we did not have.

Five variants affected the coding sequence were reported in the SeattleSNPs database data, including 10423del2 (frameshift, exon 3), 10425delG (frameshift, exon 3), 16983G>A (proline→proline, exon 5), 25025T>G (tyrosine→stop, exon 9), and 25054insAAG (premature protein truncation, exon 9) of which 10423del2 and 25025insT>G were present only in the African American population, 16983G>A and 25054insAAG were present only in the European population, and 10425delG was present in both the African American and European populations. By comparison, a total of 12 coding variants were identified in our study, including 10423del2 (frameshift, exon 3), 10465G>A (tryptophan→stop), 14701del592 (deletion of exon 4), 16983G>A (proline→proline, exon 5), 16986C>A (tyrosine→stop, exon 5), 25025T>G (tyrosine→stop, exon 9), 25048G>T (cysteine→phenylalanine, exon 9), 25849T>G (tyrosine→aspartic acid, exon 10), 25850A>T (tyrosine→phenylalanine, exon 10), 26669G>T (glycine→valine, exon 11), 26692C>T (arginine→tryptophan, exon 11), and 27309T>C (tyrosine→histidine, exon 12) of which 16983, 26669 and 27309 were present only in NHWs

and 10423, 10465, 14701del592, 16986, 25025, 25048, 25849, 25850, and 26692 were present only in Blacks. Nine of the coding variants identified in our study were non-synonymous, one was frameshift, and one resulted in the total deletion of exon 4. Of the five variants affecting the coding sequence reported in the SeattleSNPs database, we identified three in our study (10423del2, 16983G>A, and 25025T>G).

Altogether, we identified a total of 168 variants in our sequencing sample that were previously unreported in the SeattleSNPs database; 53 were present in NHWs and 138 were present in Blacks (listed in Tables 8 and 9 in sections 3.1.1 and 3.1.2, respectively). The identification of these variants not reported in the SeattleSNPs database could be due to our larger sample size allowing a higher detection rate, or it could be due to their association with extreme HDL-C levels (these variants may be unique to the extreme HDL-C groups).

The SeattleSNPs database reported 20 variants unique to their European population and 107 variants unique to their African-American population, while we identified 61 variants unique to our NHW population and 212 variants unique to our African population. As in the SeattleSNPs database, our study also identified a higher number of sequence variants in Blacks versus NHWs for the *CD36* gene.

We also identified a large number of indels in the *CD36* gene in our sequencing sample. Levy et al. (2007) sequenced the human genome and reported ~290,000 indels spanning 3 billion bases, while Sjodin et al (2010) found 3,850 indels spanning 20.3Mb. Based on these numbers, we would expect to find ~3-5.6 indels located in the 30kb that we sequenced in our study. However, we found 46 indels, which is much higher than expected, indicating that *CD36* is a gene enriched with indel polymorphisms. In addition to indels, we also identified two exonic stop codon polymorphisms: 10465 (tryptophan→stop) and 25025 (tyrosine→stop). Both were

found only in our Black population, and 25025 was relatively common with a MAF of 0.226. 250250 (rs3211938) is also one of the seven variants reported in OMIN to cause platelet glycoprotein IV deficiency. Aitman et al. (2000) reported this mutation as one of the most common in the African population he studied. Fry et al. (2009) have argued that this SNP might have recently undergone positive selection in certain African populations, possibly due to an association of the minor allele (G) with less severe malaria. It is intriguing that deleterious variants that interrupt the transcription of *CD36*, such as the rs3211938 variant identified in our Black population, do not show a severe phenotype and undergo positive selection, possibly due to some advantage caused by resistance to certain infections such as malaria.

