# LIPOPROTEIN LIPASE GENE SEQUENCING AND PLASMA LIPID PROFILE

by

**Dilek Pirim**

B.S, Istanbul University, Turkey, 2007

Submitted to the Graduate Faculty of

the Department of Human Genetics

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

Dilek Pirim

It was defended on

December 8, 2010

and approved by

F. Yeşim Demirci, M.D., Assistant Professor, Human Genetics
Graduate School of Public Health, University of Pittsburgh

Clareann H. Bunker, Ph.D., Associate Professor, Epidemiology,
Graduate School of Public Health, University of Pittsburgh

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

**LIPOPROTEIN LIPASE GENE SEQUENCING AND PLASMA LIPID PROFILE**

Dilek Pirim, M.S.

University of Pittsburgh, 2010

In the United States, coronary heart disease (CHD) is the most common cause of death and number one killer of American males and females. Several epidemiological studies have identified risk factors for CHD, like low high-density lipoprotein cholesterol (HDL-C), elevated total cholesterol and low-density lipoprotein (LDL) cholesterol, and high triglycerides (TGs), but underlying genetic variations that cause predisposition to these traits still remain unclear. Lipoprotein lipase (*LPL*) is one of the major genes involved in lipid metabolism and its gene sequence variation has already been reported to be associated with the risk of CHD and risk of other complex diseases like dyslipidemia, type 2 diabetes, essential hypertension, and Alzheimer's disease. Unraveling the unknown genetic variation in the *LPL* gene in relation to HDL-C and correlated lipid traits is critically important for public health because identification of genetic markers may lead to promising future public health interventions, like prognostic tools and therapeutic approaches to alleviate the burden of CHD in the U.S. In this study, we investigated the role of common and rare variation in *LPL* by resequencing individuals having extremely low (n=48) and high (n=47) HDL-C levels selected from a population-based non-Hispanic white (NHW) sample of 623 individuals. A total of 179 variants were identified in 95 individuals by resequencing the entire *LPL* gene, including 91 uncommon or rare variants [minor allele frequency (MAF) <0.05)] and 88 common variants (MAF ≥0.05). Of the 91 relatively uncommon or rare variants, 21 were present only in the low-HDL group and 25 were present

only in the high HDL-C group. Overall, the prevalence of uncommon or rare variants was higher in the high HDL-C than the low HDL-C group. Thirty two of the 88 common variants demonstrated significant association (P-value <0.05) between the high and low HDL-C groups. We also examined 12 common variants (MAF≥ 0.05) in the total NHW sample and identified 7 variants to be significantly associated with lipid levels.

In conclusion, our comprehensive resequencing of the LPL gene confirms that both common and rare variants in this gene are associated with interindividual variation in plasma lipid profile.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

## ACKNOWLEDGMENT

First of all, I would like to express my sincere gratitude to my thesis advisor Dr. M. I. Kamboh for giving me the opportunity to join his lab and participate in this exciting research. Thank you for your invaluable assistance, support, and encouragement during the course of my work. I am also deeply indebted to other my committee members, Drs. F. Yesim Demirci and Clareann Bunker, for their time and effort in reviewing this work.

I also wish to thank Turkish Ministry of Education for supporting me financially during my graduate study.

Special thanks to Dr. F. Yesim Demirci for all her efforts and devoted works that she has done for my project in the past two years. I am sure that without your assistance and knowledge, this work would not have been successful. I truly appreciate your understanding and gracious support.

Thank you to my coworkers and fellow students in Dr. Kamboh's lab, especially Yuee Wang, M. Irfan Kamboh, Kate Hughes, and Jatinder Singh for all of their assistance with my project and guidance during my training period. Thanks to Zaheda Radwan for sharing all joys and hard times during the time of the project.

I would like to thank all my friends who supported me whenever I needed, no matter how far the distances were.

Lastly, I wish to thank my family for all they have done for me. Thanks dad for being such a wonderful role model; your devotion to science and medicine always encourages and inspires me. Thanks mom for your efforts that make the distances disappear and make me feel as if you are with me whenever I need you.

# 1.0 BACKGROUND AND SIGNIFICANCE

## 1.1 CORONARY HEART DISEASE

### 1.1.1 Public health importance of coronary heart disease

In the United States, coronary heart disease (CHD) is the most common cause of death and major killer of American males and females; one of every six deaths is caused by CHD (Coronary Heart Disease) according to the National Vital Statistics in 2009 (Heron et al. 2009). Based on the Framingham Heart Study (FHS) in 2001 more than half of the total cardiovascular events were due to the CHD in females and males under age 75 (Lloyd-Jones et al. 2010; American Heart Association 2010). In 2006, almost 425,000 (224K males and 201K females) people died from CHD and the CHD death rate was 134.9 per 100,000 (176.3 for white males, 206.4 for black males, 101.5 for white females and 130.0 for black females (Lloyd-Jones et al. 2009). American Heart Association computation based latest available mortality data show that in every twenty five seconds an American will experience a coronary event and it will cause death of one person in about every minute. The same data computes that about 81% of Americans die of CHD are at age 65 or above (Lloyd-Jones et al. 2010). According to the data of National Health and Nutrition Examination Survey (NHANES) between 2003 and 2006, the prevalence of CHD in the United States was 17,600,000 for adults age 20 and older. It means 7.9 percent (9.1 percent for males and

7.0 percent for females) of American adults age 20 and older have CHD (Lloyd-Jones et al. 2010). The incidence is different between females and males; males at 40 and older have high risk of developing CHD than females at the same age. The percentages for lifetime risk of developing CHD are 48.6 % for males and 31.7% for women at age 40 (Ulrich et al. 1999). The latest available data of the American Heart Association estimates that the direct and indirect cost of CHD is $177.1 billion for 2009 (Lloyd-Jones et al. 2010).

### 1.1.2 Risk factors of coronary heart disease

The risk factors for CHD are classified into two categories according to their causative and quantitative contributions to CHD. Low level serum high-density lipoprotein (HDL) cholesterol, elevated serum total and low-density lipoprotein (LDL) cholesterol, elevated blood pressure, cigarette smoking, diabetes mellitus and advancing age are the major and independent risk factors. Several studies, including Framingham Heart Study, have investigated the quantitative relationship between these factors and CHD. The major and independent risk factors contribute the total risk of a person independently and total risk calculation can be done by summing of the risk imparted by each of the major risk factors. Second category includes predisposing and conditional risk factors that are associated with independent risk factors for CHD. Obesity, physical activity, ethnic characteristics, psychosocial factors, family history of premature CHD are predisposing risk factors and elevated serum triglycerides, small LDL particles, elevated serum homocysteine, elevated serum lipoprotein (a), prothrombotic factors and inflammatory markers (eg. C-reactive protein) are conditional risk factors for CHD (Grundy et al. 1999). After nine years follow up, the data of the three prospective cohort studies includes both sexes and a spectrum of adult ages shows that 87% to 100% of fatal CHD cases had prior exposure to at least

one clinically elevated major CHD risk factor which are cholesterol ≥240 mg/dL [≥ 6.22mmol/L], arterial blood pressure≥140/190 mm Hg, medication use for cholesterol and hypertension, high cholesterol, cigarette smoking or clinical diabetes; 64% to 100% of these fatal CHD cases had prior exposure to two or more major CHD risk factors at higher than favorable levels (Greenland et al. 2003). The mortality data, prevalence and incidence of CHD show that male gender and blacks has higher risk of developing a CHD (Keil et al. 1989; Cooper et al. 1992; Johnson et al. 1986).

Since family history is an important risk factor for CHD, several studies imply the importance of genetic factors in the pathogenesis of CHD due to their aggregation in families with combination of lifestyle and common environment (Jannotti et al. 2000; Tiret et al. 2002; Breslow et al. 2001). The heritability of these risk factors has been investigated in many twin studies to understand the role of shared genes and shared environment. These results strength the importance of the genetic factors in predisposition to CHD (Hans et al. 1997; Marenberg et al. 1994; Chen et al. 2009; Austin et al. 1987). Most of the known genetic variation that affects the risk of CHD underlay in the genes that have an important role in lipid metabolism, including the lipoprotein lipase (*LPL*) gene. These variations may result a change in the gene function and cause changes in the lipid levels. This emphasizes the significance of the genetic studies for unraveling the genetic variation in candidate genes that may affect the lipid levels and contribute to the etiology of CHD.

## 1.2 HIGH DENSITY LIPOPROTEIN (HDL)

High density lipoprotein (HDL) is the smallest and densest lipoprotein particle in the circulation which consists of several distinct subclasses that vary in size, shape, density (1.063 to 1.21g/ml), surface charge and composition (Rye et al. 2009; Tsompanidi et al. 2009). The other major groups of lipoproteins, in order of largest to smallest according to their lipid content, are chylomicrons, very low density lipoprotein (VLDL), intermediate-density low protein (IDL), and low density lipoprotein (LDL). They are responsible to carry the lipids like cholesterol and triglyceride (TG) in the circulation. HDL particles compose by a hydrophobic core of cholesterol esters and a small amount of triglycerides (TGs); is surrounded by phospholipids, free cholesterol and apolipoproteins (apos) (Rye et al. 2009; Florentin et al. 2008). HDL may have a discodial or spherical shape depending on the lipid composition. Mature HDL is in spherical shape and contains 45-55% apolipoproteins, 26-32% phospholipids, 15-20% esterified cholesterol, 3-5% free cholesterol and about 5% triglycerides (Tsompanidi et al. 2009). The major apolipoprotein particles of HDL are apoAI and apoAII and the minor apolipoproteins are apoIV, apoCI/CII/CIII, apoD, apoE, apoJ, apoK and apoM (Florentin et al. 2008). HDL particles also include liposoluble, antioxidants, and several enzymes like paraoxonase-I (PONI), platelet-activating factor acetylhydrolase (PAF-AH) and glutathione phospholipid peroxidase (Florentine et al. 2008). As a major component of HDL, apoAI has a role in biogenesis and functions of HDL (Tsompanidi et al. 2009).

Although, high blood cholesterol is a major risk factor for CHD, cholesterol which is carried in HDL particles is called ‘good’ cholesterol and it is beneficial for a number of reasons (Toth, 2005; Genest et al. 1999). American Heart Association considers low high density lipoprotein (HDL-C) as a major risk factor for heart disease if it is less than 40 mg/dL for men and less than 50 mg/dL for women; HDL-C levels of 60 mg/dL and above is considered as being protective against heart disease.

### 1.2.1 HDL-Metabolism

Unlike the other lipoprotein particles, HDL promotes the reverse cholesterol transport (RCT), carries the excess cholesterol from non-hepatic cells (such as macrophages) to the liver and steroidogenic organs so cholesterol can be able to use for the synthesis of lipoproteins, bile acids, Vitamin D, and steroid hormones (Genest et al. 1999; Stein et al.1999; Von Eckardstein et al 2001). There are three different pathways that reverse cholesterol transport can take place. First, LDL receptors on liver recognize the multiple copies of apoE particles on mature HDL particles so they can be taken up by liver (Bruce et al. 1998). Second, scavenger receptor class B1 (SR-B1) mediates the selective uptake of the high-density lipoprotein cholesteryl esters (HDL-CE) (Acton et al. 1996). Third, cholesteryl ester transfer protein (CETP) transfers the cholesteryl esters from HDL to the TG rich lipoprotein particles, and then they can be taken by the liver via LDL receptors (Bruce et al., 1998)

### 1.2.2 The antiatherogenic role of HDL particles

The antiatherogenic role of HDL particles is mainly related its function in reverse cholesterol transport pathway which removes cholesterol from peripheral cells (macrophages in the artery wall) (Rye et al. 2009; Florentin et al. 2008). Besides this function, HDL has anti-thrombotic and anti-inflammatory properties that contribute to its atheroprotective role (Von Eckardstein et al. 2001; Genest et al. 1991). It also prevents atherosclerotic lesion progression by inhibiting LDL oxidation, promoting endothelial repair, inhibiting the binding of monocytes to the endothelium and improving endothelial function (Rye et al. 2009; Assmann et al. 1993; Rust et al.1999)

### 1.2.3 HDL cholesterol and CHD risk

Lipid levels and their role in development of CHD have been studied since the early 1950s (Barr et al. 1951). Despite the fact that HDLs carry about 20% of the total plasma cholesterol, several clinical and epidemiological studies have demonstrated the inverse relationship between HDL-C and CHD (Gordon et al. 1989, Castelli et al. 1986, Castelli 1977). This was reported as a 1 mg/dL increase of HDL-C levels is associated with a 2-3% decreased CHD risk (Gordon et al. 1989). It is worth noting that although the direct relationship between the CHD risk and low HDL-C is clearly identified in many studies, this does not reflect a causal relationship due to other factors that affect CHD risk as well as HDL-C levels (Vergeer et al. 2010). The risk factors for CHD that also cause low level of HDL-C are: male gender, smoking, obesity, hypertriglyceridemia, insulin resistance, physical inactivity, systemic inflammation, and low socioeconomic status (Vergeer et al. 2010).

### 1.2.4 Effects of genetic factors on HDL-C levels

Since the inverse relationship between HDL-C levels and CHD have been confirmed in various epidemiological studies, interest has focused on the role of genetic factors influencing HDL-C concentration (Weissglas-Volkov et al. 2010). Based on family and twin studies, the heritability of HDL-C levels is estimated up to 80 % (Kronenberg et al. 2002; Perusse et al., 1997; Wang and Paigen 2005; Goode et al. 2007). This means, in addition to environmental factors, genetic factors play an important role in affecting the variation of plasma HDL-C levels.

The genetic inheritance of low and high HDL-C levels can be monogenic, polygenic, or can be determined by interactions between several genes and their interactions with environmental factors. Most of the researchers conduct studies by using case/controls, families, and unascertained populations to understand the role of major genes in lipid metabolism influencing HDL-C variations (Weissglas-Volkov et al. 2010). In this aspect, the mutations in the *APOA1*, *ABCA1*, *LCAT*, and *LPL* genes have been associated with low HDL-C levels and they cause monogenic disorders as apoAI deficiency, Tangier disease, familial lecithin-cholesterol-acyl transferase (LCAT) deficiency (FLD), and lipoprotein lipase (LPL) deficiency (Type I hyperlipoproteinemia or familial chylomicronemia), respectively. However, mutations in the *CETP* and *HL* genes have been associated with high HDL-C levels (Weissglas-Volkov et al. 2010).

The approaches to identify the susceptibility genes in complex polygenic diseases fall into two categories: candidate gene studies and genome wide association studies (GWAS). The candidate gene studies using resequencing and association techniques have identified several variants in some of the genes that have a known important function in HDL metabolism. Among

these genes (*CETP, LIPC, LPL, LIPG, LCAT, ABCA1, APOA1, APOC3, APOC5, APOE, SR-BI,* and *PON1),* the common variants that have been identified in *LPL, CETP*, and *LIPC* are the most significantly associated with HDL-C levels (Weissglas-Volkov et al. 2010, Sviridov et al. 2007).

There are ten GWASs published on HDL-C levels (Willer et al. 2008, Kathiresan et al. 2008a, Kooner et al.2008; Wallace et al. 2008; Kathiresan et al. 2007; Aulchenko et al., 2009, Chasman et al., 2008; Sabatti et al. 2009; Ridker et al. 2009, Heid et al. 2008). These results are consistent with the previous association studies and the strongest association signals were observed for SNPs in the *LPL, CETP,* and *LIPC* genes. However, these GWASs also identified new candidate genes influencing HDL-C levels which need further studies to understand their functional importance in the HDL metabolism.

## 1.3 LIPOPROTEIN LIPASE

The lipoprotein lipase (LPL protein; *LPL* gene) is located on chromosome 8p22 in humans (Sparkes et al. 1987). The National Center for Biotechnology Information (NCBI) reference nucleotide sequence is NC_00008.10 (http://www.ncbi.nlm.gov/sites/entrez). It is a member of the TG lipase gene family; the other genes in the same family are hepatic lipase (HL), pancreatic lipase (PL), and endothelial lipase (EL). The *LPL* gene comprises 10 exons spanning 30 kb (Deeb and Peng 1989). The gene encodes 475 amino acids and it becomes a mature protein of 448 residues after cleavage of a 27-amino-acid signal peptide (Wion et al 1987). 5'- untranslated region, the signal peptide and the first two amino acids are encoded by the first exons. The rest of the 9 exons encode the remaining 446 amino acids. The $10^{th}$ exon is the largest exon which encodes the long 3'-untranslated region of 1,948 nucleotides. Figure 1 shows the gene structure

and the basic promoter elements of the *LPL* gene which are located within 101bp upstream of the transcription starting site (Wang et al 2009).

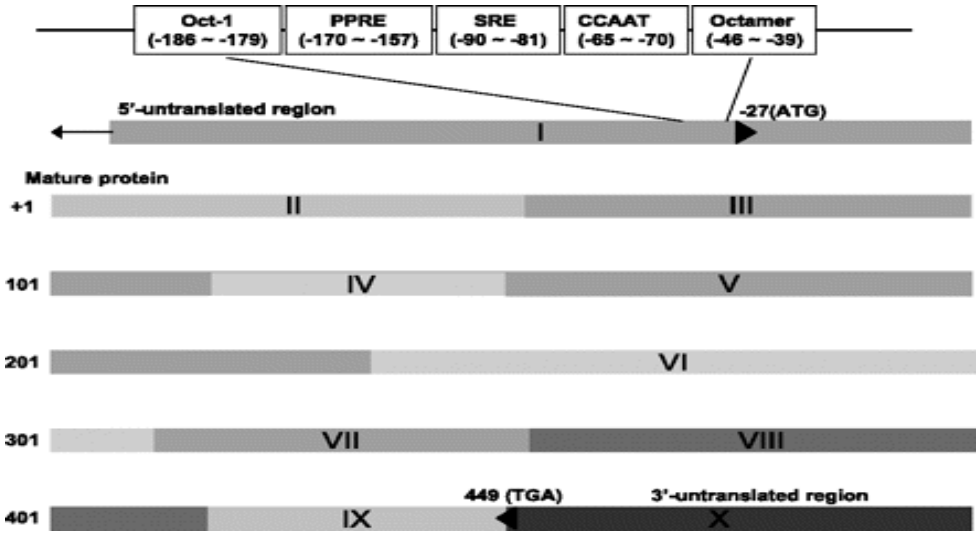



**Figure 1.** Lipoprotein lipase gene structure with promoter elements. (From Wang et al 2009, "Am Physiol Soc, used with permission")

### 1.3.1 *LPL* gene regulation

The expression and regulation of *LPL* is regulated at transcriptional, posttranscriptional, translational and posttranslational levels in a tissue specific manner. Hormonal levels and nutritional regulations can also affect the expression of *LPL* by interacting proteins or directly modify the regulation of *LPL* (Wang et al. 2009). Variations in *LPL* gene sequence, expression and regulation are important because they may influence its activity and also its contribution to disease so screening the *LPL* gene in people who have different level of HDL-C levels may explain how *LPL* variants affect HDL-C levels and CHD. HDL-C levels may be regulated in three different ways due to plasma LPL activity (Weissglas-Volkov 2010). First, some of the phospholipids and apolipoproteins are shed and transferred to the HDL particles during the hydrolysis of TG-rich lipoproteins (Klos et al. 2007). Second, LPL limits the CETP-mediated

HDL-C reduction by decreasing plasma TG because the exchange of cholesterol for TG from HDL is modulated by CETP due to the amount of VLDL particles (Rye et al. 2009). Third, HDL lipid composition is altered by LPL enzyme activity because affects the catabolic rate of apoA1 so they can be more rapidly cleared from the circulation (Barter et al. 2003).

## 1.4 LIPOPROTEIN LIPASE ENZYME

In 1943, Paul Hahn first observed clearance of postprandial lipidemic plasma heparin injection (Hahn 1943). After this finding, more studies focused on understanding the mechanism of so called heparin releasable 'clearing factor' and it was revealed that this factor is a lipolytic enzyme which has high activity against the TG component of plasma lipoproteins. The enzyme was called LPL, after discovering its role on breaking down plasma triglycerides of TG-rich lipoproteins which are VLDL and chylomicrons (Tsutsumi 2003). LPL is a major rate-limiting enzyme responsible for the hydrolysis of TG-rich particles circulating in the bloodstream, so it has a central role in overall lipid metabolism and transport. ApoCII is required for activation of LPL and there are several factors such as apoCIII, apoE, high salt conditions and fatty acids that inhibit LPL activity (Murthy et al. 1996). It is synthesized by parenchymal cells and then transferred to the luminal surface of endothelial cells where it is anchored to the surface of the cells by heparan sulfate proteoglycans (Goldberg et al. 1996).

LPL enzyme activity has been detected in several extrahepatic tissues and cells, including adipose tissues which it is predominantly found and in other tissues such as heart, skeletal muscle, lung, lactating mammary gland, brain, kidney, and macrophages (Kirchgessner et al. 1989). LPL enzyme activity has been also identified in differentiated macrophages, placenta, spleen,

pancreatic β-cells and steroidgenic tissue, but not in liver of adult animals. Due to tissue-specific activity of LPL, high levels of hepatic messenger RNA and enzyme activities are shown during suckling period in lactating mammary glands (Semenkovich et al. 1989, Yacoub et al. 1952; Merkel et al. 1998a). The high level LPL activity in adipose tissue is seen after a meal consistent with the fat storage function of adipose tissue and its level is increased by fasting and exercise training in skeletal muscle (Merkel et al. 2002).

The LPL enzyme is catalytically active in its dimer nascent form which is composed of 55 kDa subunits and it shows a head-to-toe configuration by noncovalent interactions (Zhang et al. 2005). Maturation of nascent LPL starts in the endoplasmic reticulum depending the activity of lipase maturation factor 1 (Peterfly et al. 2007). According to the three-dimensional structure of LPL, it has been shown that LPL is organized into two structurally distinct domains; a larger N-terminal domain (residues 1-312) and a smaller C-terminal domain (residues 312-448) connected by a flexible peptide (Yang et al. 1989, Bengtsson-Olivecrone et al. 1986, Lookene et al. 1993) N-terminal domain includes catalytic site covered by a lipid-binding lid and catalytic center consists of three amino acids: Ser132, Asp156, and His241. Interaction of LPL with lipoprotein substrates occurs in the C-domain and it results a conformational change that allows enzyme to conduct its catalytic function by opening of the lid in the N-terminal domain (Santamarina-Fojo et al. 1994).

There are four distinct physiological activities of LPL that have been identified in several studies which all influence the plasma lipoprotein profile and the cellular metabolism of fatty acids and lipids (Preiss-Landl et al 2002). First and major activity of LPL is its hydrolyzing function of TG-rich lipoproteins (chylomicron and VLDL) to provide fatty acids to underlying tissues (Figure 2). Second, LPL has uncatalytic activity called ‘bridging function’ which allow enzyme to interact with lipoproteins and anchoring them to the vessel wall so it facilitate

triglyceride hydrolysis and lipoprotein uptake ( Merkel et al.1998b, Merkel et al. 2002). Third, LPL has a role in the uptake of lipoproteins by acting as a ligand for the LDL receptors, VLDL receptors, megalin, and LDL receptor related protein (Medh et al. 1996, Takahashi et al. 1995, Kounnas et al. 1993, Beisiegel et al. 1991). Fourth, LPL facilitates the selective uptake of lipoproteins associated lipids (Merkel et al. 2002; Seo et al. 2000) and lipophilic vitamins such as vitamin A and vitamin E (Preiss-Landl et al. 2002; Sattler et al. 1996; Van Bennekum et al. 1999).

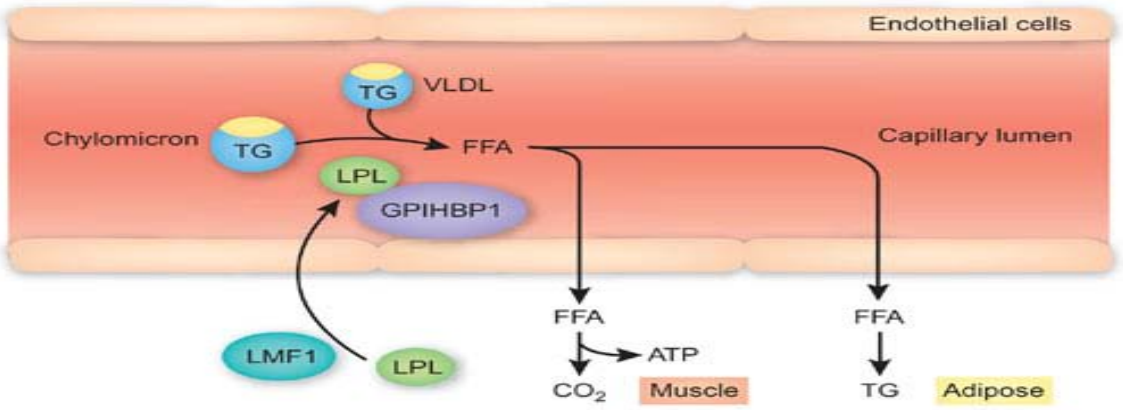



**Figure 2.** LPL activity in fatty acid transport to muscle and adipose tissues (Attie et al. 2010, used with permission).

### 1.4.1 Effects of Lipoprotein Lipase activity on lipid levels

Lipoprotein lipase is the major enzyme in lipid metabolism so its catalytic function and bridging function are both critically important for humans to maintain the lipid levels favorably. Due to its essential role in regulation of lipid metabolism, it is a candidate susceptibility gene for influencing CHD risk. Several studies show variations in the human *LPL* gene that have a role in the disease's etiology due to changes in *LPL* gene sequence, expression and regulation (Xie et al. 2010). Havel

and Gordon discovered LPL deficiency in 1960 and then the gene was cloned in 1991 followed by identification of several mutations (Havel et al. 1960, Henderson et al. 1991). Clinically abnormal lipid levels have been associated with CHD, atherosclerosis and obesity in individuals who have a number of functional DNA sequence variations in their *LPL* gene (Murthy et al. 1996; Reymer et al. 1995; Brunzell et al. 1995; Wiebusch et al. 1992). LPL deficiency, known as Type 1 hyperlipoproteinemia or familial chylomicronemia (MIM 238600), is a rare autosomal recessive disorder (1/1,000,000) caused by defects in the *LPL* gene and it is correlated with severe hypertriglyceridemia due to chylomicronemia and VLDL accumulation with very low levels of LDL-C and HDL-C levels (<20mg/dL). LPL deficiency is characterized with hepatosplenomegaly, xanthomas, acute pancreatitis and recurrent episodes of abdominal pain (Klos et al. 2007). Several studies have shown that individuals with LPL deficiency are compound heterozygotes and most of the disease causing mutations occur predominantly in exons 4, 5 and 6 of the *LPL* gene. The disease causing mutations usually result catalytically inactive LPL enzyme that is being degraded within the cell and cause little or no postheparin LPL activity (Wang et al. 2009).

## 1.5 LIPOPROTEIN LIPASE GENE SEQUENCE VARIATION

*LPL* is a candidate gene for influencing the risk of CHD due to its central role in lipid metabolism. Hypertriglyceridemia, which is the main clinical feature of LPL deficiency and dysfunction may be also a risk factor in the pathogenesis of CHD as well as dsylipidemia, type 2 diabetes (T2D), essential hypertension (EH), and Alzheimer's disease (Xie et al. 2010). It is thought to be a common biological basis for these complex diseases that structure, expression and

function of the *LPL* gene may have a role in that. This explains importance of understanding the *LPL* gene sequence variation and its association between specific phenotypes. Initially, 9.7kb of *LPL* sequence, about one-third of the *LPL* gene, was sequenced in 71 individuals, that identified 88 variable sites (79 were single nucleotide substitutions and 9 were insertion-deletion) (Nickerson et al. 1998). The average nucleotide diversity across the region, spanning from 3' end of intron 3 to the 5' end of intron 9, was found to be 0.2 %, which is almost one variant in every 500bp (Nickerson et al. 1998). Nickerson et al. (1998) sequenced a total of 9,734bp including 8,736bp in non-coding region and 9,98bp in coding region; 81 variants were identified in non-coding region and 7 variants were identified in coding region. Most of the variable sites detected in the coding region were silent and third base substitutions except one variable site that leads to stop codon and premature truncation of the protein in position 9040C>G (447Ser→Ter) and three amino-acid substitutions were observed in the coding region in positions 2849A>G (291Asn→Ser), 6176 G>A (370Val→Met), and 6203A>G (379Thr→Ala) (Nickerson et al. 1998). In another study, sequence variants were identified by resequencing all 10 exons and introns/flanking regions of the human *LPL* gene in 95 subjects, and 24 variants were identified including 9 in coding and 15 in non-coding regions (Morabia et al. 2002). Three of the 24 variants were found to be associated with amino acid changes; 291Asn291→Ser (N291S) in exon 6 and 447Ser→Ter (S447X) in exon 9 and 9Asp→Asn (D9N) in exon 2 (Morabia et al. 2002). In addition to these common coding variants, almost 100 naturally occurring mutation have been described in *LPL* gene: 61 missense mutation, 12 nonsense mutations, 10 frameshift mutations or small insertion/deletions, 3 gross mutations, 8 splicing mutations and 4 promoter variants (Merkel et al. 2002). Ser447X is the only variant associated with increased LPL activity and found in up to 20% in the general population (Merkel et al. 2002). The most common *LPL* coding variants with

their locations and percentages of carriers in related populations are briefly described in the Table 1.

**Table 1.** Most common *LPL* variants

| *LPL* Variant | Location | Carrier frequencies (%) in populations |
|---|---|---|
| Ser447X | Exon 9 | 20% in Caucasians (Humphries et al. 1998, Wittrup et al. 1999) |
| Asp9Asn | Exon 2 | 1.5% in Caucasians (Wittrup et al. 1999) |
| Asn291Ser | Exon 6 | 2-5% in Caucasians (Wittekoek et al. 1998) |
| -93T→G | Promotor | 76.4% in South African Blacks and 1.7% in Caucasians (Ehrenborg et al. 1997) |

## 1.6 *LPL* POLYMORPHISMS ASSOCIATED WITH PLASMA HDL-C LEVELS AND CORONARY HEART DISEASE

Several studies have investigated the association of *LPL* polymorphisms with HDL-C levels (Boes et al. 2009). Seven variants have showen significant association with HDL-C levels. Two of them are the common variants (Asp9Asn, N291S) that result amino acid changes and lead to decrease in enzymatic activity and so they have been projected to reduce HDL-C levels (Wittrup et al. 1999, Zhang et al. 1996; Mailly et al. 1995). While D9N (rs1801177) results in 3.2 mg/dl decrease in HDL-C levels and 20% increase in TG levels, Asn291Ser (rs268) leads to 4.6mg/dl decrease in HDL-C levels (Wittrup et al. 1999). The D9N variant is in near-complete linkage disequilibrium with the promoter variant (T93G), which may cause decreased LPL activity due to lower promoter activity but the role of the promoter variant is not yet clear (Merkel et al. 2002). The other amino acid substitution, Gly188Glu, decreases HDL-C levels (10mg/dl) and it is most

frequent in French Canadians in Quebec (Wittrup et al. 1999; Merkel et al. 2002). It is the major mutation that results nonfunctional LPL protein and has the strongest link with increased risk of CHD among all *LPL* variants (Merkel et al. 2002). The *Hind*III (rs320) polymorphism is located in intron 8 of the LPL gene and is in strong linkage disequilibrium with the Ser447X variant (Humphries et al. 1998). The *Hind*III polymorphism has been estimated to increase HDL-C levels up to 5.5mg/dl (Senti et al. 2001; Holmer et al. 2000; Corella et al. 2002; Ukkola et al. 2001; Radha et al. 2006, Ahn et al. 1993). Ser447X (rs328) is the other common variant that cause stop codon and it is also associated with increased HDL-C levels in several studies (Wittrup et al. 1999; Nettleton et al. 2007; Lee et al. 2004; Komurcu-Bayrak et al. 2007; Pallaud et al. 2001; Constanza et al. 2005; Kathiresan et al. 2008b; Talmud et al. 2002). Two single nucleotide polymorphisms (SNPs), rs326 and rs13702, have been identified to be in a strong LD with rs320 so they are also associated with HDL-C levels (Boes et al. 2009; Klos et al. 2006).

Genome wide association (GWA) studies have also identified many SNPs associated with HDL-C levels. To our knowledge, seven GWA studies have found association signals in the *LPL* gene with HDL-C levels (Boes et al. 2010). The reported SNPs with the lowest p-values are rs2083637 (Aulchenko et al. 2009), rs10503669 (Willer et al. 2008), rs331 (Chasman et al. 2008), rs328 (Kathiresan et al. 2008b), rs17482753 (Heid et al. 2008), rs17411031 (Wallacee et al. 2008), and rs326 (Kooner et al. 2008).

## 1.7 SPECIFIC AIMS

The objective of this study is to evaluate the role of common and rare genetic variation in the *LPL* gene in relation to HDL-C and correlated lipid traits in a non-Hispanic white (NHW) sample (n=623). Since *LPL* is a biological candidate gene for HDL-C levels, its common and/or rare variation is hypothesized to contribute to the variation in HDL-C levels. The following are the three aims of our study.

*Aim 1*: Resequence the *LPL* gene in a subset of phenotypically determined samples having HDL-C in the upper (n=47) and lower (n=48) $5^{th}$ percentile derived from a NHW sample of 623 individuals to identify both rare (allele frequency <0.01) and common variants (allele frequency ≥0.05).

*Aim 2:* Screen the tagSNPs of common variants of the *LPL* gene in the total NHW sample (n=623).

*Aim 3:* Evaluate the association of both rare and common *LPL* variants identified in Aim 1 and Aim 2 with HDL-C and correlated lipids (total cholesterol, triglycerides, LDL-C) levels in the NHW sample.

# 2.0 SUBJECTS AND METHODS

## 2.1 SUBJECTS

The study samples comprised 623 non-Hispanic whites (NHWs) derived from the San Luis Valley Diabetes Study in Colorado. The main features of the study subjects are depicted in Table 2. All subjects used in this current study were non-diabetics and a more detailed description of the sample population is given in Razagghi et al. (2000) and Demirci et al. (2010). Total fasting serum cholesterol was determined by using esterase-oxidase method, total HDL-C and TG were measured using enzymatic methods and LDL-C was calculated by the Friedewald equation when triglycerides levels were less than 400mg/dl (4.5 mmol/l) (Harris et al. 1998).

DNA was extracted from buffy coat using standard DNA extraction procedures.

**Table 2.** Population characteristic data (mean±SD) of entire NHW (n=623)

| Variable | Men (n=295) | Women (n=328) |
|---|---|---|
| Age (years) | 52.9 ± 0.6 | 52.4 ± 0.6 |
| BMI (kg/m2) | 26.2 ± 0.20 | 24.8 ± 0.2 |
| LDL (mg/dl) | 139.8± 2.0 | 134.7± 2.0 |
| HDL-C (mg/dl) | 43.9 ± 0.6 | 56.3 ± 0.7 |
| Triglycerides (mg/dl) | 147.6 ± 4.1 | 128.2 ± 2.9 |
| Total Cholesterol (mg/dl) | 213.7 ± 2.2 | 217.7 ± 2.1 |

A total of 95 NHW individuals whose serum HDL-C in the upper 5$^{th}$ (n=47) and in the lower 5$^{th}$ (n=48) percentiles were selected for resequencing of the entire *LPL* gene. Of the selected 95 individuals, 47 were females (23 with high HDL-C levels and 24 with low HDL-C levels) and 48 were (24 with high HDL-C levels and 24 with low HDL-C levels) males. A summary of the selected sample of 95 individuals is presented in Table 3.

**Table 3.** Population characteristic data (mean±SD) of the subset of population used for DNA sequencing

| Variable | Total n=95 | | *p*-value |
|---|---|---|---|
| | High HDL(n=47) | Low HDL(n=48) | |
| Age (years) | 55.45 ± 9.8 | 53.03 ± 10.54 | 0.25 |
| Sex (M/F) | 24/23 | 24/24 | 0.92 |
| BMI (kg/m2) | 23.17 ± 3.17 | 27.35 ± 3.90 | <0.001 |
| LDL (mg/dl) | 126.84 ± 46.95 | 136.95 ± 41.28 | 0.28 |
| HDL-C (mg/dl) | 77.68 ± 13.32 | 31.81 ± 4.37 | <0.001 |
| Triglycerides (mg/dl) | 114.09 ± 60.88 | 240.21±153.22 | <0.001 |
| Total Cholesterol (mg/dl) | 227.34 ± 51.76 | 208.81 ± 44.65 | 0.06 |

## 2.2 DNA SEQUENCING

The *LPL* gene is located on chromosome 8p22 and comprises 10 exons. Figure 3 depicts the location of the *LPL* gene on chromosome 8 with neighboring genes and indicates their orientations. The accession number for DNA reference sequence for LPL gene used in PCR and resequencing is NC_00008.10; it is derived from Genbank in NCBI site (http://www.ncbi.nlm.nih.gov). A total of 37 overlapping resequencing amplicons were sequenced in both directions and the PCR primers that were used to produce these overlapping

amplicons are given in Table 4. These amplicons cover the entire LPL gene (27,993 bp) as well as 1,196 bp in the 5' flanking region and 1kb in the 3'flanking region resulting into a total of 30,189 bp genomic fragment. Although we design most of the primers by using Pimer 3 software (http://frodo.wi.mit.edu/primer3/), we also used a subset of primers from a previous study (Nickerson et al. 1998). That sequenced only a portion of the targeted region (9.7kb), starting 3' end of intron 3 to 5' end of intron 9.

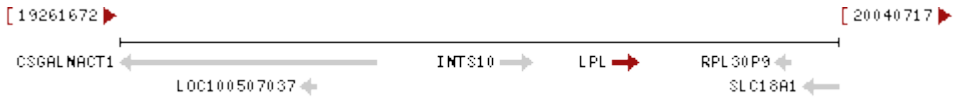



**Figure 3.** *LPL* and neighboring genes on chromosome 8.

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

| Amp. # | PCR Amp. (bp) | Forward Primer | Reverse Primer | Internal Sequencing Primer |
|---|---|---|---|---|
| 1 | 822 | 5'-GGGTTGGGGATACACTTCAT-3' | 5'-TGTTTTCCAAGGAGGGAAAG-3' | |
| 2 | 722 | 5'-TGATCCATCTTGCCAATGTT-3' | 5'-AGGGCTTTGCTCTCCATCT-3' | |
| 3 | 851 | 5'-GAAAGCTGCCCACTTCTAGC-3' | 5'-GTACTTTCTCCACCCCGACA-3' | |
| 4 | 703 | 5'-TGACCTGCAGTCACCTCTCT-3' | 5'-GCTCTCTATGCTGCTGTTGC-3' | |
| 5 | 800 | 5'-GGGGCCAAATGAGAATGTC-3' | 5'-AGTTGGCTCCTACCATCTTC-3' | |
| 6 | 1177 | 5'-GGATCAGTTTGAAAACACTGGA -3' | 5'-CATTTTGATGGCTGGAACAT-3' | |
| 7 | 1112 | 5'-TGCCTTATGCCAGATTGTTC-3' | 5'-TTGAATGAAGGGCTGTTGAG-3' | |
| 8 | 1147 | 5'-ATACCATTCTGGCTTGGATT-3' | 5'-ACTGATGTGGTCGATTTGGT-3' | |
| 9 | 1056 | 5'-AGCTGCATGTTAGAGAAGTCAA-3' | 5'-CCAAACTTCAGTCAGCTCTCC-3' | |
| 10 | 1093 | 5'-CTGCCCAATAGCAATCACAG-3' | 5'-CAATGGGTAAACACTCCAAGA-3' | |
| 11 | 1113 | 5'-TCTTGGTGGATGAATGGAT-3' | 5'-ATTACCAGTGTGAGCCATCG-3' | |
| 12 | 1022 | 5'-GCCATAGGAGTGGGAACAGT -3' | 5'-ACTGGAGGGTTGCTTGATTT-3' | |
| 13 | 787 | 5'-TCGAAAACACTTCAGAAACAAAA-3' | 5'-AGTAAATGGAGGCCCAGAGA-3' | |
| 14 | 1090 | 5'-CTGCGAGGTTGGTAAAGGAT-3' | 5'-CCTGCCTGTGCTGAAAATA-3' | |
| 15 | 1053 | 5'-TGTGATAAAATCTCAAATTCCTAAA-3' | 5'-TCCTACAGTGGCTGACATTTTT-3' | |
| 16 | 1019 | 5'-AGGGAGGGCTTCAGTTCAG-3' | 5'-TTCACAATGGGAACCCTGTA-3' | |
| 17 | 1049 | 5'-AACCCGATTTTCTTGCCTTA-3' | 5'-TGAATGCCCCCAGAAAATA-3' | |
| 18 | 1084 | 5'-AGAGTTGGGTGCCAAAACTT-3' | 5'-GGGTATATATTTTCCCATTATTCC-3' | |
| 19 | 691 | 5'-AACCAGGTAATTGGAAGTAAAAA-3' | 5'-ACAGTTCTGCCAAAAATAAAACT-3' | |
| 20 | 1061 | 5'-TGTTTACGGAAAAGTGAAACAAA-3' | 5'-GGGGCTTCTGCATACTCAA-3' | |
| 21 * | 475 | 5'-GGCCAAATGTGTATATGAAAAC-3' | 5'-CCATGACTGTAGAATAGGAGC-3' | |
| 22 * | 1783 | 5'-AGAGGACTTGGAGGTAAATATT-3' | 5'-GACTCCTTGGTTTCCTTATTTA-3' | 5´-ATGTTACTGGAACAGAAGATG-3´<br>5´-CTGGTCCACATCTGGGTAAA-3´ |
| 23 * | 1229 | 5'-AGGCTGGAGACTGTTGTAAAT-3' | 5'-CTCAGGTTTCCATCTGGATTC-3' | 5´–CTATCAACTCTGTTATGGTGGC-3´ |
| 24 | 708 | 5'-CCCTCTATGTGCTCATGCAA-3' | 5'-TGGGGCCACTGTTCTTTAAT-3' | |
| 25 | 1169 | 5'-GGAATGGTCGGAAAATGAGA-3' | 5'-AAGGAAAGGCAGCAGGACTA-3' | |