4.2 DISTRIBUTION OF *CD36* VARIANTS IN HIGH AND LOW HDL-CHOLESTEROL GROUPS

We performed a preliminary analysis of the variants identified by resequencing the *CD36* gene. In NHWs, we observed 21 relatively uncommon or rare variants (MAF<5%) present only in the low HDL-C group, and 25 relatively uncommon or rare variants present only in the high HDL-C group. Eighty-six variants were present in both high HDL-C and low HDL-C groups. For NHWs: 14 out of 47 (29.8%) individuals with high HDL-C had more than two rare variants versus 16 out of 48 (33.3%) individuals with low HDL-C; and 11 out of 47 (23.4%) individuals with high HDL-C had more than three rare variants versus 6 out of 48 (12.5%) individuals with low HDL-C:, and 8 out of 47 (17.0%) individuals with high HDL-C had more than four rare variants versus 5 out of 48 (10.4%) individuals with low HDL-C. In Blacks, we observed 59 relatively uncommon or rare variants (MAF<5%) present only in the low HDL-C group, and 32

relatively uncommon or rare variants present only in the high HDL-C groups. One hundred and ninety variants were present in both high HDL-C and low HDL-C groups For Blacks: 32 out of 48 (66.7%) individuals with high HDL-C had more than two rare variants versus 34 out of 47 (72.3%) individuals with low HDL-C; 28 out of 48 (58.3%) individuals with high HDL-C had more than three rare variants versus 30 out of 47 (63.8%) individuals with low HDL-C; 22 out of 48(45.8%) individuals with high HDL-C had more than four rare variants versus 24 out of 47 (51.1%) individuals with low HDL-C. We observed that in NHWs, the differences in percentage of rare variants in high and low HDL-C groups increased as the variant cut-off increased, with there being more individuals in the high HDL-C group with rare variants when compared to individuals in the low HDL-C. This does not seem suggestive of the "rare variant" hypothesis proposed by Cohen et al. (2004), which suggests that an excess of sequence variants in subjects with low HDL-C reflects an accumulation of damaging alleles in this group. We also did not observe any differences in the number of rare variants in individuals with high HDL-C versus low HDL-C for our Black population.

Statistically significant differences in MAF between the low and high HDL-C groups were observed for some common variants (MAF≥5%) in the sequencing data from a small sample, but have not been confirmed yet by genotyping in the entire NHW and Black populations. These variants include: 4249C>T/rs3211830 (increase, p=0.047), 13094T>A/rs3211879 (decrease, p=0.040), and 14299A>G/rs3173799 (increase, p=0.047) in NHWs; 16568A>G/rs3211899 (increase, p=0.022) in Blacks. In the NHW population, 4249C>T and 14299A>G were found to be linked to one another in LD analysis, while 13094 was located in a different bin. Out of these four variants that demonstrated significant differences in MAF between the low and high HDL-C groups, none have been reported in the literature as being

studied for an association with HDL-C. Thirteen of the common variants had a *p*-value of between 5-10%, which may be statistically significant due to the small sample size.

We used our sequencing data to compare our results with several SNPs that were reported in to be associated with HDL-C levels in the literature but not yet genotyped in our entire population: rs3211810, rs3211849, rs1054516, rs3173798, rs3211868, rs3211870, rs1358337, rs3211938, and rs3211913 (Love-Gregory et al., 2008). Table 19 is a summary of those SNPs whose minor allele is reported to be associated with increase in HDL-C levels. All minor allele associations are in the context of the minor allele in our study. Even though none of the p-values we obtained in our sequencing samples were significant, the MAFs obtained in three of our sequencing samples supported the associations reported by Love-Gregory et al. (2008): rs3211849, rs1358337, and rs3211913. The directions of the reported associations in Love-Gregory et al. (2008) along with the MAF obtained for both populations in our sequencing sample are listed in Table 19.

Table 19. Variants in literature reported to have an association with HDL-C (Love-Gregory et al., 2008)

SNP	Reported Direction of Association With the minor allele	Blacks in this study		Whites in this study	
		MAF High HDL-C	MAF Low HDL-C	MAF High HDL-C	MAF Low HDL-C
rs3211849	Decrease p=0.029	0.489	0.489	0.415	0.469
rs1358337	Decrease p=0.00066	0.447	0.479	0.415	0.458
rs1054516	Increase p=0.003	0.250	0.307	0.420	0.479
rs3211913	Increase p=0.039	0.545	0.446	0.01	0.00
rs3211810	Decrease p=0.044	0.156	0.138	--	--
rs3211870	Decrease p=0.0079	0.402	0.340	0.415	0.546
rs3211938	Increase p=0.00018	0.234	0.217	--	--
rs3173798	Decrease p=0.033	0.208	0.17	0.074	0.021
rs3211868	Decrease p=0.011	0.208	0.17	0.074	0.021