**Table 4 Continued**

| | | | | |
|---|---|---|---|---|
| 26 | 915 | 5'-CCACGCCCAACTAATTTC-3' | 5'-CCTAGAAAATGCAGACCTTGAA-3' | |
| 27 | 1057 | 5'-TGTTTTGGCCTTCTGATTTG-3' | 5'-CATGGTGAGACCCTGTCG-3' | |
| 28 | 755 | 5'-AGTAAGAAGTCCATGACAAAGTGT -3' | 5'-TTTCCTGGGTTTCCTACAAT-3' | |
| 29* | 1881 | 5'-CATCAATTACAGTCGTACCTAT-3' | 5'-TCAGCTTTAGCCCAGAATGC-3' | 5´-GAGCAGTCTTATGTTACTGGGC-3´ |
| 30* | 794 | 5'-TCATTTGCAGAAAGGAAAGG-3' | 5'-AATTCAGAACAGGAGTAGTG-3' | |
| 31 | 874 | 5'-TGCCTCTTTCCTACCTGACC-3' | 5'-ATTTTTGTAAAGGACGAAAAACAT-3' | |
| 32 | 1072 | 5'-AAAAACATGCCTATTAGGAAAAG-3' | 5'-CGCATCTGAACATTCTCTGTC-3' | |
| 33 | 1078 | 5'-CGGCCCTAGATGCAGTTTTA-3' | 5'-AGATTCGCCCAGTTTCTGAG-3' | |
| 34 | 1049 | 5'-AGAAGTCATTTGGCCCAGTC-3' | 5'-GCTGAGGATTACAGGCTCATT-3' | |
| 35 | 1046 | 5'-ACTTGGAGAGGGACGAAGAA-3' | 5'-TCACAACCCAAATCCAGAAA-3' | |
| 36 | 1044 | 5'-GCATAATTCGGAAGGGAAAA-3' | 5'-TTATCAAGGCAACCCAAAGC-3' | |
| 37 | 754 | 5'-GCCTGCATAAAGTACACAGGA-3' | 5'-CTTCTCCACATCCTCAGCAA-3' | |

*Primers from the study of Nickerson et al. (1998)

The GeneAMP® PCR System 9700 thermal cycler with a heated lid (Applied Biosystems, Foster City, CA) was used for performing the polymerase chain reaction (PCR). The PCR reactions and cycling conditions are presented in Table 5. Gel electrophoresis by using 96-well pre-cast agarose E-Gel® 96 2% with SYBR® Safe (Invitrogen Corperation, Carlsbad, CA) was performed following amplification of each of the PCR fragments to check the success of the reaction. Reamplification was done for some of the samples that were failed in the initial amplification and regular 2% agarose gel with ethidium bromide (2µl) was performed after PCR reaction of this subset of reamplified samples to confirm the amplification.

All the amplified samples were sent to a commercial sequencing lab where automated sequencing and capillary electrophoresis were performed on ABI 3730x1DNA Analyzers (Genomic Services of Beckman Coulter, Danvers, MA). Sequencing data received from the commercial lab was analyzed in our lab by Variant Reporter version 1.0 (Applied Biosystems) and Sequencher version 4.8 (Gene Codes Corporation, Ann Arbor, MI).

**Table 5.** PCR reaction and cycling conditions

| PCR Reaction (Total volume 25 µL) | | PCR conditions |
|---|---|---|
| DNA | 3.0 µL | 1. 95º C for 5 minutes<br>2. 95º C for 45 seconds<br>3. 58-60º C for 45 seconds<br>4. 72º C for 1 minute<br>- Repeat steps 2-4 for 40 cycles<br>5. 72º C for 10 minutes<br>6. Cool to 4º C |
| dH2O | 12.25-13.75 µL | |
| 10x BufferGold | 2.5 µL | |
| MgCl2 (25 mM) | 1-3.5 µL | |
| dNTPs (1.25mM) | 3.8 µL | |
| Forward Primer (20mM) | 0.4 µL | |
| Forward Primer (20mM) | 0.4 µL | |
| AmpliTaqGold (5U/µL) | 0.15 µL | |

## 2.3 GENOTYPING

While the sequencing was underway, we used the Hapmap database and Haploview program to determine the number of common tagSNPs that cover the entire gene and ~1kb flanking regions in whites. Seventeen tagSNPs were identified in Haploview (MAF≥4%, $r^2 \geq 0.7$) using SNP genotype data of CEU population (Utah residents with Northern and Western European ancestry) (www.hapmap.org). To date, we have genotyped 12 of them using TaqMan SNP genotyping assays. Table 6 lists the 48 variants that were captured by genotyping of 12 tagSNPs; 12 tagSNPs are highlighted in yellow. Table 7 lists the assay IDs of the genotyped 12 SNPs in the entire NHW sample with their refSNP IDs.

**Table 6.** Tagger results using SNP genotype data of CEU population provided from HapMap (MAF≥ 4%, $r^2 \geq 0.7$)

| *Bin* | *Variants captured* |
|---|---|
| 1 | rs3916027, rs295, rs297, rs291, rs13702, rs326, rs327, rs331, **rs320 (HindIII),** rs301 |
| 2 | rs264, rs3779787, rs271, rs256, rs255, rs3779788, rs263 |
| 3 | rs11570891, rs15285, rs3735964, rs1059611, rs12679834, **rs328 (Ser447X),** rs325 |
| 4 | rs312, rs4922115, rs11570892, rs316, rs4921684, rs330 |
| 5 | rs10099160, rs319, rs3200218 |
| 6 | rs17410577, rs13266204 |
| 7 | rs258, rs253 |
| 8 | rs1534649, rs10104051 |
| 9 | rs270 |
| 10 | rs281 |
| 11 | rs9644636 |
| 12 | rs285 |
| 13 | rs248 |
| 14 | rs249 |
| 15 | rs3289 |
| 16 | rs283 |
| 17 | rs343 |

**Twelve screened variants are highlighted in yellow.**

Table 7. TaqMan SNP Genotyping Assays

| **Assay ID** | ***LPL* refSNP IDs** | **Location** | **Position** | **Taqman Assay Type** |
|---|---|---|---|---|
| C__27500004_10 | rs3779787 | Intron 1 | 2335 | Pre-made |
| C__11856397_10 | rs13266204 | Intron 1 | 4424 | Pre-made |
| C___9642884_10 | rs1534649 | Intron 1 | 4060 | Pre-made |
| C__12104326_10 | rs249 | Intron 4 | 15425 | Pre-made |
| C___1842993_10 | rs253 | Intron 4 | 15836 | Pre-made |
| C__12104296_20 | rs264 | Intron 5 | 17599 | Pre-made |
| C___1842996_10 | rs270 | Intron 6 | 18095 | Pre-made |
| C__12104268_10 | rs283 | Intron 6 | 19517 | Pre-made |
| C__12104236_10 | rs312 | Intron 7 | 22416 | Pre-made |
| C___1843006_20 | rs327 | Intron 8 | 23955 | Pre-made |
| C___8804467_10 | rs3289 | Exon 10 3'UTR | 27611 | Pre-made |
| C___8804485_10 | rs1059611 | Exon 10 3'UTR | 28982 | Pre-made |

384-well plates containing dried whole genome amplified DNA was used in TaqMan genotyping. PCR amplification was performed using a PTC-200 MJThermal Cycler (Biorad) or a GeneAmp 9700 (Applied Biosystems). After thermal cycling, Real-Time PCR system (ABI Prism 7900HT Sequence Detection Systems) was used for endpoint fluorescence reading of the plates. TaqMan reaction and thermal cycler conditions are given in Table 8. The remaining common SNPs are currently being genotyped by the medium-throughput Sequenom IPLEX genotyping assays.

**Table 8.**TaqMan reaction and thermal cycler conditions

| TaqMan Reaction (total volume of 5 μL) | | PCR Cycling Conditions |
|---|---|---|
| $dH_2O$ | 2.43 | 1. 95º C for 10 minutes<br>2. 95º C for 15 seconds<br>3. 60º C for 1 minute<br>-repeat steps 2-3 50x |
| TaqMan Master Mix | 2.50 μL | |
| TaqMan Assay Mix | 0.06 μL | |

## 2.4 STATISTICAL METHODS

Allele and genotype frequencies were determined by direct counting. Concordance of the genotype distribution to Hardy-Weinberg equilibrium was tested using a $\chi 2$ goodness-of-fit test for each variant. The variants identified by sequencing were analyzed by using Haploview to determine allele frequencies and their distributions among high and low HDL groups and their LD patterns. In the subset of sample that was used for sequencing, the $\chi 2$ test was used to compare the allele frequencies between the low and high HDL groups. For those SNPs that were genotyped in the entire sample, linear regression was performed to test for the effects of genotypes on the means of plasma lipoprotein lipid levels. The HDL-C and TG levels were transformed using normal log transformation to reduce the effects of non-normality. The significant covariates were identified using stepwise regression in both directions. The additive and dominant models were used for data analysis. The covariates included in the final model were sex, age, BMI and smoking. 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. A P-value of less than 0.05 under one of these models was considered as suggestive evidence of association and a P-value between 0.05 and 0.1 was considered as marginally significant ($0.05 \leq P \leq 0.01$).

# 3.0 RESULTS

## 3.1 DNA RESEQUENCING

A total of 179 variants, single base substitutions or indels plus a microsatellite (tetranucleotide repeat marker) in intron 6, were identified in our study by complete resequencing of the *LPL* gene in 95 American NHW individuals falling in the upper (n=47) and lower (n=48) $5^{th}$ percentile of HDL-C distribution. Table 9 shows a summary of the *LPL* variants identified in our NHW population sample.

Of the 179 variants identified, 105 were found to be already reported in the Chip Bioinformatics which currently uses dbSNP build 130. Among these 179 variants, 88 had a MAF $\geq$0.05, 54 had a MAF 1-5% and 37 had MAF $\leq$0.01. Seventeen of those variants were insertions or deletions, and remaining 162 were single nucleotide substitutions. One hundred forty two variants were located in the introns (including the microsatellite), 7 were located in the 3' flanking region, 8 were located in the 5'flanking region and 23 were located in the exons. Of the 23 exonic variants, 3 resulted in non-synomous changes; aspartate9->asparagine (D9N) in exon 2, asparagine291->serine (N291S) in exon 6, and serine447->stop codon in exon 9, and 3 resulted in synonymous changes; valine108->valine in exon 3, glutamic acid118->glutamic acid in exon 4 and threonine361->threonine in exon 8.

The remaining 17 exonic variants were located in the last exon 10 that codes 3'UTR region (Table 9). Of the 17 insertions and deletions, only one of the insertions was located in the exonic region but it was in 3'UTR and so did not affect the protein sequence. The range of indels sizes was 1-20 base except a 697-base deletion identified in intron 2.

We identified 74 variants that are not reported in the dbSNP build 130; 22 of these new identified variants had MAF $\geq$0.05, 52 had MAF <0.05 and 33 had MAF $\leq$0.01. Of 74 new identified variants, 4 were located in 5' flanking region, 62 were in introns, 5 were in 3'UTR and 3 were in the 3' flanking region.

**Table 9.** *LPL* variants identified in our study for the NHW population

| ***LPL* variant** | **Alleles** | **Location** | **refSNP ID** | **Amino Acid Change** | **MAF** | **Call Rate %** |
|---|---|---|---|---|---|---|
| 208 | T>C | 5' flanking | rs1470186 | | 0.016 | 98.90 |
| 351 | C>A | 5' flanking | | | 0.005 | 98.90 |
| 428 | G>A | 5' flanking | rs73667465 | | 0.016 | 98.90 |
| 549 | C>T | 5' flanking | rs17091742 | | 0.016 | 98.90 |
| 958 | G>A | 5' flanking | | | 0.005 | 95.80 |
| 1088 | G>T | 5' flanking | | | 0.005 | 97.90 |
| 1090 | T>G | **5' flanking** | **rs1800590 (-T93G)** | | 0.011 | 98.90 |
| 1130 | G>C | 5' flanking | | | 0.005 | 97.90 |
| 2335 | G>T | Intron 1 | rs3779787 | | 0.122 | 98.90 |
| 2913 | T>C | Intron 1 | | | 0.005 | 95.80 |
| 3558 | G>A | Intron 1 | rs34309063 | | 0.247 | 100.00 |
| 3964 | G>C | Intron 1 | rs17410577 | | 0.239 | 98.90 |
| 4060 | G>T | Intron 1 | rs1534649 | | 0.430 | 97.90 |
| 4424 | A>G | Intron 1 | rs13266204 | | 0.263 | 97.90 |
| 4621 | C>G | Intron 1 | | | 0.021 | 100.00 |
| 4948 | C>G | Intron 1 | rs6997330 | | 0.022 | 97.90 |
| 5094 | C>G | Intron 1 | | | 0.005 | 97.90 |
| 5107 | C>T | Intron 1 | | | 0.005 | 96.80 |
| 5118 | A>T | Intron 1 | | | 0.022 | 95.80 |
| 5200 | C>T | Intron 1 | | | 0.006 | 91.60 |
| 5531 | G>A | Intron 1 | rs1031045 | | 0.016 | 96.80 |
| 5772 | A>G | Intron 1 | rs60633545 | | 0.016 | 98.90 |
| 5949 | T>G | Intron 1 | | | 0.137 | 100.00 |
| 6383 | G>T | Intron 1 | | | 0.005 | 100.00 |
| 6435 | G>C | Intron 1 | | | 0.005 | 97.90 |
| 6477 | T>C | Intron 1 | | | 0.005 | 98.90 |
| 6553 | C>T | Intron 1 | rs59254395 | | 0.016 | 100.00 |
| 6554 | A>G | Intron 1 | rs56043715 | | 0.016 | 100.00 |
| 6821 | C>T | Intron 1 | rs10104051 | | 0.426 | 98.90 |
| 7130 | T>C | Intron 1 | rs28615996 | | 0.021 | 73.70 |
| 7131 | T>G | Intron 1 | | | 0.007 | 75.80 |
| 7313 | G>A | Intron 1 | rs28645722 | | 0.016 | 98.90 |
| 7388 | C>G | Intron 1 | rs28575919 | | 0.016 | 100.00 |

**Table 9(Continued)**

| | | | | | | |
|---|---|---|---|---|---|---|
| 7503 | T>C | Intron 1 | rs6999612 | | 0.016 | 100.00 |
| 7512 | C>T | Intron 1 | rs3779788 | | 0.121 | 100.00 |
| 7556 | T>C | Intron 1 | rs59811201 | | 0.016 | 89.50 |
| 8221 | A>C | Intron 1 | rs7000460 | | 0.018 | 84.20 |
| 8250 | G>A | Intron 1 | rs59630933 | | 0.019 | 97.90 |
| 8415 | T>A | Intron 1 | rs56321069 | | 0.183 | 100.00 |
| 8467 | C>T | Intron 1 | | | 0.005 | 100.00 |
| 8516 | delG | Intron 1 | | | 0.011 | 50.50 |
| 9015 | A>G | Intron 1 | rs28445964 | | 0.021 | 68.40 |
| 9024 | T>C | Intron 1 | | | 0.008 | 98.90 |
| 9130 | T>A | Intron 1 | rs13252357 | | 0.005 | 98.90 |
| 9411 | A>C | Intron 1 | rs28689946 | | 0.016 | 98.90 |
| 9418 | G>A | Intron 1 | rs28582042 | | 0.016 | 98.90 |
| 9589 | C>T | Intron 1 | | | 0.016 | 98.90 |
| 9696 | G>T | Intron 1 | rs73667468 | | 0.016 | 97.90 |
| 9914 | T>G | Intron 1 | rs73667469 | | 0.016 | 100.00 |
| 10127 | G>A | **Exon 2 (non-synonymous)** | **rs1801177(D9N)** | **Aspartate9> Asparagine** | 0.016 | 100.00 |
| 10632 | C>T | Intron 2 | | | 0.005 | 94.70 |
| 10912 | A>G | Intron 2 | | | 0.022 | 94.70 |
| 10987 | C>A | Intron 2 | | | 0.122 | 96.80 |
| 11050 | T>C | Intron 2 | rs7016529 | | 0.016 | 96.80 |
| 11090 | C>G | Intron 2 | rs8176337 | | 0.190 | 98.90 |
| 11228 | T>C | Intron 2 | | | 0.005 | 96.80 |
| 11574 | G>A | Intron 2 | rs34123038 | | 0.049 | 100.00 |
| 11600 | G>C | Intron 2 | | | 0.011 | 100.00 |
| 11760 | A>C | Intron 2 | rs73667470 | | 0.016 | 100.00 |
| 11888_11889 | insA | Intron 2 | | | 0.011 | 95.80 |
| 12224_12920 | del697 | Intron 2 | | | 0.005 | 98.90 |
| 12449 | G>A | Intron 2 | | | 0.117 | 97.90 |
| 12484 | C>A | Intron 2 | | | 0.048 | 97.90 |
| 12550 | G>A | Intron 2 | | | 0.118 | 95.80 |
| 12810_12829 | dup20 | Intron 2 | | | 0.126 | 95.80 |
| 12853_12854 | Ins16 | Intron 2 | | | 0.055 | 95.80 |
| 12861_12864 | del4 | Intron 2 | | | 0.060 | 95.80 |
| 12878_12889 | del12 | Intron 2 | | | 0.005 | 100.00 |

**Table 9(Continued)**

| | | | | | | |
|---|---|---|---|---|---|---|
| 12884_12887 | del4 | Intron 2 | | | 0.016 | 100.00 |
| 13003 | G>T | Intron 2 | | | 0.053 | 100.00 |
| 13639 | G>A | Intron 2 | | | 0.011 | 98.90 |
| 13854 | G>A | **Exon 3 (synonymous)** | **rs1121923(V108V)** | **Valine108> Valine** | 0.016 | 100.00 |
| 14114 | T>C | Intron 3 | rs73667472 | | 0.080 | 100.00 |
| 15206 | C>A | Intron 3 | rs343 | | 0.042 | 100.00 |
| 15245 | G>A | **Exon 4 (synonymous)** | **rs248(E118E)** | **Glutamic acid118> Glutamic acid** | 0.047 | 100.00 |
| 15425 | T>C | Intron 4 | rs249 | | 0.095 | 98.90 |
| 15449_15450 | Ins2 | Intron 4 | | | 0.063 | 98.90 |
| 15653 | delA | Intron 4 | | | 0.410 | 92.60 |
| 15836 | C>T | Intron 4 | rs253 | | 0.410 | 100.00 |
| 16316 | C>G | Intron 5 | rs254 | | 0.074 | 97.90 |
| 16320 | T>C | Intron 5 | rs255 | | 0.105 | 96.80 |
| 16386 | C>T | Intron 5 | rs256 | | 0.118 | 97.90 |
| 16442 | G>C | Intron 5 | | | 0.005 | 97.90 |
| 16563 | T>A | Intron 5 | | | 0.016 | 93.70 |
| 16671 | G>C | Intron 5 | rs258 | | 0.409 | 100.00 |
| 17231 | C>T | Intron 5 | rs263 | | 0.135 | 100.00 |
| 17476 | A>C | Intron 5 | | | 0.005 | 95.80 |
| 17599 | G>A | Intron 5 | rs264 | | 0.116 | 60.00 |
| 17948 | A>G | **Exon 6 (non-synonymous)** | **rs268 (N291S)** | **Asparagine291> Serine** | 0.033 | 92.60 |
| 18065 | T>G | Intron 6 | | | 0.009 | 91.60 |
| 18086 | T>G | Intron 6 | rs269 | | 0.119 | 96.80 |
| 18095 | C>A | Intron 6 | rs270 | | 0.201 | 100.00 |
| 18121 | G>A | Intron 6 | rs271 | | 0.109 | 100.00 |
| 18297 | A>C | Intron 6 | | | 0.005 | 97.90 |
| 18395_18396 | InsT | Intron 6 | | | 0.121 | 100.00 |
| 18462 | T>G | Intron 6 | | | 0.005 | 98.90 |
| 18621 | C>T | Intron 6 | | | 0.005 | 100.00 |
| 18708 | T>C | Intron 6 | rs276 | | 0.021 | 95.80 |
| 18822 | T>C | Intron 6 | rs277 | | 0.232 | 97.90 |
| 18942 | G>A | Intron 6 | rs278 | | 0.211 | 97.90 |
| 19442 | A>T | Intron 6 | rs281 | | 0.242 | 95.80 |

**Table 9(Continued)**

| | | | | | | |
|---|---|---|---|---|---|---|
| 19445 | C>G | Intron 6 | rs282 | | 0.129 | 96.800 |
| 19517 | C>T | Intron 6 | rs283 | | 0.161 | 100.00 |
| 19608 | C>T | **Intron 6** | **rs285 (PvuII)** | | 0.407 | 97.90 |
| 19675 | A>T | Intron 6 | rs286 | | 0.060 | 95.80 |
| 19815 | G>A | Intron 6 | | | 0.005 | 97.90 |
| 19975 | A>G | Intron 6 | rs287 | | 0.172 | 100.00 |
| 20038 | T>C | Intron 6 | rs289 | | 0.159 | 100.00 |
| 20080 | C>T | Intron 6 | | | 0.005 | 100.00 |
| 20271 | T>C | Intron 6 | rs291 | | 0.163 | 98.90 |
| 20363 | A>T | Intron 6 | | | 0.053 | 100.00 |
| 20505_20506 | insA | Intron 6 | | | 0.163 | 100.00 |
| 20544 | T>C | Intron 6 | rs294 | | 0.080 | 100.00 |
| 20657 | A>C | Intron 6 | rs295 | | 0.163 | 98.90 |
| 20663 | G>A | Intron 6 | rs296 | | 0.005 | 97.90 |
| 20790 | T>C | Intron 6 | rs297 | | 0.163 | 91.60 |
| 21125_21128 | del4 | Intron 6 | | | 0.005 | 87.40 |
| 21353 | T>C | Intron 7 | rs301 | | 0.172 | 86.30 |
| 21780 | T>G | Intron 7 | rs304 | | 0.103 | 86.30 |
| 21820 | A>G | Intron 7 | rs305 | | 0.108 | 97.90 |
| 21895 | T>G | Intron 7 | rs308 | | 0.012 | 94.70 |
| 21965 | C>T | Intron 7 | rs310 | | 0.006 | 93.70 |
| 22044_22047 | del4 | Intron 7 | | | 0.075 | 92.60 |
| 22416 | G>C | Intron 7 | rs312 | | 0.072 | 100.00 |
| 22461 | G>A | Intron 7 | rs314 | | 0.180 | 92.60 |
| 22514 | T>C | Intron 7 | rs315 | | 0.011 | 92.60 |
| 22855 | C>A | **Exon 8 (synonymous)** | **rs316 (T361T)** | **Threonine361-> Threonine** | 0.079 | 88.40 |
| 23190_23191 | del2 | Intron 8 | | | 0.500 | 86.30 |
| 23192 | G>T | Intron 8 | | | 0.068 | 88.40 |
| 23388 | C>G | Intron 8 | rs318 | | 0.018 | 90.50 |
| 23395 | A>C | Intron 8 | rs319 | | 0.293 | 93.70 |
| 23496 | T>G | **Intron 8** | **rs320 (*Hind*III)** | | 0.185 | 94.70 |
| 23573 | T>C | Intron 8 | | | 0.017 | 95.80 |
| 23636 | A>C | Intron 8 | rs322 | | 0.170 | 98.90 |
| 23747 | T>C | Intron 8 | rs325 | | 0.067 | 98.90 |
| 23858 | A>G | Intron 8 | rs326 | | 0.206 | 100.00 |

**Table 9(Continued)**

| | | | | | | |
|---|---|---|---|---|---|---|
| 23955 | T>G | Intron 8 | rs327 | | 0.192 | 90.50 |
| **24143** | **C>G** | **Exon 9 (non-synonymous)** | **rs328 (Ser447X)** | **Serine447-> Stop codon** | 0.005 | 100.00 |
| 24505 | A>G | Intron 9 | rs329 | | 0.016 | 98.90 |
| 24573 | T>C | Intron 9 | | | 0.005 | 100.00 |
| 24815 | G>A | Intron 9 | rs330 | | 0.111 | 100.00 |
| 24824 | G>A | Intron 9 | rs331 | | 0.122 | 100.00 |
| 24852 | T>C | Intron 9 | rs12679834 | | 0.080 | 100.00 |
| 24899 | C>T | Intron 9 | | | 0.016 | 100.00 |
| 25005 | A>G | Intron 9 | | | 0.005 | 100.00 |
| 25049 | G>A | Intron 9 | | | 0.063 | 98.90 |
| 25320 | C>T | Intron 9 | | | 0.005 | 95.80 |
| 25335 | C>T | Intron 9 | | | 0.079 | 98.90 |
| 25352 | A>C | Intron 9 | | | 0.079 | 91.60 |
| 25844 | T>G | Intron 9 | | | 0.074 | 95.80 |
| 26201 | T>G | Intron 9 | | | 0.077 | 96.80 |
| 26234 | T>G | Intron 9 | rs10099160 | | 0.282 | 98.90 |
| 27000 | C>T | Intron 9 | | | 0.029 | 98.90 |
| 27160 | T>A | Intron 9 | | | 0.027 | 98.90 |
| 27229 | C>T | Intron 9 | rs11570891 | | 0.082 | 97.90 |
| 27249 | G>A | Exon 10-3' UTR | rs4922115 | | 0.085 | 98.90 |
| 27611 | T>C | Exon 10-3' UTR | rs3289 | | 0.016 | 95.80 |
| 27688 | C>T | Exon 10-3' UTR | | | 0.005 | 95.80 |
| 27783 | A>T | Exon 10-3' UTR | | | 0.011 | 95.80 |
| 28036 | A>G | Exon 10-3' UTR | rs11570892 | | 0.101 | 100.00 |
| 28067 | A>T | Exon 10-3' UTR | rs3208305 | | 0.203 | 100.00 |
| 28093 | C>T | Exon 10-3' UTR | rs1803924 | | 0.082 | 100.00 |
| 28382 | C>T | Exon 10-3' UTR | rs1059507 | | 0.110 | 100.00 |
| 28407 | C>A | Exon 10-3' UTR | | | 0.026 | 100.00 |
| 28464 | C>A | Exon 10-3' UTR | rs3735964 | | 0.079 | 100.00 |
| 28490 | A>G | Exon 10-3' UTR | rs3200218 | | 0.279 | 100.00 |
| 28524 | C>T | Exon 10-3' UTR | | | 0.005 | 100.00 |
| 28911 | T>C | Exon 10-3' UTR | rs13702 | | 0.200 | 100.00 |
| 28982 | T>C | Exon 10-3' UTR | rs1059611 | | 0.079 | 97.90 |
| 29046_29047 | Ins2 | Exon 10-3' UTR | | | 0.079 | 100.00 |
| 29086 | C>T | Exon 10-3' UTR | rs15285 | | 0.200 | 100.00 |
| 29088 | C>A | Exon 10-3' UTR | rs3866471 | | 0.105 | 100.00 |

**Table 9(Continued)**

| | | | | | | |
|---|---|---|---|---|---|---|
| 29287 | G>A | 3' flanking | rs3916027 | | 0.184 | 100.00 |
| 29315 | T>G | 3' flanking | rs9644636 | | 0.339 | 100.00 |
| 29474 | C>T | 3' flanking | | | 0.026 | 98.90 |
| 29487 | T>A | 3' flanking | rs4921683 | | 0.105 | 98.90 |
| 29547 | C>T | 3' flanking | rs4921684 | | 0.105 | 98.90 |
| 29557_29558 | InsA | 3' flanking | | | 0.005 | 98.90 |
| 29716 | T>C | 3' flanking | | | 0.026 | 95.80 |