4.3 COMPARISON OF OUR STUDY RESULTS WITH PUBLISHED LITERATURE

So far we have screened a total of 19 variants have been screened in the entire NHW (n=623) and Black (n=788) populations with TaqMan SNP genotyping assays (Table 6). No association with fasting HDL-C levels was observed either the NHW or Black samples for these 19 variants. Nine out of these 19 variants genotyped in our study were investigated in the literature including: rs1334511, rs1049654, rs3211822, rs3211842, rs3173804, rs7755, rs1924, rs1527483, rs1537593 (Ma et al., 2004; Love-Gregory et al., 2008). The literature reported no association of SNPs rs1924, rs1527483, rs1537593, rs1334511, rs3211822, rs3173804, and rs7755 with HDL-C levels, and our study confirmed these findings. Reports of the association of rs1049654 with HDL-C levels were inconsistent, with Ma et al (2004) reporting no association with HDL-C in Caucasians, and Love-Gregory (2008) reporting an increase in HDL-C levels associated with the minor allele (C) in African-Americans. Our results were consistent with Ma et al. (2004) for rs1049654, and did not support the association reported by Love-Gregory et al. (2008). Love-Gregory also reported that rs3211842 was associated with a decrease in HDL-C levels, but our data did not support that conclusion (2008). This comparison of SNPs genotyped in our entire samples so far with the published literature is shown in Table 20, indicating whether the study found a statistically significant association with an increase in HDL-C, a statistically significant association with a decrease in HDL-C, or no statistically significant association with HDL-C levels and the SNPs of interest.

Table 20. Comparison of our TaqMan genotyping results with published literature.

SNP ID	Study		
	Ma et al. (2004)	**Love-Gregory et al. (2008)**	**Our Study**
	585 Caucasians, male and female	**2020 African Americans: 737 males, 1283 females**	**95 NHW, 95 Black**
rs1049654	No association	Minor allele with high HDL-C (p= 0.0028)	No association
rs1924	No association	--	No Association
rs7755	No association	No association	No Association
rs1527483	No association	--	No Association
rs1537593	No association	--	No Association
rs3211822	--	No Association	No Association
rs3173804	--	No Association	No Association
rs1334511	--	No Association	No Association
rs3211842	--	Minor allele with low HDL-C (p= .00074)	No Association

There are several SNPs that have been reported as being associated with HDL-C levels in the literature that we haven't genotyped and do not have sequencing data because they fall in the region that is ~44 kb 5' further upstream (*) or ~3kb 3' further downstream (**) from our targeted region for sequencing, corresponding to the largest *CD36* sequence available in Genbank (NC_000007.13) that harbors additional alternative noncoding exons. These SNPs are: rs10499859, rs9784998, 13438282, rs3211909, and rs13246513.

4.4 CONCLUSIONS AND FURTHER DIRECTIONS

Heart disease is a major public health concern, with over 16 million people living with CHD, and 80 million people living with CVD. One risk factor for these health problems is decreased HDL-C levels. Although the reported GWA studies have not implicated *CD36* as a main candidate

gene affected HDL-C levels, recent studies have implicated the locus on chromosome 7 harboring *CD36* with components of MetS, including HDL-C. Few studies have been done looking at common SNPs and their association with HDL-C levels and the data from these studies has been inconsistent. This lack of consistency prompts further studies to elucidate the role of *CD36*. Also, to date there have been no efforts to deeply sequence the *CD36* gene to identify both common and rare variants. We undertook this study in order to investigate the reported associations of *CD36* with HDL-C levels, as well as develop a comprehensive catalogue of *CD36* variation through deep sequencing.