### 3.1.1 *LPL* Annotated Sequence

Figure 4 depicts the variants identified in *LPL* within a color FASTA representation of the annotated reference sequence adapted and modified from the CHIP Bioinformatics database (http://snpper.chip.org). The color code for the reference sequence is as follows: green for 5' and 3' flanking regions, grey for introns, black for exons. Capital letters are used to show coding bases and small letters show the bases in the UTR and flanking regions. The variants identified in this study also listed in dbSNP build 130 are shown in blue font with refSNP ID; the variants that are identified in only our study are shown in red font. The 197 variants which were already reported in public databases but not identified in this study are shown in grey font. The small deletions and the insertions identified in our study are shown by highlighting the region in yellow. We identified a large deletion in intron 2 which was not reported before in any public databases and it is highlighted in *italics*.

```
                  1          2          3          4          5
            1234567890 1234567890 1234567890 1234567890 1234567890

19,839,852      ataaat atagtagatt ggaggttctg atttgatgag ccagtttctc
19,839,898  agccataaac tgagaggggg ttggggatac acttcattgt ccttcctggc
19,839,948  taatgtaaat cccttatatt taaaaagata tttaaaagta ttccaagcat
19,840,008  tttggcagaa aagcatagta tctaatgtta ttttttttctt attttatgtg
19,840,058  catgcctctt atccatttaa aaatagcttt actgacctat aatttacaca p.208/rs1470186-[T/C]
19,840,108  ctatataatt ctcccattga aagtgcataa ttctgttgct tttagtatat
19,840,158  ttacagagtt gtgcagcatc agcataatgt aatctagaac attgtcatca
19,840,208  actacccccca aatctctatt cttcccttcc cctattaatt acccagcccc p.351-[C/A]
19,840,258  aggcaagcac tgatctactt ttggtctcta tggatttgtc tatttgtgga p.428/rs73667465-[G/A]
19,840,308  cactttaaat ggaatcatac aatatgtgtc ttttgcgact atcttctttc
19,840,358  acttatcata actcaatacg gctttagatt atttgacctc gatgttctgc
19,840,408  ctctgaacat aaaatattat ccttgcattc cttgatgagt ttgaggattg p.549/rs17091742-[C/T]
19,840,458  agaataattt gcatgagaca aaaattagaa actagttaga gcaagtaggc
19,840,508  ttttctccat cacataagct gatccatctt gccaatgtta aaacaccaga
19,840,558  ttgtacaagc acaagctggg acgcaatgtg tgtccctcta tccctacatt
19,840,608  gactttgcgg gggtggggat ggggtgcggg gtgagtgagg gaggactgca
19,840,658  agtgacaaac aggattcgtc aaaagagagg tgtattaaag tgccgatcaa
19,840,708  atgtaattta acagctaaac tttccctcct tggaaaacag gtgattgttg
19,840,758  agtatttaac gtgaatcgat gtaaacctgt gtttggtgct tagacagggg
19,840,808  gcccccgggt agagtggaac cccttaagct aagcgaacag gagcctaaca p.958-[G/A]
19,840,858  aagcaaattt ttccgtctgc cctttccccc tcttctcgtt ggcagggttg
19,840,908  atcctcatta ctgtttgctc aaacgtttag aagtgaattt aggtccctcc p.1088-[G/T];p.1090/rs1800590-[T/G]
19,840,958  ccccaactta tgattttata gccaataggt gatgaggttt atttgcatat p.1130-[G/C]
19,841,008  ttccagtcac ataagcagcc ttggcgtgaa aacagtgtca gactcgattc
19,841,058  cccctcttcc tcctcctcaa gggaaagctg cccacttcta gctgccctgc Exon 1
19,841,108  catccccttt aaagggcgac ttgctcagcg ccaaaccgcg gctccagccc
19,841,158  tctccagcct ccggctcagc cggctcatca gtcggtccgc gccttgcagc
19,841,208  tcctccagag ggacgcgccc cgagATGGAG AGCAAAGCCC TGCTCGTGCT
19,841,258  GACTCTGGCC GTGTGGCTCC AGAGTCTGAC CGCCTCCCGC GGAGGGGTGG
19,841,308  CCGCCGCCGA CCGTAAGTTT TGCGCGCAAA CTCCCCTCCA CCTGCAGACC Intron 1 rs11570895
19,841,358  CGGCGGGTGG CCACTGCCAC CCGAACTGAG GATGAGAAGA AGGAAGTTGG
19,841,408  AAGGGGCGGT GGATGCGCCC AGGGACTCTC CCAGCCTGGG CTCTAGCCCC
19,841,458  GAAACGGTCC CCGGAGTGGG ATCCAGGAGG GGCCGGGAGG GAATCTCCTC
19,841,508  CCGATCGTGA AGCGGCGGCG CCCAGTTCCC GCTTTTTCTC TCTGCCGGGT
19,841,558  TCCCGCGCTA TCCCTTCCAC TCTGGCTGGG ACCGCGTTCC CGGGCTCGCA
19,841,608  GGCTCCGCCG GGGAGGTTCC GGGGTGTGGG GGCCGGGACG GCGGAGGCGG rs2898492
19,841,658  GGAGTAAGGG CCCGGCTGGC GGTGACCTGC AGTCACCTCT CTGCCGGAGG
19,841,708  GGCCCTGGAA TGAAAGGCGC GCGGGCCAAG GTGACCTCGC CTTGGTTGGC
19,841,758  ACTGCGGCTC AGCCCCCGCC CGGGGACTCG CGGGCCGACT GTGGCCCCTT
19,841,808  CTGGGGAAGC CGGGGCGCGG GGAGGCGTTC CGGGCATCTC AGCCGCACGG
19,841,858  GGTACGCTCG CCCTCGGCGG GGCCCCTCGC TCCGCTGTGG GAGTGGCAGT
19,841,908  GGGTGTCGGG GTGGAGAAAG TACGCGTGGC GCGGAGTCCT GGGGACGCGG rs7839976
19,841,958  CGTCCCACCC GCTCTGGGGA GCCCCGGACT CTCTCCAGCT TCCAGGCTCG
19,842,008  CATGCCCCTC TTTTCTTAGT GCCCTGAGAA CCCAGCGAGG GGCTGACCCT
19,842,058  CCCGAAACCG TGGCGCAGCC ACCAGCAATC TGTGGTCGCC GACTCGGGGG
19,842,108  GTTGCCAGGT CTGCGTTTGG CCACCCTTTC TGTCCTGGGG GCTGAGGTCA
19,842,158  GCTCCGGGCG CCCGGCCCCG CCGCGCGGCT GCGAGCACGT GGGGTTGACG p.2335/rs3779787-[G/T]
19,842,208  GGCGCCGCGT GGAGGCAGCG AGCACAACGG TGGTCACCGC CGCCAGGGAA
19,842,258  CCGCCCGCTC GCTGGGGTCC AGGCGTTCGG GGCCAAATGA GAATGTCTCA
19,842,308  GACCTGTCCG CAATGGAGGC AGCCTGCTTA ATTCGAACCT CGATTCAGTA
19,842,358  AACATGCAAC AGCAGCATAG AGAGCAGCTG AAGCCATTCA TAACACGGGA
19,842,408  CAACATTTCC TTTTTTCTTC CATGCTGGAA TTGCAATTAG GGCGGTGTCG
19,842,458  CTTGGATGTG CTCTCAGGCG GCACGTCCCC AGCGGTTCAA GTTATAATAG
19,842,508  AGTCTCTCCA TAGCTTTGAT GGCCGCTAAA CGTTTGTTTT ATTTTGGCAT
19,842,558  TAATTTGTGA AACATTTTTG TTAGGTTAAA AAACAAAAAG TTGGCAGGGA
19,842,608  GCAGTGGCTC ACCCATGTAG TCCCAGCACT TCGGGAGGCC GAGGCCGGAA
19,842,658  GACAGCTTGA GCCCAGAAGT TCGAAATCAG CCTGGGCAAC ATAGGGAGAC
19,842,708  ACCGTCTCTA TAAAAAAAAA ATACAAAAAT TAGCCGGGCG TGGTGGCCTG
19,842,758  TGCCTGTGGT CCCAGATTAGT CAGCAGGCTG AGAGATCACT TGAGCCAGGG p.2913-[T/C]
19,842,808  AGTTTCAGTA GACTGCAGTG AGCTGTGATT GCCCCACTGC ACTTCATCCT
19,842,858  GGGTGACAGT GAGACCCTGT CTCATAAAAT TAAAACAAAA CAAAACAAAC
19,842,908  ACTAAGTATA GATTCCATCA AAGCAAAATT GGATAAGAAA AAAAGTATCT
19,842,958  TTTCTATTGG ATCAGTTTGA AAACACTGGA GAGTTGATGA GAAAGTCTTC
19,843,008  AACATCTTAG GTGGGGTATG TTTCGTATGT TTCCCTTCGT ACTGCTTAAT
19,843,058  GCTGACAAGA AGATGGTAGG AGCCAACTCC AAATTCTTAT TTCAGAAAGC
19,843,108  ACACCATAGA ATAACGTCAT TTTCATTGCA AAACAAGCAC CGAAATATGT
19,843,158  CATCACATTC AAGTTTTTCC TAGGCTCCCT TACAGGTTCA AGATCCTAAA
```

```
19,843,208  TTCTTGGGAG TCAGTGTCAC CTCTCTGGGT TTAGGTTCCT CAACTCTGCA
19,843,258  ATGAGTTTGG ATGAGGCCAA TGTTCTCTGA GCCTGGTGTA ACTCTTGCCT
19,843,308  CTTTAAGTGG ACACTTATGT GATTAATTAG TTTAATTGAG TTGTAGCCAA
19,843,358  CACATGCTTT TCCTAGCTGT AAATATATTA AGGAAGGATT ATTTCCAAGT
19,843,408  AGACTGGAAA CGATGCCCTC CCATCCCCTC CACTTTCACT CTACTCACCC p.3558/rs34309063-[G/A]
19,843,458  AATATATCAT GCCTCTCCCA TCACAGCAAC TTTCTCCCTC TTTCTCCTCC
19,843,508  AGATGCATTC ATCTAGGAAG GTAAGAATTT CAGGGAGAGA AAGATGTCAC
19,843,558  CGTGGTAGAA AGACAGGGAT CACCTCCTCT GGGCTCTTGA GTTTACTTAT
19,843,608  TCATTCTGGA TTCTTTCTAA CAAGAATATG AGGACAAGAG GCACTGTCCT
19,843,658  CAGGCACTTC GTCCTGGGAG CCACCACCAT CTCTGCATGG CCCCAATTAG
19,843,708  GAAACGTGAA GAGCTAGGAG AGGGAGAGTA TGGTCAGTGC TTAGCAGCTG
19,843,758  AAGTTCCACT TGCCTGGCCA TCGTGAATTT CCAGGCTGTC TTCTGAGTTG
19,843,808  AACATGATGG CAAAGGAGAG CAAAATAGCA GATGTCACTG AAGGAGAGCT p.3964/rs17410577-[G/C]
19,843,858  CAGCGAGGGA GTGATTGATT AATAGCTGTA TTGAAAGGTG GGAGTCAGGT
19,843,908  ACGGGGGAAG AGCGGCGATG GAAAATTTTC GCTTTCTTTC AGCAGCTTAT p.4060/rs1534649-[G/T]
19,843,958  TTTTAACTCA GCTTTCTGTT CTTGCTTTAT TATGGAGGAA AAATTGGGCC
19,844,008  ATAGAGTTTA CTGCCTTATG CCAGATTGTT CAAGAAAATG CCTTGCAACT
19,844,058  TACAATATTT TGCAGCTAGT TTCTTCCGTG ACCACCACAA AGACTGCATT rs73667466
19,844,108  GACTTAAATA TGAAGATGTT CCAGCCATCA AAATGATGGT TGGTGATGAT
19,844,158  TTTGGATCAC AAAGTGTAAG GAAAGTATTC AAGACATGAG TATCATGATT
19,844,208  TTTTAAGGTC TGGATGAAGA GACCATTTGG ATTTACTAAT AAGGTAAATT
19,844,258  CCAACTTTTA TGGCAATAAA AACAACAAAA ACACTTATCA GTGTAAAGCT p.4424/rs13266204-[A/G]
19,844,308  TTGGGATCAT CTATCCATTA AATGAGTTTG TCACCACAGT GAACTAAATA
19,844,358  CCTTTTATCA ACAGAGACTT TCTAACCTGG GAGTAAAATT CTTGTCACAG
19,844,408  TGCTTTTGTC ACATTCTGTC TTTGCAAAAG TTGAAGGCTC CAATAGTTTC
19,844,458  TGAAGGACTA ATAGGATAGG GTTCCAACTT ACTCAGAGGC TAAGAGTTTG p.4621-[C/G]
19,844,508  AAATTTACTC TGAACGATGT CTGTTCACTA GACTGCGTGA CTGCAGTTTG rs58877654
19,844,558  CTGTCTTGTG GCCATTTTAA GATTGTCTGT GTGCACTGAC ACCATTTGCG
19,844,608  TACTCAACAA CAGGTATCTA CTAGGAAGGA AGGAATGGAT TATCTTAGGT rs10503668
19,844,658  GCTATATATA TATGTAAGTT CTGCCGGAAG TGAGCCTATT AAACTTGTGC
19,844,708  CAATTCATTC CTTTTGCTCT TCCCAGTCTG TGCTCTCAGA ATGATCAAAT
19,844,758  GCTATCTAAG TAGTGTGGTG ATTTTGACTG TTTGAATGAA TAAACGAAAA
19,844,808  CGTCTCCCAC CATTATTTGA CAGTAGAATA GAACAAAGAT GACCCAGTTG p.4948/rs6997330-[C/G]
19,844,858  GCGTGGCCAC ACTGTGTCTA ATCCAGCCAC CTTGCTCACA GCACTCAACC
19,844,908  CACTTTGGTG TGTGCGCTCA TCTGTATTTT CGTAAATGTT GAAGTCTCTT p.5094-[C/G]
19,844,958  TCTATGCAGT CAGGTAGGGG TAATACCATT CTGGCTTGGA TTAATTTGAG p.5107-[C/T]; p.5118-[A/T]
19,845,008  TGTAGTACAA AGTTTAGAGA AGCTTTTAAG TAGCATGAAA AGTCAGATGC
19,845,058  TTTCGGAGGG ATGGTGGTGA ATGTAGGTAA ATGGATCTGC ACTTAGAGAT p.5200-[C/T]
19,845,108  CCTCAACAGC CCTTCATTCA AGATACAGCT ATCATGACAT CAATAATGTA
19,845,158  CCTTTGAAAA AAAATTGGCT GCAGAATTAA TTACGGTGAA AAACAGTATA
19,845,208  GGAGTTAGCA CTTACATTTT AACTAAAAAG AATAGCGTCC CCATGTTTAT
19,845,258  TCAGCCCTCC CTCCAATAAA ACAATTGTTG GCAAAGTAAT CATGGACTTT
19,845,308  CATTGTGTTA GTTTGAGACA ACTGGATGTT TCCATTTGCC ATCCTCAGCA
19,845,358  ACAAGAAGAA AGTAGTCTCA GATTAACCAG GTAAGTATTT TGTATTTCAC p.5531/rs1031045-[G/A]
19,845,408  TTAAAAAATC TTGACTGAAA TATGCCTATG TGATGACACT CAACCAAATT
19,845,458  ATTTCATTAA CCAGTCCACA AGATCTCCTT AAATAATGAT GGCTTATTCA
19,845,508  CACTTGATGG TCTCATTCAG TGGGGCAATT TTAATACACA TCTCTGAACC
19,845,558  TATTTTTTAA CCCCTCTTTT TCAGTAGTGT GGAAGGTTAG CCCTAATATT
19,845,608  GGAGAAAATT CAGGGTAAAA TTCAGATGAT TCATACAGGA TTTATTTTTC p.5772/rs60633545-[A/G]
19,845,658  CTATTCCATT AATAAAACAA CTTTTATAAA AATAAAAAGT AGGCTGGCAC
19,845,708  AGTGGCTCAC TCCTGTAATC CCAGCATTTT GGGAGGCCGA GGCGGGTGGA
19,845,758  TCATGAGGTC GGGAGTTCAA GACCAGCGTG GCCAGGATGG TGAAACCCCA
19,845,808  TCTTCTACTAA AAATACAAAA ATTAGCCAGG CGTGGTGGCA GGCACCTGTA p.5949-[T/G]
19,845,858  ATCCCGGCTA CTCGGGAGGC TGAGGCAGAG AATGGCTTGA ACCCAGGAGG
19,845,908  CATAGGTTGC AGTGAGTCAC GATCGTGCCA CTGCACTCCA GCCTGGGTGA
19,845,958  CAGAGCAAGA CTCCGTTTCA AAAAATAATA ATAAAATAAA ATGAAATAAA
19,846,008  GTAAAGCTGC ATGTTAGAGA AGTCAAGAGC ATTACTTACG TTAGAATATC
19,846,058  TGAACAGACC AATCAATTCA GTCTGATCAT GATATTGATG TTTTTCTTCA
19,846,108  CCAAATCGAC CACATCAGTA ATTCACTTTG TTCTTCGATA TCCTACAGAC
19,846,158  ACTGCCTGAG TCGATAACTA TGACTATCAG TCTCAGAGAG CAAATGAATT
19,846,208  ACTGAGGAAG CCCTGTAGGA GTGAGAGAAA AGGGGGGAGA GAGAGAGAAA p.6383-[G/T]
19,846,258  GGGGTGGGGG GATAACAGGA GAACAGAAAT TCCAAAGAGA ATTGCATTCT p.6435-[G/C]
19,846,308  CATTGAGTTC TTGTACCTCA TGTCATTGCA TTAAATGTTCA TCTTACTCAC p.6477-[T/C]
19,846,358  GTGATGACTT TGATCTGCCT TTAAAGCACC ATCTGCTGCT TTCCTGGGAT
19,846,408  GCTCAACACT TCCCTCTTTC TAGCAACAAG AATTACCACT CTTCCCCTCT p.6553/rs59254395-[C/T];p.6554/-
rs56043715-[A/G]
19,846,458  ATACATTTAT CTTTCTCTAC GTGCTTTAAC TTCTCAGCCT AATTTCGTCT
19,846,508  CTGTGAGTTA TTATCTATGT TAGAATAAAT TCTTTGTCTT TGTTTACACA
19,846,558  CTCAGATTTG TAGTTATTTA TTTAGGAATT TAGGAATAAA GATTCCATAG
19,846,608  TCAGGAAAGG CACAATTTAT AACTTGCGTG TTACCCAAAA CTCTCCCCTA rs61274012
```

```
 19,846,658  AGGGCTTAAT ATGGACATTT CTGACGAGGC CTGATGGGCA GGTGGTACGG p.6821/rs10104051-[C/T];rs73667467
 19,846,708  TGATGCTAAG TTAAATTCAG AATGAAGGCC TGCCTTTCCT TCCCTCCTTC
 19,846,758  CTTCCCCTTC CCCTTCTTCC TTCCTTCCTT CCTTCCCTCC ATCCCTCCCT
 19,846,808  CCTGTACTCC TCTTCTTTCT CAATTCTAAG GTGGCCTTAA TTTCTAAGGG
 19,846,858  ACATGGCAAA AGACAGTCTA GTTGGATGAG TGCAGTCACT AATATTATTT
 19,846,908  CCATGTATGG AAAATAACTG TTTCCTTAGT AACAATTGCA TCAAATCAGT
 19,846,958  TCACCTGCTG CCCAATAGCA ATCACAGGAT GCATTTGGGAC AAATAAATAT p.7130/rs28615996-[T/C];p.7131-[T/G]
 19,847,008  ACTGACTGCC CCACAGCCAC ATGGTCTAAG TCAGTTACTG GAGAGCTGAC
 19,847,058  TGAAGTTTGG GAAGCATATT CATCTTACGA CACTGAGACA TCCTCGGGGG
 19,847,108  GTTGCAAACA CAGGTAGTGT GAAAATTATC AGAACATCCA AGAAAAGGAA
 19,847,158  AGTTTGACTA AGTGCCGATA AGATTTATGA TGTCATGTCT GACATAGAAT p.7313/rs28645722-[G/A]
 19,847,208  TGAAACCATC ACAGAGCACA TAGAGTGGTA TATTTTCCTG TCAAATGAAA p.7388/rs28575919-[C/G]
 19,847,258  ATCATTTTCT TTAAAAGTGA AATGAAAGTC TCTAAATACA AATTTACTAG
 19,847,308  AGGATGTGTA AATTTCCTAC TTTTCATTAC ATACTCTGGA CCCAACAGAG
 19,847,358  GGAAATTGGA GCTGTCAGTG AGCCATACAT GCAATCTGGT ACAGGATCTA  p.7503/rs6999612-
[T/C];p.7512/rs3779788-[C/T]
 19,847,408  TGGATTGAAT AGACTTTTTT TCATGGAACT ACACAAAGCC AGTCTTAGTC p.7556/rs59811201-[T/C]
 19,847,458  ATTTCAAAGA AAATTTGTGG TCATTTCAAA ACCACCAGCA ATTCCAGGGA
 19,847,508  CACCAAGTTG CATAATTCTA GGGGAAAGTG GACTAAAAGT GAATGGCAGC
 19,847,558  CTCTGGAGTT ATACTGAGCA TTATTCTTAA AATGTCAATT TGGCAAATAG
 19,847,608  GTGGTAAGCG AGATCTGTCT GCCAGATTGT TCACATCATC TCTGCTTTAA
 19,847,658  AAAGATTGAT CATAGAATAT GTTAAAATAA GACCTGTGGA GAGGAGGTAT
 19,847,708  GAGCTATTTA AGGTGGAAAG GTGTGGGAGA GGGTGAAATT AGTTTTAAAT
 19,847,758  TTTCTAAAGC ACTTTTTAAC AGGAAAAGAA GTTCTTGGGC ACTGAAGGCA
 19,847,808  GAATTAGATT AAAAGTATTC AATACTCTTC CATTATCAGA GAAATAGTAA
 19,847,858  AGCTACTAGA GTGCTTTCTG GTTGGGAAGG AAGAAGGCTA CCAACATCTT
 19,847,908  CAAATAAAGA TGAGAAGTAG AGCATATTCT TGGTGGATGA ATGGATTCAC
 19,847,958  GTGTACTTTT ACTATATGCA AATGAAAAGA GCTTTAAAGA TCAATAGTTT
 19,848,008  AGCAAGACTC TTATTTGAAT AAGTGATCTT GGAGTGTTTA CCCATTGTAA
 19,848,058  CAAAACTCCT TTTCTCATAA TTAGAATCAT TAAGAAATTT CTCCTACATT p.8221/rs7000460-[A/C]
 19,848,108  TTAGATGACG GAGGTGTGTG TGCACGTGTG TGTGTATGTG TGTGTGATCA p.8250/rs59630933-[G/A]
 19,848,158  ACTCATTGCA TATTTCTTAA ACCACAATAT TGTTTAAGGA ATTTTAGAAA
 19,848,208  ATAATTACTC AATAGGAAAA TGTTGGCCAA TCCTCAGATA TTTAGATAAG
 19,848,258  GCTGATTCAA ATGCCTATTC TTTCACTGTT CTCACTATGA CACTCTTATT p.8415/rs56321069-[T/A]
 19,848,308  TTTATTATTG GCCTGACTTC CGCAGTTATT TTGAAGTTAC AGATTTTTAA p.8467-[C/T]
 19,848,358  ATTTTGAGTT GAAAAAAAAG AGCAAATTTA GATTAAGGAA TGAGAAGTAG p.8516-[delG]
 19,848,408  TCCTCGCAGC CTCATGAATC TCCTGAAATT TCGAACGGCA AAATCTAAAA
 19,848,458  TCTACAAGTT ATTACCTTCT TACAGTAAAT AGGTGGGTGT TATGGGTCGT
 19,848,508  TTTCTTTAAC TTCTTTACTT GAAAAGGAAT TAAATGATTT CCCTTTAACA
 19,848,558  TAACTTCCCT TTGATTGTGC TCTGCTTCAT GAAGTCTGAT TTTATTTGCA
 19,848,608  ATATAATTTA CTTCAACTAT TCACTGTACC CCATGAAGAT TCAGCAGCAT
 19,848,658  TTATAACTAT TGTCCATAGT TTCAAAAACT AGGTTGTCTT TCTCTTCTCA
 19,848,708  CCAAGTTTGG GATTAACTAT GAAGAACCAA AGTGAACCCT TTCAACAACA
 19,848,758  AGGTTTGCTG TGGTTTTCAA GTTTTGCCTT TGTGTGGAAC ATTGTAATGA
 19,848,808  CATAGTGGGA AAAGAAATAT TTGGGGAGAG AATTAACCAT GGCTGATACA
 19,848,858  TAGCACGGGT ATTTCTGAAC AACCTACTAA ATTATTTCTT AGAACATTTT rs73601656;p.9015/rs28445964-[A/G];
p.9024-[T/C]
 19,848,908  GAAGTATATC TTGCCATAGG AGTGGGAACA GTTTCATACA AAAGCCTCCT
 19,848,958  CATGCTTCCA ACTTTTCTTT AAAAAATTTT TTTTAAATTA TTTTATTAAA p.9130/rs13252357-[T/A]
 19,849,008  AATAGAGGCC CGGCCCGATG GCTCACACTG GTAATCCCAG CACTTTAGGA
 19,849,058  GGCTGAGGTG GGCAAATCAC TTGAGGCCAG GAGTTTGAGA ACAGCCTGGC
 19,849,108  CAACATGGTG AAACCTCATC TCTACTAAAA ATCCAAAAAT TAACCAGGCC
 19,849,158  CGGTGGCTCA TGCCTGTAAT CCCAGCATTT TAAGAGGCTG AGGCGGGTGG
 19,849,208  ATCACTTGAG CCCAGGAGAT AGCGACCACC CTGGGCAACA TGGCCAAACT
 19,849,258  TCATCTCTAC AAGAAATACA AGTTAGCCTG GCGTGGTGGC ACGCACCTGT p.9411/rs28689946-
[A/C];p.9418/rs28582042-[G/A]
 19,849,308  GGTCCCAGCT ACTCAGGAAG CTGAGGTGGG AGGATCACTT GAGTTCGAGG
 19,849,358  GTGCAGTGAA CCAAGATCGC ACCACTGCAC TCCTTTGGCC TGGGACACAG
 19,849,408  AACAAGACCC TGTCTCAAAA AACAAAACAA AACAAAACAA CCCCCCCCCC p.9589-[C/T]
 19,849,458  GCCCCACACA CACACAAATA GTGGAACTAT AGCACACAAG AGCCATGCAT
 19,849,508  GAGTCAGTGT TCTCCACGAA AGCAAGCTTC AAAGTGGAAC TGAGAGACCG p.9696/rs73667468-[G/T]
 19,849,558  GGCTCTGATC CTCACCCTCC CACTAATACC AGGGTAGCCT CAGCCAAGTC
 19,849,608  ACTTAAATTC ATGTCTTGAG AGAAGACAGA ATTAACTAAG GAATCCCCAA
 19,849,658  ACAGCAGTTC CCTTAGTTGA AATCACAAAC GTACACACAC ATACCCTGAC