We recently completed our sequencing, and we'll be using this data to complete screening of all common TagSNPs and validate the rare variants that we have identified. Completion of this comprehensive screening will reveal the extent to which common and rare variants in *CD36* contribute to HDL-C regulation in NHWs and Blacks. However, final determination of causative effects will involve performing functional studies.

Preliminary analysis of our data did not confirm any previous associations with *CD36* and HDL-C levels through TaqMan genotyping analysis; however, our sequencing data did support one association with a marginally significant p-value. We identified several common variants with significant MAF differences between our high and low HDL-C levels, and additional genotyping of these variants in our entire populations needs to be completed in order to determine the validity of these associations. Confirmation of these findings is necessary by other groups because we are the first to undergo such a comprehensive study. Additional studies of *CD36* with larger population sizes are needed as well to analyze variants that may only have a small effect of HDL-C levels, and further studies of rare variation in this gene are required to better understand the genetics of HDL-C cholesterol in relation to the rare allele hypothesis.

BIBLIOGRAPHY

Aitman T, Cooper L, Norsworthy P, Wahid F, Gray J, Curtis B, McKeigue P, Kwiatkowski D, Greenwood B, Snow R, Hill A, Scott J. Malaria susceptibility and CD36 mutation. *Nature*. 2000;405:1015-6.

Alessio M, Ghigo D, Garbarino G, Geuna M, Malavasi F. Analysis of the human *CD36* leucocyte differentiation antigen by means of the monoclonal antibody NL07.*Cell. Immunol.* 1991;137:487-500.

American Heart Association. Heart and Stroke Facts. 2003. www.americanheart.org.

Applied Biosystems. TaqMan® GenotypingMaster Mix Protocol. 2007.

Armesilla A, Vega M. Structural Organization of the Gene for Human *CD36* Glycoprotein. *The Journal of Biological Chemistry*. 1994;269:18985-18991.

Aulchenko YS, Ripatti S, Lindquist I, Boomsma D, Heid IM, Pramstaller PP, et al. Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. *Nat. Genet.* 2009;41:47–55.

Berg K. Genetics of coronary heart disease and its risk factors. *Ciba Found Symp* 1987;130:14-33.

Boes E, Coassin S, Kollerits B, Heid IM, Kronenberg F. Genetic-epidemiological evidence on genes associated with HDL cholesterol levels: A systematic in-depth review. *Experimental Gerontology* 2004;44:136–160.

Bruce C, Chouinard R, Tall AR. Plasma lipid transfer proteins, high density lipoproteins, and reverse cholesterol transport. *Annu. Rev. Nutr*. 1998;18:297–330.

Bunker CH, Ukoli FA, Matthews KA, Kriska AM, Huston SL, Kuller LH. Weight threshold and blood pressure in a lean black population. *Hypertension* 1995;26:616-623.

Bunker GH, Ukoli FA, Okoro FI, Olumu AB, Kriska AM, Huston SL, Markovic N, Kuller LH. Correlates of serum lipids in a lean black population. *Atherosclerosis* 1996;123;215-225.

Calvo D, Gómez-Coronado D, Suárez Y, Lasunción MA, Vega MA. Human *CD36* is a high affinity receptor for the native lipoproteins HDL, LDL, and VLDL. *J Lipid Res*. 1998; 39:777-788.

Castelli WP, Garrison RJ, Wilson PWF, Abbott, Kalousdian S and Kannel WB, Incidence of coronary heart disease and lipoprotein cholesterol levels: the Framingham Study, *JAMA* 1986;256: 2835–2838.

Cavelier C, Lorenzi I, Rohrer L, Von Eckardstein A. Lipid efflux by the ATPbinding cassette transporters ABCA1 and ABCG1. *Biochim. Biophys. Acta*. 2006;1761:655–666.

Chasman DI, Paré G, Zee RYL, Parker AN, Cook NR, Buring JE, et al. Genetic loci associated with plasma concentration of low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, triglycerides, apolipoprotein A1, and apolipoprotein B among 6382 white women in genome-wide analysis with replication. *Circ. Cardiovasc. Genet*. 2008; 1:21–30.

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

Daviet L, Buckland R, Puente Navazo MD, McGregor JL. Identification of an immunodominant functional domain on human *CD36* antigen using human–mouse chimaeric proteins and homologue-replacement mutagenesis. *Biochem. J*. 1995;305**:**221–224.

Deo RC, Reich D, Tandon A, Akylbekova E, Patterson N, Waliszewska A, Kathiresan S, Sarpong D, Taylor HA Jr, Wilson JG. Genetic differences between the determinants of lipid profile phenotypes in African and European Americans: the Jackson Heart Study. *PLoS Genet.* 2009;5:e1000342.

Corpeleijn E, van der Kallen C, Kruijshoop M, Magagnin M, de Bruin T, Feskens E, Saris W, Blaak E. Direct association of a promoter polymorphism in the *CD36*/FAT fatty acid transporter gene with Type 2 diabetes mellitus and insulin resistance. *Diabet. Med*. 2006;23:907–911.

Endemann G, Stanton L, Madden K, Bryant C, White RT, Protter A. *CD36* is a Receptor for Oxidized Low Density Lipoprotein. *The Journal of Biological Chemistry*. 1993; 268:11811-11816.

Febbraio M, Abumrad N, Hajjar D, Sharma K, Cheng W, Pearce S, Silverstein R. A Null Mutation in Murine *CD36* Reveals an Important Role in Fatty Acid and Lipoprotein Metabolism. *J Biol Chem*. 1999; 274:19055-19062.

Febbraio M, Guy E, Silverstein RL. Stem cell transplantation reveals that absence of macrophage *CD36* is protective against atherosclerosis. *Arterioscler Thromb Vasc Biol*. 2004;24:2333–2338.

Febbraio M, Podrez EA, Smith JD, Hajjar DP, Hazen SL, Hoff HF et al. Targeted disruption of the class B scavenger receptor *CD36* protects against atherosclerotic lesion development in mice. *J Clin Invest.* 2000;105:1049–1056.

Fernandez-Ruiz E**,** Armesilla L**,** Sanchez-Madrid F, Vega M. Gene Encoding the Collagen Type I and Thrombospondin Receptor *CD36* Is Located on Chromosome 7q11.2.*Genomics.* 1993;17:759-761.

Fry A, Ghansa A, Small K, Palma A, et al. Positive selection of a CD36 nonsense variant in sub-Saharan Africa, but no association with severe malaria phenotypes. Hum Mol Genet. 2009;18:2683-2692.

Gelhaus A,Scheding A,Browne E,Burchard GD,Horstmann RD. Variability of the *CD36* Gene in West Africa. *Human Mutation*. 2001;18:444-50.

Gordon T, Castelli WP, Hjortland MC, et al. High density lipoprotein as a protective factor against coronary heart disease. The Framingham Study. Am J Med 1977;62:707–14.

Goyenechea E, Collins LJ, Parra D, Liu G, Snieder H, Swaminathan R, Spector TD, Martínez JA, O'Dell SD. CD36 gene promoter polymorphisms are associated with low density lipoprotein-cholesterol in normal twins and after a low-calorie diet in obese subjects.*Twin Res Hum Genet.* 2008;11:621-8.

Greenwalt DE, Lipsky RH, Ockenhouse CF, Ikeda H, Tandon NN, Jamieson GA. Membrane glycoprotein *CD36*: A review of its roles in adherence, signal transduction, and transfusion medicine. *Blood.* 1992;80**:**1105–1115.

Gruarin P, Thorne R, Dorahy D, Burns G, Sitia R, Alessio M. *CD36* Is a Ditopic Glycoprotein with the N-Terminal Domain Implicated in Intracellular Transport. *Biochemical and Biophysical Research Communications*. 2000;275;446-454.

Guy E, Kuchibhotla S, Silverstein RL, Febbraio M. Continued inhibition of atherosclerotic lesion formation in long term western diet fed *CD36*/ apoE8 mice. *Atherosclerosis.* 2007;192:123–130.

Harris MR, Bunker CH, Hamman RF, Sanghera DK, Aston CE, Kamboh MI. 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* 1998; 137:187-195.