 19,849,708  TAAGGAGAGC TTCAAGAAAG GCTGGAAGAC TTAGGAGAGG TCAATGGTGA
 19,849,758  CAATCTTAAT TCAGAGTTAA GGTTGTCTCT CTGTCAACTT TGCTCAACGT p.9914/rs73667469-[T/G]
 19,849,808  TGGAGCATCT GTTGTTCTCT TGCAATCCAC ATTCGTTTTC GAAAACACTT
 19,849,858  CAGAAACAAA AATAGCATCA GCGGTGGTTG CCTGTGAACC TAAAACATAT
 19,849,908  CATTCCAATG AATAAAATCA AGCAACCCTC CAGTTAACCT CATATCCAAT
 19,849,958  TTTTCCTTTC CAGAAAGAAG AGATTTTATC GACATCGAAA GTAAATTTGC Exon 2 p.10127/rs1801177-[G/A]
```

```
 19,850,008  CCTAAGGACC CCTGAAGACA CAGCTGAGGA CACTTGCCAC CTCATTCCCG
 19,850,058  GAGTAGCAGA GTCCGTGGCT ACCTGTCATT TCAATCACAG CAGCAAAACC rs11542065
 19,850,108  TTCATGGTGA TCCATGGCTG GACGGTAAGG GAGGCTCTTT GGGGAAGAGT Intron 2
 19,850,158  GGATTGGGGT GGTGAGGTAT CCTGACTGGC CTGCCCAATT GTTGGGGACC
 19,850,208  CAGTGATGGG TCCGCACCCC ACATCTCACG TGGATCTCCT TACACTTGAA rs59054859 rs11570898
 19,850,258  TAAAGACAGT TCTGGCTCAG GTGGGATCTG AAGCCACAGG TTCATGAGAA
 19,850,308  CTCCCCCTAG GCAGTGCCAG CCTTCATTTT AACACTGTAC CTGGTTGGTG
 19,850,358  CCCTTGAGCC AGAGCTTCCT GCGAGGTTGG TAAAGGATGC TCTGCCCAGC
 19,850,408  TACTGAGCAG AAGATAGGTG ATTGCTGTGG GGAACCGGTG GAACCCTGGC
 19,850,458  ATGATCCCGC ATCACCCAGC ACATTGCCAG GAGAACCTTT CTAAAGAAGA p.10632-[C/T]
 19,850,508  CAGCATGGAA GAGTGAGGAG AGGGCTGGAG TGGAGTGAGG AAGTGTGGCG
 19,850,558  TCCATGCTGG CTTTGCCATT CTCCAGCTCT GTGCTGTGGA TCAAGTCACT
 19,850,608  TGTCATCTCT GGGCCTCCAT TTACTGATCT GTAAAGCAGA GGTTGCACTA
 19,850,658  GATGTCCCTA AAAACACTCT ACTTCCAAAA TTTTCCAGTT GTAAACTTTA rs6991305
 19,850,708  GGGAGCCTTT CTGGAAATTA AAAAAAAAAA GGGCTGGGCA GGGAACTGAC
 19,850,758  ATGCTGACAT GCCAGATGAT TAGAAAAAGT GAAACTGTGA TTAAATTTAC p.10912-[A/G]
 19,850,808  TTTAAAAAAT TGTTAAAAGT CCCCTCTCTA ATATGTCACA CAGTCACTTT p.10987-[C/A]
 19,850,858  AGATTATGTA GTTCTTGCTT AGTTTGTTTT AAAATGTTTC AATCCCAAGT
 19,850,908  GAGTAGACTG TGTAATTTAT TTAAAAGGAT TTGTGTAAAA TGGCTTGTTA p.11050/rs7016529-[T/C];
p.11090/rs8176337-[C/G]
 19,850,958  AAGATATATT ATCTATCCAT TATGTCAGAG CTTGTGAATA TTATAATACT
 19,851,008  AATTTCTAGT GAGAACTAAC TAAAAATAAC ATTCATCTAG TCCACATTTC
 19,851,058  TTTTCCTACA TAATTTAAGT GGCTTTGTTA TTGACTGAAG AGGACAATCG p.11228-[T/C]

 19,851,108  CTACAATTTT TATAGAGATG AGACTACTAC TAACAGAATT CACACAGATG
 19,851,158  TTTAACAAAA TATAGCCATT TACTATGTTA ACATAGTAAA TTCAATATTT
 19,851,208  TGTGATAAAA TCTCAAATTC CTAAACATAA ATCCTTTAAT ATTTTAATAG
 19,851,258  GTGTAAGTAG GAAAAAGGAA TCTTTTTCTT TTTTGAGATG GAGGCTCACT
 19,851,308  CTATTGCCCA GGCTGGAGTG CAGTGGCGCG ATCTCGGCTC ACTGCAACCT
 19,851,358  CCGCCACCTG GGCATAAGAG ATTCTCCTGC CTCAGCCTCC CGAGTAGCTG
 19,851,408  GGACTACAGG CGCCCACCAC CACACCCGGC CAATTTTTAT ATTTTCAGCA p.11574/rs34123038-[G/A]
 19,851,458  CAGGCAGGGT TTCACCCTGT TGGCCAGGCT GGTCTCGAAC TCCTGACCTC p.11600-[G/C]
 19,851,508  AAGTGATCTA CCCACCTTGG CCTCCCAAAG TGTTGGGATG ACAGGCGAGA
 19,851,558  GCCACTGCAT CTGGCCCAGA AAAAGGAATC TTTAGAGGCC CTTTGGTTAT
 19,851,608  ATACAGAACT TGGATTTTTAA AATTTTTCAA AAAATCTGAG CTTAATAAAC p.11760/rs73667470-[A/C];rs60772381
 19,851,658  GCATTTATAC AACAGAAATA AATTGAGTAT CTCAGTCATC TCAATCTTAT
 19,851,708  CCCTGAGAGA ATTTTATACT TTGGGAAGTT GTGGGAAAAA AATGGTTCTT p.11888_11889-[InsA]
 19,851,758  TTTCATACTA AGATGACATG ACCAACCCAA TATCAACAGA CGGTGCCACT
 19,851,808  TCCTATCATT GGTCCTACTG CCTCACTAAG CCCACCTGTA TCTTTCACAT
 19,851,858  CATGTGTCCT GACATTTTGA GTGCTTGAGC ACAGAGACTG CTGTCTGGCT
 19,851,908  GTGGGACTGA GTTGGGTCTG TGCAAGAACT AAGCCAGCCA CACTGATCTT
 19,851,958  GATTATCTCA GTGAACTCAC TGGCAGGGTC AGGTGGCCCA CCTGGTATAG
 19,852,008  GCAGCAGGGA GGGCTTCAGT TCAGCTGCGT GTCTGAAAAC CAAAGATTTA
 19,852,058  AAACATAGTA ATTATTGAAC CTCAGAAGAA AAACTCAGAT TGAAAGAACT p.12224_12920-[del697}
 19,852,108  TAGAATAAGA CCCTTTTTGA GTTGAGAAAG GTGAGTACTT AGATTTTTCA
 19,852,158  TTTGCTTTGT TTGGGATTAC TTACATCAGT ATTTTATGTT GATCAGAAAG
 19,852,208  AAAGGATTCA ATTAGCTATT GTTCGGTTAA TAAAAATGTC AGCCACTGTA rs73667471
 19,852,258  GGAGTAAGTT GGATGTCCAG CCTTTTTAGA TTGCTTAACT TGGAAACACT
 19,852,308  GGGCTGGGAG CGGTGGCTCA TGCCTGTGAT CCCAGCACTC TGGGAGGCCA p.12449-[G>A];rs57357723;rs10095748
;p.12484-[C/A]
 19,852,358  AGGCAGGCAG ATCACTGGAG GTCAGGAGTT TGAGACCAAC CTGGCCAACA
 19,852,408  TGGGGAAACC CAATCTCTAC TAAAAAAATA CAAAAAAATT AGCCAGGTGT p.12550-[G/A];rs58934202
 19,852,458  GGTGGTATGT GCTTGTAGTC CCAGCTACTC AGGAGGCTGA GGCAGGAGAA
 19,852,508  TCGCTTGAAC CAAGGAGGCG GAGGTTGCAC TGAGCTGAGA TCATGCCACT
 19,852,558  GCACTCCAGC CTGGGCGAGA CAGTAAGACT CTGCCAAAAG AAAGGAAAGA
 19,852,608  AAGGAGGGAA AGAAAGGAAG GAATGAAAGG AAGGAAGGGA AGGAGGAAGG
 19,852,658  GAAGGAGGGA GGGAGGAAGG AAGGAAGGAA TGAAGGAAGG AAGGAAGGAA p.12810_12829-[dup20]
 19,852,708  TGAAGGAAGG AAGGAAGGGA GGGAGGAAGA AAGGAAGGAA GAACAAAGAA p.12853_12854-{Ins16};p.12861_12864-
[del4];p.12878_12889-[del12]
 19,852,758  AAGAGAAACA CTGGTAGTAC AGAAAAACTT CTGATAGAGG CCTAGAGTAA
 19,852,808  ACCCGATTTT CTTGCCTTAT CTGAAATAAG CTGCCTGGGG ACTCACAGGC rs11570894
 19,852,858  ACAGACGAAG GGAAATGAGG AGGCTCTCCA GCTGTGTCAT GAGACACCCA p.13003-[G/T]
 19,852,908  AAGGAATGCT TAGCATGTAG CATGCATGTG ATACATCCCA GCAGGTTGCT
 19,852,958  TAGACACAGC TATCTTGGAG CTTTGCCACT TGCTTGGATG TCACTGGCTT
 19,853,008  TAAGTACAGG GTTCCCATTG TGAAGTAGGG GATCCTGGCT GAAACAGGGA rs7002728
 19,853,058  GACATTAACA TTACATTCTG AAGAAATGAC ATCAACCTCT CCTGATCTTG
 19,853,108  AAAGCCAACT ACAAAGGGTG CCCAACACCC CAACCTTGAA GGGAGGCGAA
 19,853,158  GGTGAGTGGG ACTGGACCAA TTCAACAGGG TTCTGCTCCT AGCCAGGTGC
 19,853,208  TCCTGCTAGT TTCCTCAAAG ACCCACTTTG CATTCAGACC AATCTTTCCT
 19,853,258  TTTAATAGTA TAAATGATCA AAATTTTATT GAATGTCTAA AATATACTTT
```

```
19,853,308  TTAAATGGGA AGCATGGTGA ACCCCAATCT GCCGTTCCTC AACTCAACTC
19,853,358  AATGCCTTCC TGGCTTACTT AGATCTGCCT TGGAAGGGAC AGACCTGTCT
19,853,408  CTGAACACTG TTCTGTTATT TGATTTTTCT ATCTGTGCCA ATGGGTTTCC
19,853,458  AATCAAGTTT GTTTTTTCCA TTTCATGCAG GTGTATTGGG CTGATGTATC p.13639-[G/A]
19,853,508  TATGACAAGT GGTAGGTGGG TATTTTAAGA AAGCTTGTGT CATCATCTTC
19,853,558  AGGTAACAGG AATGTATGAG AGTTGGGTGC CAAAACTTGT GGCCGCCCTG Exon 3
19,853,608  TACAAGAGAG AACCAGACTC CAATGTCATT GTGGTGGACT GGCTGTCACG
19,853,658  GGCTCAGGAG CATTACCCAG TGTCCGCGGG CTACACCAAA CTGGTGGGAC
19,853,708  AGGATGTGGC CCGGTTTATC AACTGGATGG AGGTAAGACT GGGAGAAGGA Intron 3 p.13854/rs1121923-[G/A]
19,853,758  GACTTATGTG TCCAAAACAG TGTTTTTGAC TGGAGCCAGA AAACCGGCTG rs3735959
19,853,808  TTCTTTCTTC CTTTTCTCTT AGATTTAAAT ATTTTCTGGG GGCATTCAAA
19,853,858  TCTTCAGAAT CAGCGTGGAT ATTATTTTAT ATCCAAAAGC AACATTTTGA
19,853,908  TAAGAATAGA CTATAAGGCC AATAAATAGT CCTGCCCTGC TCTATCGTTT
19,853,958  GATATTTTCA CAGTGAATAG ATCTGCTGAG ACCAATAACT AAGTGGGCCC p.14114/rs73667472-[T/C]
19,854,008  AACAGAAAAA AGCTTGTGAT TTTCGGACAG AGGAAAGATG GCATGTTCAG
19,854,058  CCAGACCCTT CCCTACCAGT TGGCTGGCCT GTGGACATCT TATCCTCACC rs58670071
19,854,108  TTGACATACC AATCTCTTTC ATGAAAATAT TAATAGTACT TATTCTTTAG
19,854,158  TGTGAAATAG GATGAACGTT TTTGTTGAGC ATTGGGAGAG TGATGGAATT
19,854,208  GAGCTAGGAA GATGTTGGAA GGGAAGGTGC ATAGGATAGA AGGGAACTGA
19,854,258  GGTCTGGAGT TCTGACTTAG CTGCAAGAGA CCCACTTTTT CACACGATCC
19,854,308  CTTGAGAAGT ACCTCTGAAA AGTATCTTGG GGTTGGAAAG AAGCTGATAC
19,854,358  TCTGACCAAG GCAAATTATT TTAACCAGGT AATTGGAAGT AAAAAATAAG
19,854,408  CTGTGTTTAT TAGACTGATC ATAAAAGACA AAAGTTCTTT TCTTGTCTTT
19,854,458  TTTGCTGACC AGGCAAATGA ACATTGGGCAA ACGGGATCAC CTCCCTGGGG rs8176338
19,854,508  CTCAGGCTTC TCACCTGTTA AATGAGGGGC TGGACCATGT ACCTCTGGTC
19,854,558  CCTTCCACTG AATGTTTCCT GAGTCTGTCA TTGCTTGGCT AACCTTCAAT
19,854,608  GATAAAGTGA TACAGATATT TAGAGTAAGG AATAATGGGA AAATATATAC
19,854,658  CCATATACTA TACATTCAAA CATACACACA TATACATATA TGCATGCATA
19,854,708  TAAATGTATA CGCATATGTA TATGTGTATA TGTTTGTACG TATAGTATAT
19,854,758  GGTTATATAT GCAAATACAT ATACATATAC AAACATATGC ATATATTATA
19,854,808  TACATAAATA ATACTATTTC AGATGCATGG AAAAACTTTG TAATTTAAAT
19,854,858  CTGCTATTAA AGAAAGAGAA AATCAATTCT GGATTTGTTT ACGGAAAAGT
19,854,908  GAAACAAAAG AAAAAGACAA TTTTAACACT AGAGAATATT TTCTCTCTCT
19,854,958  TACCTGTAAC ACAAAATTAA AATAAGTAGA ATTAGTTTTC AGTATTTCCT rs57345602
19,855,008  ATATTTGGAA AACAATATTT ATATTCATTT TGTTTCTTTT AGTTTTATTT
19,855,058  TTGGCAGAAC TGTAAGCACC TTCATTTTCT TTTTCTTCCA AAGGAGGAGT Exon 4 p.15206/rs343-[C/A];rs247;
rs11570897;p.15245/rs248-[G/A]
19,855,108  TTAACTACCC TCTGGACAAT GTCCATCTCT TGGGATACAG CCTTGGAGCC
19,855,158  CATGCTGCTG GCATTGCAGG AAGTCTGACC AATAAGAAAG TCAACAGAAT
19,855,208  TACTGGTAAG AAAGCAATTT CGTTGGTCTT ATCATAAGAG GTGAAAAGAC Intron 4
19,855,258  TGTCATTCTG AGAGAGAATC AGAACAAATT TTGTTAAATA CCCACATGTG p.15425/rs249-[T/C]
19,855,308  TGGTGTTCTT CCCGGAGACA TGACCAGCAC TTGATTATCT CATTGTAGGG p.15449_15450-[InsTG]
19,855,358  CTCTTTATTA GGGATAAGAA AAAACACAGA CGCTCTCACT GGCTTACTAT
19,855,408  CCACTGGCAA TAGCACAGAA ATAAAGCATA ATTACACACA ATGCCTGCAG rs251
19,855,458  ATTTCTCTGG GAAGCCTGTT TCCTCCCACT CTCAGCTCTG TGTTTTAGTA
19,855,508  GTGTAAATGC ACATCAGTAC TAGGAGAAAA GAAGAAGGAC CAATTCCAGA p.15653-[delA]
19,855,558  GGCCACTTCG AAAGAAGACC GTCATCTAGG CAAAGGTGTG GCATACACAC
19,855,608  AGAGAGAAAG AACCCACCAC TGTTTATACA TCTTCTCGAC ATATTCAGAA
19,855,658  ATAATCTACA AAAGGAAATC CAGCCATCCT GAGTGGAAAC TGCTGCATAA p.15836/rs253-[C/T]
19,855,708  GGCTAGTTTA AGAGACTCAA ATTCATTTTA GAAGGAGCCA AGCCTCCTTT
19,855,758  TATGTCTCTC TAAGTAAAGA TACCATGACT GTAGAATAGG AGCTAATAAG rs11570899
19,855,808  AATCTAAATA GCTGCCAGTG CATTCAAATG ATGAGCAGTG ACATGCGAAT
19,855,858  GTCATACGAA TGGAAATTTA CAAATCTGTG TTCCTGCTTT TTTCCCTTTT
19,855,908  AAGGCCTCGA TCCAGCTGGA CCTAACTTTG AGTATGCAGA AGCCCCGAGT Exon 5
19,855,958  CGTCTTTCTC CTGATGATGC AGATTTTGTA GACGTCTTAC ACACATTCAC
19,856,008  CAGAGGGTCC CCTGGTCGAA GCATTGGAAT CCAGAAACCA GTTGGGCATG rs45607438
19,856,058  TTGACATTTA CCCGAATGGA GGTACTTTTC AGCCAGGATG TAACATTGGA
19,856,108  GAAGCTATCC GCGTGATTGC AGAGAGAGGA CTTGGAGGTA AATATTATTT Intron 5
19,856,158  AGAAGCGAAT TAAATGTGAC TCTTATCCTT AACCCTTATT GACCCAATGT p.16316/rs254-[C/G];p.16320/rs255-
[T/C]
19,856,208  CCTACTCAGT AGCTTCAAAG TATGTAGTTT TCATATACAC ATTTGGCCAA p.16386/rs256-[C/T]
19,856,258  ATTATGTTTC TGAAGAATTC TGCAATGTTC AGCATGACCA CCTTAGAGCC p.16442-[G/C]
19,856,308  AGGCAGACAG CCATTTTATC TTTTATTTAC TATACTGTAG GCTACACTGA
19,856,358  GCAGTGCACT TACAGTAGCA AGAGAAAAAG GTGGGATTTT AGACAGGAAG
19,856,408  ACTCCACTGA CCTCAATAAT GGCATCATAA AATGCTATCT GGCCACATGT p.16563-[T/A]
19,856,458  TGTCATACCT TGAATGTAGC TGCAAAGCCA ATGGAAAGAT TTTAGATGTT
19,856,508  ACTGGAACAG AAGATGTTAA TTAGGATAAA TCTTCCAAAA TGTTCAGAAC rs257;p.16671/rs258-[G/C];
19,856,558  ATAATGTTAG CTTAATGTTT TACTTTAATA ATGTTAGCTT GTGTTAAATT
19,856,608  TATGATTTTT GTTTGTTTGT TTTTTGAGAT AGAGTCTTAT TCTATTGCCC
19,856,658  AAGCTGGGGT GCAGTCACAC AATCACAGGG ACTTGCAATG TTGCCCAGGC
```

```
 19,856,708  TGGTCTCAAA CTCCTGGCCT CAAGTGATCC TCCTGCCTCA GCCTCCCAAA rs259
 19,856,758  GTTCTGGGAT TGCAGCTGTG AGCCACCACG CCCAGTTTAC GATTTATTTT rs260
 19,856,808  TAAGAGCCCC TTGCATACTT TATAGACATT GGGACCTACC TAGGATATTC
 19,856,858  TCGTTATTTT TGTGCACGTA ATAGAACTTA GAGCATATTG TTACTATTTT rs261
 19,856,908  CGATTGTCCT AAAAACTTAC AAGGAATTCA TTCTTATGGC ATTGCTGATT rs262
 19,856,958  ATTTCTATGT TCATTTGATA TAAAAGAGTG TTAGTAGGGG CAGAACCCTC
 19,857,008  AATTGTACAT AATATCAATG ATAAAATACA ATTCATTTAA CAATTACCCT
 19,857,058  CTTAAGATGT GGTTTCTAGA AATACAAATT GTCCCTAACT TACAGTTTTC p.2713/rs263-[C/T]
 19,857,108  CAACTTTACA ATTGGGCTGT AACACCATTT TAAGTTGAGA AGCACGTGAT rs9282782
 19,857,158  GGTTTGACTT AAAACTTTTT GACATTATGA TGGGTTTTGG GGGTATTAAG
 19,857,208  TGCATTTTGA CTTACAGTAT TTTTGACTTA TGAAGAATTT ATTGTAAGGC
 19,857,258  AAGGGGCAGG TATATGTTTC TAGAAGCACC TAGAAGTGTT AGACACTTTC
 19,857,308  AATGTAAGAG AAGGATGAGA TAAACAAGGA AATCAACCTC CACCTTGGAG p.17476-[A/C]
 19,857,358  GCTTATTACA GCTTCATAAA CATACTCATA AATATAAGAA GCACAAAAGT
 19,857,408  CAAAAATTCC CTGTGAACTT GCAACTTTCA CTCTCTTGAA GGTGGGTGGG
 19,857,458  CCGCTACCAC CAAGAATATC TCCTGAAATA GGGCCTACAA TCATAAATGC p.17599/rs264-[G/A]
 19,857,508  ACAGGACTAT ATCCTTGGGT GATTCTACTC TAACACCACA TCTCACCTAT rs265
 19,857,558  TTTAGACATG CCAAATGAAA CACTCTTTGT GAATTTCTGC CGAGATACAA rs266;rs267
 19,857,608  TCTTGGTGTC TCTTTTTTAC CCAGATGTGG ACCAGCTAGT GAAGTGCTCC Exon 6
 19,857,658  CACGAGCGCT CCATTCATCT CTTCATCGAC TCTCTGTTGA ATGAAGAAAA rs28934893;rs35414700
 19,857,708  TCCAAGTAAG GCCTACAGGT GCAGTTCCAA GGAAGCCTTT GAGAAAGGGC rs1800011
 19,857,758  TCTGCTTGAG TTGTAGAAAG AACCGCTGCA ACAATCTGGG CTATGAGATC
 19,857,808  AATAAAGTCA GAGCCAAAAG AAGCAGCAAA ATGTACCTGA AGACTCGTTC p.17948/rs268-[A/G]
 19,857,858  TCAGATGCCC TACAAAGGTA GGCTGGAGAC TGTTGTAAAT AAGGAAACCA Intron 6 rs59184895
 19,857,908  AGGAGTCCTA TTTCATCATG CTCACTGCAT CACATGTACT GATTCTGTCC p.18065-[T/G];p.18086/rs269-[T/G];
rs2075651; p.18095/rs270-[C/A]
 19,857,958  ATTGGAACAG AGATGATGAC TGGTGTTACT AAACCCTGAG CCCTGGTGTT p.18121/rs271-[G/A]
 19,858,008  TCTGTTGATA GGGGGTTGCA TTGATCCATT TGTCTGAGGC TTCTAATTCC
 19,858,058  CATTGTCAGC AAGGTCCCAG TGCTCAGTGT GGGATTTGCA GCCTTGCTCG
 19,858,108  CTGCCCTCCC CTGTAAATGT GGCCATTAGC ATGGGCTAGG CTATCAGCAC
 19,858,158  AGAGCTCAGA GCTCATTTGG AACCATCCAC CTCGGGTCAA CAAACTATAA p.18297-[A/C]
 19,858,208  CCCTTGTGCC AAATCCAGCC TACTTCCTGC TTTTGTAAAT AGTTTTTTTA rs272;p.18395_18396-[InsT]
 19,858,258  AAACTTTTAA GTTCAGGGGT ACGTATGTAG GTTTGCTAAA AAGGTAAACT
 19,858,308  TGTGACATGG GAGTTTGTTG TCCAGAATAT TCCATCACCC AGGTATTAAG p.18462-[T/G]
 19,858,358  CTTAGTACCC ATTAGTTACT TTTCCTGAAG CTCTCCCTCC TCCCACCCTC
 19,858,408  TGGGAGGCCC CAGTGTCTGT TGTTCCCCTC TATGTGCTCA TGCAAAGTTT
 19,858,458  TATTAGGACA CAGCCACACA CATTCATTAC CATATTGTCA AAGGCTGGTT p.18621-[C/T]
 19,858,508  TCATGCCACC ATAACAGAGT TGATAGCCCA CAGAGCCTAA AATATTTACT rs275
 19,858,558  CCCTGGCCCT TTACAGAATG TTCACAACTT ACATAAAGGC AAGGACCATC p.18708/rs276-[T/C]
 19,858,608  TGTCTTATTT ATTTATTTAT TTAATTTGAG ATGAAGTCTA GCTTTCTCCT
 19,858,658  AGGCTGGAGG AGAGGGGCAT GATCTTGGCT CACCACAACC TCTGCCTCCC p.18822/rs277-[T/C]
 19,858,708  GGGTTCAAAT GATTCCCCTG CCTCAGCCTC CGGAGTAGCT GGGATAACAG
 19,858,758  GCATGCACCA TCATGCCCAG CTAATTTTTG TATTTTTAGT AGAGAGGGGG p.18942/rs278-[G/A]
 19,858,808  TTTCACCGTG TTGACCAGGC TGGTCTCGAA CTGCTGACCT CAGGTGATCT
 19,858,858  GCCCTCCTTG GCCTCATCTG TCTTTTTAAA TGCAACTATT CCTGGAAGGC
 19,858,908  AAGAATATCT CACACCTTCT AAGATACTGC CATTTTGCCA GGAGTTTGTT
 19,858,958  TCACACTTGA ATTTCAAGCT TGGCCTCTTG TTTAGAGGCA GACCTAAAGG rs279
 19,859,008  AATGGTCGGA AAATGAGAGA GGAGGTCTTC GGATAAATCC GGTGAGAGGG
 19,859,058  ACCAACTTCA GGAAGGGTGG CTTTTGTGGA ATCCAGATGG AAACCTGAGG
 19,859,108  GAAGGGATGA TATTAAAGAA CAGTGGCCCC AGGTAAAACA TATGGCACCC
 19,859,158  ATGTGTAAGG TGATTCTTAG AATCTGTAGA GGTGTCTTTC GTGGTATAGA rs280
 19,859,208  GGTTGAGGCA CCTGTGCTTC AAGGAAACCT TAACTCTTCA AAATCAGGCA
 19,859,258  ATGCGTATGA GGTAAAGAGA GGACTGTGGG ACCATAATCT TGAAGACACA rs17091775;p.19442/rs281-
[A/T];p.19445/rs282-[C/G]
 19,859,308  GACAGGCTTC ACTCATCCCT GCCTCCTGCA CCAGTGGGTT CAAGGCTCTG
 19,859,358  TCAGTGTCCC CTAGGGGCAC CTCACCACTC CCAGCTTCTT CAGCTCTGGC p.19517/rs283-[C/T];rs284
 19,859,408  CTGTCCTGCT GCCTGCAAGG GTTTTGCTTA ATTCTCAATT CAATGTCTCT
 19,859,458  TCATCTTTTA GCAGCTGTGG GGTTTTGTTG TTGTTCTTCT GTTTTTGCTT p.19608/rs285-[C/T]
 19,859,508  AGTATCTGAC TACTTTTTAA TTATAAAAAG AGATGTATCT AAACAAAATA p.19675/rs286-[A/T]
 19,859,558  GAGATTGTTA TCAGAAGTTC ACAACATTTA TTAAAAATTT TTTCACCTGG
 19,859,608  ACAAGAGTCT AAAGCAGCAT AAAAATATGG TCTGCTATAT TCTAAACCAT
 19,859,658  CAGTCTTAAG AGATCTGTGT CTCAGCTTAA GAGAAAATAC ATTTAATAGA p.19815-[G/A]
 19,859,708  CAGTAACACA AATAAGAAAA AAATCTGACC AAGGATAGTG GGATATAGAA
 19,859,758  GAAAAAACAT TCCAAGAATT ATTTTATTTA TTTATTTATT TATTTATTTA
 19,859,808  TTTATTTATT TATTTATTTT TGAGACACAG TCTCGCTCAG TTACCCAGGC p.19975/rs287-[A/G]
 19,859,858  TGGAGTGCAG CGGCGCAATC TTAGCTCACT GCAACCTCTG CTTTCCGGTT rs288;rs2583659;p.20038/rs289-[T/C]
 19,859,908  CAAGCGATTC TCCTGCCTCA GCCTCCTGAG TAACTGGGAT TACAGGCACC p.20080-[C/T]