Hamman RF, Marshall JA, Baxter J, Kahn LR, Mayer EJ, Orleans M, Murphy JR, Lezotte DC. Methods and prevalence of non-insulin dependent diabetes mellitus in a biethnic Colorado population. The San Luis Valley Diabetes Study. *Am Epidemiol* 1989;129:295-311.

Heid IM, Boes E, Müller A, Kollerits B, Lamina C, Coassin S, et al. A genomewide association analysis of HDL cholesterol in the population-based KORA Study sheds new light on intergenic regions. *Circ Cardiovasc Genet*. 2008;1:10–20.

Kashiwagi H, Honda S, Tomiyama Y, Mizutani H, Take H, Honda Y, Kosugi S, Kanayama Y, Kurata Y, Matsuzawa Y. A novel polymorphism in Glycoprotein IV (Replacement of proline-90 by serine) predominates in subjects with platelet GPIV deficiency. *Thromb Haemostasis*. 1993;69:481-484.

Kashiwagi H, Tomiyama Y, Honda S, Kosugi S, Shiraga M, Nagao N, Sekiguchi S, Kanakura Y, Kurata Y, Matsuzawa Y. Molecular basis of *CD36* deficiency. *J Clin Invest*. 1995;95:1040–1046.

Kashiwagi H, Tomiyama Y, Kosugi S, Shiraga M, Lipsky RH, Kanakura Y, Kurata Y, Matsuzawa Y. Identification of molecular defects in a subject with type I *CD36* deficiency. *Blood.* 1994;83:3545–3552.

Kashiwagi H, Tomiyama Y, Nozaki S, Honda S, Kosugi S, Shiraga M, Nakagawa T, Nagao N, Kanakura Y, Kurata Y, Matsuzawa Y. A single nucleotide insertion in codon 317 of the *CD36* gene leads to *CD36* deficiency. *Arterioscler Thromb Vasc Biol*. 1996;16:1026–1032.

Kathiresan S, Manning AK, Demissie S, D'Agostino RB, Surti A, Guiducci C,et al. A genome-wide association study for blood lipid phenotypes in the Framingham Heart Study. *BMC Med. Genet*. 2007;8(Suppl.1):S17.

Kathiresan S, Melander O, Guiducci C, Surti A, Burtt NP, Rieder MJ, et al. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat Genet*. 2008b;40:189–197.

Kathiresan S, Pirruccello J. Genetics of lipid disorders. Current Opinion in Cardiology. 2010;25:000-000.

Klos KL, Kullo IJ. Genetic determinants of HDL: monogenic disorders and contributions to variation. *Curr. Opin. Cardiol*. 2007;22:344–351.

Kontush A, Chapman MJ. Functionally defective high-density lipoprotein: a new therapeutic target at the crossroads of dyslipidemia, inflammation, and atherosclerosis. *Pharmacol. Rev*. 2006;58:342–374.

Kooner JS, Chambers JC, Guilar-Salinas CA, Hinds DA, Hyde CL, Warnes GR, et al. Genome-wide scan identifies variation in MLXIPL associated with plasma triglycerides. *Nat. Genet*. 2008;40:149–151.

Krauss R. What can the genome tell us about LDL cholesterol? *The Lancet*. 2008;371:450-452.

Leprêtre F, Cheyssac C, Amouyel P, Froguel P, Helbecque N. A promoter polymorphism in *CD36* is associated with an atherogenic lipid profile in a French general population. *Atherosclerosis*. 2004;173:375-377.

Leprêtre F, Linton K, Lacquemant C, Vatin V, Samson C, Dina C, Chikri M, Ali S, Scherer P, Se´ron K, et al. Genetic study of the *CD36* gene in a French diabetic population. *Diabetes Metab*. 2004; 30:459–463.

Leprêtre F, Vasseur F, Vaxillaire M, Scherer P, Ali S, Linton K, Aitman T, Froguel P. A *CD36* nonsense mutation associated with insulin resistance and familial type 2 diabetes. *Hum. Mutat*. 2004; 24;104.

Levy S, Sutton G, Ng PC, Feuk L, Halpern, A, et al. The diploid genome sequence of an individual human. PloS Biol. 2007;5:e254.

Lewington S, Whitlock G, Clarke R, Sherliker P, Emberson J, Halsey J, 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.* 2007;370: 1829–1839.

Lloyd-Jones DM, Adams R, Carnethon M, et al. Heart Disease and Stroke Statistics—2009 Update: A Report From the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation*. 2009;119:e21-e181.

Lloyd-Jones DM, Nam B-H, D'Agostino RB Sr, Levy D, Murabito JM, Wang RJ, Wilson PWF, O'Donnell CJ. Parental cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults: a prospective study of parents and offspring. *JAMA*. 2004;291:2204 –2211.

Ma X, Bacci S, Mlynarski W, Gottardo L, Soccio T, Menzaghi C, Iori E, Lager RA, Shroff AR, Gervino EV, Nesto RW, Johnstone MT, Abumrad NA, Avogaro A, Trischitta V, Doria A. A common haplotype at the *CD36* locus is associated with high free fatty acid levels and increased cardiovascular risk in Caucasians**.** *Human Molecular Genetics*. 2004. Oct 1;13:2197-205.

Madden J, Carrero JJ, Brunner A, Dastur N, Shearman CP, Calder PC, Grimble RF. Polymorphisms in the *CD36* gene modulate the ability of fish oil supplements to lower fasting plasma triacyl glycerol and raise HDL cholesterol concentrations in healthy middle-aged men. *Prostaglandins Leukotrienes and Essential Fatty Acids*. 2008;78:327-35.

Miyaoka K, Kuwasako T, Hirano K, Nozaki S, Yamashita S, Matsuzawa Y. *CD36* deficiency associated with insulin resistance. *Lancet*. 2001 Mar 3;357:686-7.

Masuda J, Ross R. Atherogenesis during low level hypercholesterolemia in the nonhuman primate. II. Fatty streak conversion to fibrous plaque. *Arteriosclerosis*. 1990;10:178-187.

Moore KJ, Kunjathoor VV, Koehn SL, Manning JJ, Tseng AA, Silver JM, et al. Loss of receptor-mediated lipid uptake via scavenger receptor A or *CD36* pathways does not ameliorate atherosclerosis in hyperlipidemic mice. *J Clin Invest.* 2005;115:2192–2201.

Murabito JM, Pencina MJ, Nam BH, D'Agostino RB Sr, Wang TJ, Lloyd-Jones D, Wilson PW, O'Donnell CJ. Sibling cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults. *JAMA.* 2005;294:3117–3123.

Nagy L, Tontonoz P, Alvarez JGA, Evans RM. Oxidized LDL regulates macrophage gene expression through ligand activation of PPARg. *Cell*. 1998;93:229–240.

National Center for Health Statistics (NCHS). *Health, United States, 2007: With Chartbook on Trends in the Health of Americans.* Hyattsville, Md: National Center for Health Statistics; 2007. Available at: http:// www.cdc.gov/nchs/hus.htm.

Nofer JR, van der GM, Tolle M, et al. HDL induces NO-dependent vasorelaxation via the lysophospholipid receptor S1P3. *J Clin Invest*. 2004;113:569–81.

Nozaki S, Kashiwagi H, Yamashita S, Nakagawa T, Kostner B, Tomiyama Y, Nakata A, Ishigami M, Miyagawa J, Kameda-Takemura K, et al. Reduced uptake of oxidized low density lipoproteins in monocyte-derived macrophages from *CD36*-deficient subjects. *J Clin Invest*. 1995;96:1859-65.

Oquendo P, Hundt E, Lawler J, Seed B. *CD36* directly mediates cytoadherence of Plasmodium falciparum parasitized erythrocytes. *Cell* 1989;58:95-101.

Pearce SF, Wu J, Silverstein RL. A carboxyl terminal truncation mutant of *CD36* is secreted and binds thrombospondin: evidence for a single transmembrane domain**.** *Blood.* 1994;84:384-389.