 19,859,958  CGCCACCACG CCCAACTAAT TTCTGTATTT TTCTTAGTAG AAACAGGGTT rs34951282
 19,860,008  TCACCATGTT GGCCAAGCTA GTCTCAAACT CCTGACCTCA GGTGATTCAC rs290
 19,860,058  CCACCAAGGC CTCCCAAAGT GCTGGGATTA CAGGCATGAG CCACCATGCC
```

```
 19,860,108  TGGCCTCCAA GAACTCTTTT TTCCTCCATC ATCATGGTTC TATTTTAGTC rs2698206;p.20271/rs291-[T/C]
 19,860,158  CTGCTGCCTT TCCTTTTAAC CTCTCCCCAG GCCCATTTGC TCAGGGTTTT
 19,860,208  TGGTAGAGAC CAGAGGAGGG GCAGGGAGGA GATATAGAAG TTCAACTACC p.20363-[A/T]
 19,860,258  TGCTTCCAGA GGCTGTCCCT AGTATAGAAT ACTTTAGGGG CTGGCTTTAC
 19,860,308  AAGGCAGTCC TTGTGGCCTC ACTGATGGCT CAATGAAATA AGTTCTTTTT rs292
 19,860,358  TAAAAAAAAT TTTATTTATT TCCATAGGTT ATTGGGGGAA CAGGTGGTGT p.20505_20506-[InsA];p.20544/rs294-
[T/C]
 19,860,408  TTGGTTACAT GAGTAAGTTC TTTAGTAGTG ATTTGTGAGA TTTTGGTGTG
 19,860,458  CCCATTACGG AATGGAAAAA TCAACGAAAT AAGTTCTATG ATGCACCTAC
 19,860,508  TAGACACCTA ATCTGCGCTA GATGGTGGGG GAATTAAGAG CATGGGCATG p.20657/rs295-[A/C];p.20663/rs296-
[G/A]
 19,860,558  ATCCTGTGAC CGGAAGCCCG CTTACAGTCA GGGTGGAGGA CAGACCTACT
 19,860,608  CATGAAACAA ACACAGTGAC ATATAGTGAC ACAGAAGCAA ATGTCAAATA p.20790/rs297-[T/C]
 19,860,658  TGCTTGCTCC AGATGCTAAG GCACAAGATG GCCAAGGATG GCGGAGTTCA
 19,860,708  TGGAGAAAGC ATCATGAGTG TTTTGGCCTT CTGATTTGAT CTCCCTAGCA
 19,860,758  CCCCTCAAAG ATGGCTACTT CCTAATGCTG CTTGGCAATT CAGACACATT
 19,860,808  TGGGTTTTTC CTATGCATAT AACCACACTT TTCTGAAAGG GAGTAGAATT
 19,860,858  CAAGGTCTGC ATTTTCTAGG TATGAACACT GTGCATGATG AAGTCTTTCC
 19,860,908  AAGCCACACC AGTGGTTCCA TGTGTGTGCA CTTCCGGTTT GAGTGCTAGT
 19,860,958  GAGATACTTC TGTGGTTCTG AATTGCCTGA CTATTTGGGG TTGTGATATT p.21125_21128-[delGACT];rs73601683
 19,861,008  TTCATAAAGA TTGATCAACA TGTTCGAATT TCCTCCCCAA CAGTCTTCCA Exon 7
 19,861,058  TTACCAAGTA AAGATTCATT TTTCTGGGAC TGAGAGTGAA ACCCATACCA
 19,861,108  ATCAGGCCTT TGAGATTTCT CTGTATGGCA CCGTGGCCGA GAGTGAGAAC rs298
 19,861,158  ATCCCATTCA CTCTGTGAGT AGCACAGGGG GGCGGTCATC ATGGCACCAG Intron 7 rs299 rs300
 19,861,208  TCCCTCTCCT GCCATAACCC TTGGTCTGAG CAGCAGAAGC AGAGAGCGAT p.21353/rs301-[T/C];rs302;rs34500595
 19,861,258  GCCTAGAAAA CAAGTCTTTA GTTAAAAAAA TCAGAATTTC AAAATTGAGG
 19,861,308  TCTTTCCTCT ATTTGATATT GAGAAAAAAA TGCTTCAAAT TGGCCATTTT
 19,861,358  ATTTTCACTT ACTAGTTATA TTTTTTTATT TATCATCTTA TATCTGTTTA
 19,861,408  TTTCTTTTAT AAAGCTGCTG TTAAACAATA TAATTAAACT ATCTCAAAAG
 19,861,458  GTTTGACATT AAAGAAAATG AGCAATGGTA ACAGGAAACC ACTCTATAGA
 19,861,508  TGTACATATA ATATGTACAG AAAATATAAG TAGTAAGAAG TCCATGACAA
 19,861,558  AGTGTTAGCT CTTTTTTTTT TTTTTTTTT TTTTTTTTT GAGATGGAGT rs303;rs13257355
 19,861,608  CTCTCTCTAT TGCCCAGGCT GGAGTGCAGT GATTCGATCT CAGCTCACTG p.21780/rs304-[T/G]
 19,861,658  CAACCTCTAC CTCCCGAGTT CAAACAATTC TTCTGTCTCA GCCTCCCGAG p.21820/rs305-[A/G]
 19,861,708  TAGCTGGGGC TGCAGGTGCC CACCACCATG CCCAGCTAAT TTTTGTATTT rs306;rs307; p.21895/rs308-[T/G]
 19,861,758  TTAGTAGCGA CAGGGTCTCA CCATGTTGGC CAAGCTGGTC TTGAATTCCT rs309
 19,861,808  GATCTCAGGT GATCCACCCG CCTCGGCCTC CCAAAGTGCT GGGATTACAG p.21965/rs310-[C/T]
 19,861,858  GTGTGAGCCA CCATGCCCAG CCTACCCTTT ACTACTAATC AAAGAAATAA p.22044_22047_[delTAAA]
 19,861,908  AAGTAAGGCA ACTTGATACT TTTACAATTA CTAGATGAAC AAATCTTTAA
 19,861,958  AAATAGCCAG TGCAGACAAG GTGGTGAAGC AGAACATGCG AACCTACCAT
 19,862,008  GCATCATTCA CGGCTAGAAC CCTCCAGGTG CGGAAGGTAG TATTTTAATA
 19,862,058  ACTTTCCATA GCTACAAAAT ATTATTACAT AGAAGGGAGT GATTTTTTTC
 19,862,108  TAATATTTAT CCTAAAGAAA TAGTCAACAA ACATTTTTAA AAACATCAAT
 19,862,158  TACAGTCGTA CCTATACTAG CATAAATTAG AAACCCAGTA TCCAACATTG
 19,862,208  AGGCAGTGGG TAAATGAATC GTGGTTTATC AAGTCATTAA AATCAATCTA
 19,862,258  GCCTTTAAAA ACTATAATTG TAGGAAACCC AGGAAAACAT AGTAAAAAAT p.22416/rs312-[G/C];rs313
 19,862,308  GGAATATAAA ATCTGAAGAG AATAAAGAAT AGAGAATCGT ATGTGTGCTA p.22461/rs314-[G/A]
 19,862,358  TGATTGTAGC TAAATAATGT TCAAGTATCA ACACAAATTG AAAAGGAATA p.22514/rs315-[T/C]
 19,862,408  CATGAAAATG AAAATTATAT TTCTGAATGA TTGACTTCAG GATTTTCTTT
 19,862,458  TAGAATTGTA TTAAATAGTT CATGTCATTA GGATAAATGC TGGAATGTGG
 19,862,508  ATATAATTTA AAATATACTA AATGCCATCG ACCTTCATTT TGAGTTCTTT
 19,862,558  GTTGGACATT TTTGTGCATT TTTAAAATAT CCCCTAAATA ATAAAGCTAT
 19,862,608  TTATATTTGG AGAGGAGAAA AAAAAGTGGG GGGCAGGGAG AGCTGATCTC
 19,862,658  TATAACTAAC CAAATTTATT GCTTTTTTGT TTAGGCCTGA AGTTTCCACA Exon 8
 19,862,708  AATAAGACCT ACTCCTTCCT AATTTACACA GAGGTAGATA TTGGAGAACT p.22855/rs316-[C/A]
 19,862,758  ACTCATGTTG AAGCTCAAAT GGAAGAGTGA TTCATACTTT AGCTGGTCAG
 19,862,808  ACTGGTGGAG CAGTCCCGGC TTCGCCATTC AGAAGATCAG AGTAAAAGCA rs5934
 19,862,858  GGAGAGACTC AGAAAAAGTA ATTAAATGTA TTTTTCTTCC TTCACTTTAG Intron 8
 19,862,908  ACCCCCACCT GATGTCAGGA CCTAGGGGCT GTATTTCAGG GGCCTTCACA
 19,862,958  ATTCAGGGAG AGCTTTAGGA AACCTTGTAT TTATTACTGT ATGATGTAGA
 19,863,008  TTTTCTTTAG GAGTCTTCTT TTATTTTCTT ATTTTTGGGG GGCAGGGGGG p.23190_23191-[delAG];p23192-[G/T]
 19,863,058  GGGAAGTGAC AGTATTTTTG TATTTCATGT AAGGAAAACA TAAGCCCTGA
 19,863,108  ATCGCTCACA GTTATTCAGT GAGAGCTGGG ATTAGAAGTC AGGAATCTCA
 19,863,158  GCTTCTCATT TGGCACTGTT TCTTGTAAGT ACAAAATAGT TAGGGAACAA
 19,863,208  ACCTCCGAGA TGCTACCTGG ATAATCAAAG ATTCAAACCA ACCTCTTCAA p.23388/rs318-[C/G];p23395/rs319-[A/C]
 19,863,258  GAAGGGTGAG ATTCCAAGAT AATCTCAACC TGTCTCCGCA GCCCCACCCA
 19,863,308  TGTGTACCCA TAAAATGAAT TACACAGAGA TCGCTATAGG ATTTAAAGCT p.23496/rs320-[T/G]
 19,863,358  TTTATACTAA ATGTGCTGGG ATTTTGCAAA CTATAGTGTG CTGTTATTGT
 19,863,408  TAATTTAAAA AAACTCTAAG TTAGGATTGA CAAATTATTT CTCTTTAGTC p.23573-[T/C]
 19,863,458  ATTTGCTTGT ATCACCAAAG AAGCAAACAA ACAAACAAAA AAAAAAAGAA rs321;p.23636/rs322-[A/C];rs323
```

```
19,863,508  AAAGATCTTG GGGATGGAAA TGTTATAAAG AATCTTTTTT ACACTAGCAA
19,863,558  TGTCTAGCTG AAGGCAGATG CCCTAATTCC TTAATGCAGA TGCTAAGAGA
19,863,608  TGGCAGAGTT GATCTTTTAT CATCTCTTGG TGAAAGCCCA GTAACATAAG p.23747/rs325-[T/C]
19,863,658  ACTGCTCTAG GCTGTCTGCA TGCCTGTCTA TCTAAATTAA CTAGCTTGGT
19,863,708  TGCTGAACAC CAGGTTAGGC TCTCAAATTA CCCTCTGATT CTGATGTGGC p.23858/rs326-[A/G];rs7005541
19,863,758  CTGAGTGTGA CAGTTAATTA TTGGGAATAT CAAAACAATT ACCCAGCATG
19,863,808  ATCATGTATT ATTTAAACAG TCCTGACAGA ACTGTACCTT TGTGAACAGT p.23955/rs327-[T/G]
19,863,858  GCTTTTGATT GTTCTACATG GCATATTCAC ATCCATTTTC TTCCACAGGG Exon 9
19,863,908  TGATCTTCTG TTCTAGGGAG AAAGTGTCTC ATTTGCAGAA AGGAAAGGCA
19,863,958  CCTGCGGTAT TTGTGAAATG CCATGACAAG TCTCTGAATA AGAAGTCAGG p.24143/rs328-[C/G]
19,864,008  CTGGTGAGCA TTCTGGGCTA AAGCTGACTG GGCATCCTGA GCTTGCACCC Intron 9
19,864,058  TAAGGGAGGC AGCTTCATGC ATTCCTCTTC ACCCCATCAC CAGCAGCTTG
19,864,108  CCCTGACTCA TGTGATCAAA GCATTCAATC AGTCTTTCTT AGTCCTTCTG
19,864,158  CATATGTATC AAATGGGTCT GTTGCTTTAT GCAATACTTC CTCTTTTTTT
19,864,208  CTTTCTCCTC TTGTTTCTCC CAGCCCGGAC CTTCAACCCA GGCACACATT
19,864,258  TTAGGTTTTA TTTTACTCCT TGAACTACCC CTGAATCTTC ACTTCTCCTT
19,864,308  TTTTCTCTAC TGCGTCTCTG CTGACTTTGC AGATGCCATC TGCAGAGCAT
19,864,358  GTAACACAAG TTTAGTAGTT GCCGTTCTGG CTGTGGGTGC AGCTCTTCCC p.24505/rs329-[A/G]
19,864,408  AGGATGTATT CAGGGAAGTA AAAAGATCTC ACTGCATCAC CTGCAGCCAC p.24573-[T/C]
19,864,458  ATAGTTCTTG ATTCTCCAAG TGCCAGCATA CTCCGGGACA CACAGCCAAC
19,864,508  AGGGCTGCCC CAAGCACCCA TCTCAAAACC CTCAAAGCTG CCAAGCAAAC rs12544438
19,864,558  AGAATGAGAG TTATAGGAAA CTGTTCTCTC TTCTATCTCC AAACAACTCT
19,864,608  GTGCCTCTTT CCTACCTGAC CTTTAGGGCT AATCCATGTG GCAGCTGTTA
19,864,658  GCTGCATCTT TCCAGAGCGT CAGTACTGAG AGGACACTAA GCATGTGACC p.24815/rs330-[G/A;p.24824/rs331-[G/A]
19,864,708  TTCACTACTC CTGTTCTGAA TTCCAGGAAT ATGCCCTTTT CAACCCTCCA p.24852/rs12679834-[T/C]
19,864,758  CACATCCCCT GCCAGACAGC AAGTGCTAAT GGGTTACAGG AACAAAGGGG p.24899-[C/T]
19,864,808  AGAAATATTA GATCATGTCA TACAAGCCAG TGACACAAGA AATGAAGGGA
19,864,858  AAGGCTAGAC ACAGTGTCAT CTGGAAACAG GAAAAGCAAT TGCTTTTGGT p.25005-[A/G]
19,864,908  TTGTTCTTTT CCTAGTTTGC ATTTGGGACA AATGTATAGA ATAAGAATTG p.25049-[G/A];rs28681081
19,864,958  CCTTCATGCC TGCAATCCCA GCACTTTGGG AGGCTGAGGC AGGTGGATCA
19,865,008  CCTGAGGTCA GGAGTTTGAG ACCAGCCTGG CCAACGTGGC GAAACCACCT
19,865,058  CTCTACTAAA AATATAAAAA TTAGCTGGGT GTGGCGGCAC ATGCCTGTAA
19,865,108  TCCCAGCTAC TCGGCAGGCT GAGGCGGGAG AATTGCTTGA ACCGGGGAGG
19,865,158  CAGAGGTTGC AGTGAGATGA GATCGCGCCA TTATATTCCA GCCTGGGCAA p.25320-[C/T];p25335-[C/T]
19,865,208  CAGAGAAAGA CTCCATCTCA AAAAAAAAAA AAACATGCCT ATTAGGAAAA p.25352-[A/C];rs28578146
19,865,258  GTATATTAAA GACCCTATGT GTAACATCTT TAATGTTTTT AAATTCTACT rs28599962
19,865,308  TTATAATAGA TTTTATACAT GTTTACTATA AATAGATTAG GAAAAATAAG rs13261181
19,865,358  CAAAAATAAA ATAAAATCAC TGTGACCATA TCACTCAGAG ACAACCCCAA
19,865,408  TTAACGTTTT TATTTATATT CTTTCGGACT TTATATATAC ATAATATTTA
19,865,458  TATGTTTTTC GTCCTTTACA AAAATAGAAT TATGGTGTAT ATACTCTGAA
19,865,508  TGACTAGATG AGAACATCTG GATCAAAAGC ATTAATGTAA GAGCATTCAG rs17116619 rs28439839
19,865,558  GATAAACTCA AAATGGAGAA TAGTTAGTGG TATTGAGCCA GGCAAAATAA
19,865,608  CGCAATTCTT ATCTAACTGG AGACTTTTCT TCTAAGAGGT TATTACGTTG
19,865,658  TTTTTCCTCA TCACAAATCT GAGGCAATAT CATACTTTCT TCAGTTCTTA rs28424158;p25844-[T/G]
19,865,708  GAAAGAGACT TTTAGATGAA GTTTTTTTTG TTTGTTTTGG TTTTTTTTTT rs28716400
19,865,758  CTTGAGATGG AGTTTTGCTC TTGCTGCCCA GGCTGGAGTG TAGTGGCTCG
19,865,808  ATCTCAGCTC ACTGCAACCT CCACCTCCTG GGTTCAAGCA ATTCTCCTGC
19,865,858  CTCAGCCTCC CAAGTAGCTG GGATTACACG TGTCCGCCAC CACACCTGGC
19,865,908  TAATTTCGTA TTTTTAGTAG AGAAAGGGTT TCACCATGTT GGTCAGGCTG rs4416836
19,865,958  GTCTTGAACT CCTAACCTCA GGTTATCCAC CTGCCTCGGC CTCCCAAAGT
19,866,008  GCTGGGATTA TAGGTGTGAG TCACCACACC CGGCCCTAGA TGCAGTTTTA
19,866,058  TACATGCATT TGTATTACAC ATAAATAGCA TGCATATTCT GCCAGAGCAT p.26201-[T/G];p.26234/rs10099160-[T/G]
19,866,108  CTACAACTTT AAATCTACAT GTGAATGTGA AAATAAAACC TCATTAAATT
19,866,158  AGTAAATAAC TCTAGCTGCT TGTAAAGCAC GTCCAGTCGT ATTTTTTATA
19,866,208  TGTTACAAGA CTTTATCTGA GAAAGCCTAA TGAAGCATTC CTTGTCTGAT
19,866,258  TATAGGATTA CTGACAGAAC AGTTATTTAG ACAGAGAATG TTCAGATGCG
19,866,308  TTTTATTTTT ATTTTTTACT TTTATTTATT TTTGAGACAG TCTCGTTCTG
19,866,358  TTGCCCAGGC TAGAATGTGG TGGCGTGATC TCGGCTCAAT GCAACTCTGC
19,866,408  CTCCCGGGTT CAAGTGATTC TTGTGCCTCA GCCTGACAAG TAGCTGGGAT rs58844409
19,866,458  TATAGGTGCC CGTTACCATG CCCAGCTAAT TTTTCTGTTT TTAGTAGAGA
19,866,508  CGGAGTTTCA CCATATTGGC CAGGCTGGTC ATTGAACTCC TAACCTCAGG
19,866,558  TGATGTGCCT GTCTCAGCCT CCCAATGTGC TGGGATTACA GGCATGAGCC
19,866,608  ACAGCACCCA GCCAGATGCA TTTTTAAAAA CGTACCTGAA CTTTATCTAG
19,866,658  GAGGTAATTA TAAATTAGAC TAATAATCTT CTACAGTTTC TTTCTTCTGT
19,866,708  GATTAAAATC AATCAAATCA AAGATTCTCT TTCTCACACC TTCTGCTAAC
19,866,758  TCCTCAGAAA CCTCATATCA CAAGAAATGA AATGGAACAG GCCTTTCGTT
19,866,808  TGATACATTT TAGAATAAGA AATCCTCTAA ATTTAGAAGT CATTTGGCCC
19,866,858  AGTCCTCCAA AAATGATGCA CCTTATTGGG ACGGGGCTAA ATAGTTGCTC p.27000-[C/T]
19,866,908  CAGTGTCTTC CATTCCTACA AACCTGCCAT TCTCTGATCC ATTATACACA
19,866,958  TCTCCCCTGG GTTTATTCTC ACAACCTTTG TTCTGAAATT CCATTTGAAG rs10283151
```

```
 19,867,008  GCTTTTTCCA TCCTAAAACC AGTGGGGGAC AGGCGGGAAT TGTAAAACAC p.27160-[T/A]
 19,867,058  TCAGAAGATA ATAAATTGCC CTTTTTCCTG TGCTTTTTCT CAGAaactgg Exon 10 p.27229/rs11570891-[C/T]
 19,867,108  gcgaatctac agaacaaaga acggcatgtg aattctgtga agaatgaagt p.27249/rs4922115-[G/A];rs7818177
 19,867,158  ggaggaagta acttttacaa aacataccca gtgtttgggg tgtttcaaaa
 19,867,208  gtggattttc ctgaatatta atcccagccc tacccttgtt agttatttta
 19,867,258  ggagacagtc tcaagcacta aaaagtggct aattcaattt atggggtata
 19,867,308  gtggccaaat agcacatcct ccaacgttaa aagacagtgg atcatgaaaa
 19,867,358  gtgctgtttt gtcctttgag aaagaaataa ttgtttgagc gcagagtaaa
 19,867,408  ataaggctcc ttcatgtggc gtattgggcc atagcctata attggttaga
 19,867,458  acctcctatt ttaattggaa ttctggatct ttcggactga ggccttctca p.27611/rs3289-[T/C]
 19,867,508  aactttactc taagtctcca agaatacaga aaatgctttt ccgcggcacg p.27688-[C/T]
 19,867,558  aatcagactc atctacacag cagtatgaat gatgttttag aatgattccc
 19,867,608  tcttgctatt ggaatgtggt ccagacgtca accaggaaca tgtaacttgg p.27783-[A/T]
 19,867,658  agagggacga agaaagggtc tgataaacac agaggtttta aacagtccct rs11542064
 19,867,708  accattggcc tgcatcatga caaagttaca aattcaagga gatataaaat
 19,867,758  ctagatcaat taattcttaa taggctttat cgtttattgc ttaatccctc
 19,867,808  tctccccctt ctttttttgtc tcaagattat attataataa tgttctctgg
 19,867,858  gtaggtgttg aaaatgagcc tgtaatcctc agctgacaca taatttgaat p.28036/rs11570892-[A/G]
 19,867,908  ggtgcagaaa aaaaaaaaga aaccgtaatt ttattattag attctccaaa rs1059497;p.28067/rs3208305-[A/T];
rs11570896;p.28093/rs1803924-[C/T]
 19,867,958  tgattttcat caatttaaaa tcattcaata tctgacagtt actcttcagt
 19,868,008  tttaggctta ccttggtcat gcttcagttg tacttccagt gcgtctcttt rs11570893
 19,868,058  tgttcctggc tttgacatga aaagataggt ttgagttcaa attttgcatt
 19,868,108  gtgtgagctt ctacagattt tagacaagga ccgtttttac taagtaaaag
 19,868,158  ggtggagagg ttcctggggt ggattcctaa gcagtgcttg taaaccatcg
 19,868,208  cgtgcaatga gccagatgga gtaccatgag ggttgctatt tgttgttttt p.28382/rs1059507-[C/T]
 19,868,258  aacaactaat caagagtgag tgaacaacta tttataaact agatctccta p.28407-[C/A]
 19,868,308  tttttcagaa tgctcttcta cgtataaata tgaaatgata aagatgtcaa p.28464/rs3735964-
[C/A];p.28490/rs3200218-[A/G]
 19,868,358  atatctcaga ggctatagct gggaacccga ctgtgaaagt atgtgatatc p.28524-[C/T]
 19,868,408  tgaacacata ctagaaagct ctgcatgtgt gttgtccttc agcataattc
 19,868,458  ggaagggaaa acagtcgatc aagggatgta ttggaacatg tcggagtaga rs1803923
 19,868,508  aattgttcct gatgtgccag aacttcgacc ctttctctga gagagatgat rs58998793
 19,868,558  cgtgcctata aatagtagga ccaatgttgt gattaacatc atcaggcttg
 19,868,608  gaatgaattc tctctaaaaa taaaatgatg tatgatttgt tgttggcatc
 19,868,658  ccctttatta attcattaaa tttctggatt tgggttgtga cccagggtgc
 19,868,708  attaacttaa aagattcact aaagcagcac atagcactgg gaactctggc
 19,868,758  tccgaaaaac tttgttatat atatcaagga tgttctggct ttacatttta p.28911/rs13702-[T/C]
 19,868,808  tttattagct gtaaatacat gtgtggatgt gtaaatggag cttgtacata p.28982/rs1059611-[T/C]
 19,868,858  ttggaaaggt cattgtggct atctgcattt ataaatgtgt ggtgctaact rs17091815;p.29046_29047-[InsTT]
 19,868,908  gtatgtgtct ttatcagtga tggtctcaca gagccaactc actcttatga p.29086/rs15285-
[C/T];p.29088/rs3866471-[C/A]
 19,868,958  aatgggcttt aacaaaacaa gaaagaaacg tacttaactg tgtgaagaaa
 19,869,008  tggaatcagc ttttaataaa attgacaaca ttttattacc aca
 19,869,051  ctaagtcatt attttgtatc atttttaaag taaatttatt cttaggtcag rs71510671
 19,869,101  attcactcag catattttga ctaagtaacc actgtactta gtaaaccgaa p.29287/rs3916027-[G/A]
 19,869,151  gagcttctga gaattatagt gtacctatta gatatttttа acatttatat p.29315/rs9644636-[T/G]
 19,869,201  ttgtataaag ctaaagaaag ccttacatat cctttaaact gactatagaa
 19,869,251  gaaaatgata cagaattttg cctgcataaa gtacacagga ctattcttgc
 19,869,301  ctacaatatg ctttttcaca agcaaaatgt tagactaata taaggcatct p.29474-[C/T];p.29487/rs4921683-[T/A]
 19,869,351  ttggccattt tatagtgtac atcatctcta tttctgaggc ctcattgtta
 19,869,401  gctgtaacgc aagtagcatt tgtgcaataa aatgaactat ttgggatggg p.29547/rs4921684-[C/T];p.29557_29558-
[InsA]
 19,869,451  agggtacatt ttttagaact ttgctttggg ttgccttgat aattaatagc
 19,869,501  atatagtcca tttatgcagc taagtaggga ttgcttctta gtacagtcag
 19,869,551  gaagaattta gcccagaaaa caattattc aatggccact gacccaaact p.29716-[T/C]
 19,869,601  tccaggctga agagcaatgg cgtgatcatg gctcactgca cctccacctc
 19,869,651  ccaggctcaa gtgattctcc tgcctcagcc tcccaagtag atggtactac
 19,869,701  aagcacacgc cactgcaccc agctaatttt tgtatttttt gtagagatgg
 19,869,751  gggtttcacc atgttgccca ggctggtctt aaattcctgg cctcaagtgt
 19,869,801  ctgccccccc tggcctccca aagtgctgga attacaggca tgagccacca
 19,869,851  tgtccagcct tgacccaaac ttttattgtc agttagctat tggggcttct
 19,869,901  ggagtttggg tctcccctga caggaggggg ctccccagtt cacacttggc
 19,869,951  cactgcccat caattcctgt tgatatgatc aacaagatag acaattgcaa
 19,870,001  atgttgctga ggatgtggag aagtgtgaac ctgtgtaagt ggctgatggg
```