Rewers M, Shetterly SM, Hoag S, Baxter J, Marshall J, Hamman RF. Is the risk of coronary heart disease lower in Hispanics than in non-Hispanic whites? The San Luis Valley Diabetes Study. *Ethnicity Dis* 1993;3:44-54.

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).

Sabatti C, Service SK, Hartikainen AL, Pouta A, Ripatti S, Brodsky J, Jones CG, Zaitlen NA, Varilo T, Kaakinen M, Sovio U, Ruokonen A, Laitinen J, Jakkula E, Coin L, Hoggart C, Collins A, Turunen H, Gabriel S, Elliot P, McCarthy MI, Daly MJ, Järvelin MR, Freimer NB, Peltonen L. Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. *Nat Genet*. 2008;41:35-46.

Shea M, Benjamin E, Dupuis J, Massaro J, Jacques P, D'Agostino R, Ordovas J, O'Donnell C, Dawson-Hughes B, Vasan R, Booth S. Genetic and non-genetic correlates of vitamins K and D. *Eur J Clin Nutr*. 2009 Apr;63(4):458-64.

Sjödin P, Bataillon T, Schierup M. Insertion and Deletion Processes in Recent Human History. *PLoS ONE.* 2010;5:e8650.

Take H, Kashiwagi H, Tomiyama Y, Honda S, Honda Y, Mizutani H, Furubayashi T, Karasuro T, Nishiura T, Kanayama Y, Kurata Y, Matsuzawa Y. Expression of GPIV and Naka antigen on monocytes in Naka-negative subjects whose platelets lack GPIV. *Brit. J. Haematol*. 1993;84:387-391.

Talle M, Rao P, Westberg E, Allegar N, Makowski M, Mittler R, Goldstein G. Patterns of antigenic expression on human monocytes as defined by monoclonal antibodies. *Cell. Immunol*. 1983;78:83-99.

Tao N, Wagner SJ, Lublin DM. *CD36* Is Palmitoylated on Both N- and C-terminal Cytoplasmic Tails. *J. Biol. Chem*. 1996;271:22315-22320.

Thorne RF, Dorahy DJ, Wilkinson RM, Burns GF. The biology of *CD36*. *In* Thrombosis and Vessel Wall (Berndt, M., Eds.). 2000; Harwood Academic Publishers, Chur, Switzerland.

Tontonoz P, Nagy L, Alvarez JGA, Evans RM. PPARg promotes monocyte/macrophage differentiation and uptake of oxidized LDL. *Cell*. 1998;93:241–252.

Tso C, Martinic G, Fan WH, et al. High-density lipoproteins enhance progenitor-mediated endothelium repair in mice. *Arterioscler Thromb Vasc Biol*. 2006;26:1144–9.

Tsompanidi EM, et al. HDL biogenesis and functions: Role of HDL quality and quantity in atherosclerosis. *Atherosclerosis*. 2009. doi:10.1016/j.atherosclerosis.2009.05.034

Von Eckardstein A. Differential diagnosis of familial high density lipoprotein deficiency syndromes. *Atherosclerosis* 2006;186:231–239.

Wallace C, Newhouse SJ, Braund P, Zhang F, Tobin M, Falchi M, et al. Genome-wide association study identifies genes for biomarkers of cardiovascular disease: serum urate and dyslipidemia. *Am. J. Hum. Genet*. 2008;82:139–149.

Willer CJ, Sanna S, Jackson AU, Scuteri A, Bonnycastle LL, Clarke R, et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat. Genet*. 2008;40:161–169.

Yanai H, Chiba H, Morimoto M, Abe K, Fujiwara H, Fuda H, Hui SP, Takahashi Y, Akita H, Jamieson GA, Kobayashi K, Matsuno K. Human *CD36* deficiency is associated with elevation in low-density lipoprotein-cholesterol**.** *American Journal of Medical Genetics*. 2000;93:299-304.

Zannis VI, Kypreos KE, Chroni A, Kardassis D, Zanni EE. In: Loscalzo J, editor. *Molecular mechanisms of atherosclerosis*. New York, NY: Taylor & Francis;2004. p. 111–74.