**Figure 4.** *LPL* Annotated Sequence

## 3.2 DISTRIBUTION OF *LPL* VARIANTS IN HIGH AND LOW HDL-C GROUPS

Of the 178 identified variants (excluding microsatellite), 88 had MAF ≥0.05 and 9 1 had a MAF<0.05. The MAF range for the 74 variants which have not been previously identified in the dbSNP build 130 was 0.005-0.500%. Among common variants, 32 had a statistically significant difference (P-value<0.05) when comparing the allele frequencies between the high HDL-C and low HDL-C groups; the P-value range was 0.048-0.002. Figure 6 shows the LD pattern of these significant variants in Section 3.3. In addition to 32 variants, 5 of the common variants have a P-value of between 0.05 and 0.10.

Of the 91 relatively uncommon or rare variants (MAF<0.05), 21 were present only in the low HDL-C group and 25 were present only in the high HDL-C group and the remaining were present in both groups. Forty of 47 (85.1%) individuals with high HDL-C had minimum one rare variant versus 35 out of 48 (72.9%) individuals with low HDL-C; 23 of 47 (72.9%) individuals high HDL-C had minimum two rare variants versus 16 out of 48 (33.3%) individuals with low HDL-C; 17 of 47 (36.1%) individuals with high HDL-C had minimum three rare variants versus 8 out of 48 (16.6%) individuals with low HDL-C. There were some individuals that had low genotyping call rates so it is worth mentioning that they may be underestimated in the analysis of rare variants.

Among 23 identified exonic variants, 10 were relatively uncommon or rare variants (MAF <0.05). Of these 10 relatively uncommon or rare exonic variants, two were found only in the high HDL-C group (24143C>G, Ser447X) and (27783A>T, P=0.142) which are located in exon 9 and untranslated region of the exon 10, respectively. On the other hand, two were found only in low HDL-C group (27688C>T, P=0.326 and 28524C>T, P=0.321), however, they are located in the

untranslated region of exon 10. There was a one base insertion in the 3' flanking region which was found only in the low HDL-C group (29557_29558InsA, P=0.321) but the remaining 19 exonic variants were found in both the high and low HDL-groups. Table 9 summarizes the distribution of common *LPL* variants in high and low HDL-C groups. The relatively uncommon or rare *LPL* variants in high and low HDL-C groups are listed in Table 10 and the variants which were found only in high or low HDL-C groups are specifically highlighted; the 25 variants that were found in only high HDL-C group are highlighted in yellow and the 21 variants that were identified in only low HDL-C group are highlighted in blue.

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

| ***LPL* variant (Nucleotide #)** | **Alleles** | **Location** | **refSNP ID** | **All MAF (n=95)** | **Assoc. Allele** | **%MAF of High HDL-C** | **%MAF of Low HDL-C** | **P value** |
|---|---|---|---|---|---|---|---|---|
| 2335 | G>T | Intron 1 | rs3779787 | 0.122 | T | 0.163 | 0.083 | 0.096 |
| 3558 | G>A | Intron 1 | rs34309063 | 0.247 | A | 0.160 | 0.333 | **0.006** |
| 3964 | G>C | Intron 1 | rs17410577 | 0.239 | C | 0.160 | 0.320 | **0.010** |
| 4060 | G>T | Intron 1 | rs1534649 | 0.430 | T | 0.383 | 0.478 | 0.189 |
| 4424 | A>G | Intron 1 | rs13266204 | 0.263 | G | 0.200 | 0.330 | **0.038** |
| 5949 | T>G | Intron 1 | | 0.137 | G | 0.160 | 0.115 | 0.367 |
| 6821 | C>T | Intron 1 | rs10104051 | 0.426 | T | 0.404 | 0.447 | 0.555 |
| 7512 | C>T | Intron 1 | rs3779788 | 0.121 | T | 0.160 | 0.083 | 0.107 |
| 8415 | T>A | Intron 1 | rs56321069 | 0.183 | A | 0.228 | 0.138 | 0.113 |
| 10987 | C>A | Intron 2 | | 0.122 | A | 0.140 | 0.106 | 0.498 |
| 11090 | C>G | Intron 2 | rs8176337 | 0.190 | G | 0.239 | 0.146 | 0.109 |
| 12449 | G>A | Intron 2 | | 0.117 | A | 0.152 | 0.083 | 0.142 |
| 12550 | G>A | Intron 2 | | 0.118 | A | 0.156 | 0.083 | 0.128 |
| 12810_12829 | dup20 | Intron 2 | | 0.126 | G | 0.163 | 0.089 | 0.132 |
| 12853_12854 | Ins16 | Intron 2 | | 0.055 | C | 0.054 | 0.056 | 0.972 |
| 12861_12864 | del4 | Intron 2 | | 0.060 | C | 0.054 | 0.933 | 0.727 |
| 13003 | G>T | Intron 2 | | 0.053 | T | 0.053 | 0.052 | 0.973 |
| 14114 | T>C | Intron 3 | rs73667472 | 0.080 | C | 0.065 | 0.094 | 0.471 |
| 15425 | T>C | Intron 4 | rs249 | 0.095 | C | 0.085 | 0.104 | 0.654 |
| 15449_15450 | Ins2 | Intron 4 | | 0.063 | C | 0.096 | 0.031 | 0.068 |
| 15653 | delA | Intron 4 | | 0.410 | C | 0.394 | 0.426 | 0.656 |
| 15836 | C>T | Intron 4 | rs253 | 0.410 | C | 0.394 | 0.426 | 0.656 |
| 16316 | C>G | Intron 5 | rs254 | 0.074 | G | 0.060 | 0.913 | 0.487 |
| 16320 | T>C | Intron 5 | rs255 | 0.105 | C | 0.106 | 0.104 | 0.960 |
| 16386 | C>T | Intron 5 | rs256 | 0.118 | T | 0.133 | 0.104 | 0.538 |
| 16671 | G>C | Intron 5 | rs258 | 0.409 | C | 0.380 | 0.436 | 0.440 |
| 17231 | C>T | Intron 5 | rs263 | 0.135 | T | 0.148 | 0.122 | 0.618 |
| 17599 | G>A | Intron 5 | rs264 | 0.116 | A | 0.128 | 0.104 | 0.613 |
| 18086 | T>G | Intron 6 | rs269 | 0.119 | G | 0.144 | 0.093 | 0.293 |
| 18095 | C>A | Intron 6 | rs270 | 0.201 | A | 0.140 | 0.261 | **0.045** |
| 18121 | G>A | Intron 6 | rs271 | 0.109 | A | 0.130 | 0.087 | 0.343 |
| 18395_18396 | insT | Intron 6 | | 0.121 | G | 0.138 | 0.104 | 0.471 |
| 18822 | T>C | Intron 6 | rs277 | 0.232 | C | 0.181 | 0.281 | 0.101 |
| 18942 | G>A | Intron 6 | rs278 | 0.211 | A | 0.170 | 0.250 | 0.177 |
| 19442 | A>T | Intron 6 | rs281 | 0.242 | T | 0.207 | 0.278 | 0.262 |
| 19445 | C>G | Intron 6 | rs282 | 0.129 | G | 0.133 | 0.125 | 0.866 |
| 19517 | C>T | Intron 6 | rs283 | 0.161 | T | 0.167 | 0.156 | 0.847 |
| 19608 | C>T | Intron 6 | **rs285 (PvuII)** | 0.407 | T | 0.444 | 0.370 | 0.304 |
| 19675 | A>T | Intron 6 | rs286 | 0.060 | T | 0.111 | 0.011 | **0.004** |
| 19975 | A>G | Intron 6 | rs287 | 0.172 | G | 0.244 | 0.104 | **0.011** |
| 20038 | T>C | Intron 6 | rs289 | 0.159 | C | 0.221 | 0.104 | **0.032** |
| 20271 | T>C | Intron 6 | rs291 | 0.163 | C | 0.234 | 0.094 | **0.009** |
| 20363 | A>T | Intron 6 | | 0.053 | T | 0.043 | 0.062 | 0.538 |

**Table 10 (Continued)**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 20505_20506 | insA | Intron 6 | | 0.163 | G | 0.234 | 0.094 | **0.009** |
| 20544 | T>C | Intron 6 | rs294 | 0.080 | C | 0.087 | 0.073 | 0.723 |
| 20657 | A>C | Intron 6 | rs295 | 0.163 | C | 0.234 | 0.094 | **0.009** |
| 20790 | T>C | Intron 6 | rs297 | 0.163 | C | 0.234 | 0.094 | **0.009** |
| 21353 | T>C | Intron 7 | rs301 | 0.172 | C | 0.223 | 0.120 | 0.061 |
| 21780 | T>G | Intron 7 | rs304 | 0.103 | G | 0.174 | 0.034 | **0.002** |
| 21820 | A>G | Intron 7 | rs305 | 0.108 | G | 0.179 | 0.037 | **0.003** |
| 22044_22047 | del4 | Intron 7 | | 0.075 | A | 0.087 | 0.064 | 0.550 |
| 22416 | G>C | Intron 7 | rs312 | 0.072 | C | 0.080 | 0.065 | 0.711 |
| 22461 | G>A | Intron 7 | rs314 | 0.180 | A | 0.244 | 0.120 | **0.031** |
| 22855 | C>A | **Exon 8** | **rs316(T361T)** | 0.079 | A | 0.074 | 0.083 | 0.821 |
| 23190_23191 | del2 | Intron 8 | | 0.500 | A | 0.456 | 0.453 | 0.228 |
| 23192 | G>T | Intron 8 | | 0.068 | T | 0.122 | 0.012 | **0.004** |
| 23395 | A>C | Intron 8 | rs319 | 0.293 | C | 0.282 | 0.302 | 0.776 |
| 23636 | A>C | Intron 8 | rs322 | 0.170 | C | 0.227 | 0.117 | **0.048** |
| 23747 | T>C | Intron 8 | rs325 | 0.067 | C | 0.102 | 0.033 | 0.067 |
| 23858 | A>G | Intron 8 | rs326 | 0.206 | G | 0.278 | 0.133 | **0.017** |
| 23955 | T>G | Intron 8 | rs327 | 0.192 | G | 0.261 | 0.122 | **0.018** |
| 23496 | T>G | **Intron 8** | **rs320 (*Hind*III)** | 0.185 | G | 0.244 | 0.128 | 0.053 |
| 24815 | G>A | Intron 9 | rs330 | 0.111 | A | 0.128 | 0.094 | 0.456 |
| 24824 | G>A | Intron 9 | rs331 | 0.122 | A | 0.154 | 0.096 | 0.247 |
| 24852 | T>C | Intron 9 | rs12679834 | 0.080 | C | 0.138 | 0.021 | **0.003** |
| 25049 | G>A | Intron 9 | | 0.063 | A | 0.043 | 0.083 | 0.248 |
| 25335 | C>T | Intron 9 | | 0.079 | T | 0.138 | 0.021 | **0.003** |
| 25352 | A>C | Intron 9 | | 0.079 | C | 0.138 | 0.021 | **0.003** |
| 25844 | T>G | Intron 9 | | 0.074 | G | 0.130 | 0.021 | **0.004** |
| 26201 | T>G | Intron 9 | | 0.077 | G | 0.136 | 0.021 | **0.004** |
| 26234 | T>G | Intron 9 | rs10099160 | 0.282 | G | 0.250 | 0.312 | 0.341 |
| 27229 | C>T | Intron 9 | rs11570891 | 0.082 | T | 0.144 | 0.021 | **0.002** |
| 27249 | G>A | **Exon 10-3' UTR** | **rs4922115** | 0.085 | A | 0.087 | 0.083 | 0.929 |
| 28036 | A>G | **Exon 10-3' UTR** | **rs11570892** | 0.101 | G | 0.117 | 0.085 | 0.468 |
| 28067 | A>T | **Exon 10-3' UTR** | **rs3208305** | 0.203 | T | 0.278 | 0.130 | **0.014** |
| 28093 | C>T | **Exon 10-3' UTR** | **rs1803924** | 0.082 | T | 0.141 | 0.022 | **0.004** |
| 28382 | C>T | **Exon 10-3' UTR** | **rs1059507** | 0.110 | T | 0.120 | 0.100 | 0.673 |
| 28464 | C>A | **Exon 10-3' UTR** | **rs3735964** | 0.079 | A | 0.138 | 0.021 | **0.003** |
| 28490 | A>G | **Exon 10-3' UTR** | **rs3200218** | 0.279 | G | 0.255 | 0.302 | 0.472 |
| 28911 | T>C | **Exon 10-3' UTR** | **rs13702** | 0.200 | C | 0.277 | 0.125 | **0.009** |
| 28982 | T>C | **Exon 10-3' UTR** | **rs1059611** | 0.079 | C | 0.138 | 0.021 | **0.003** |
| 29046_29047 | Ins2 | **Exon 10-3' UTR** | | 0.079 | G | 0.138 | 0.021 | **0.003** |
| 29086 | C>T | **Exon 10-3' UTR** | **rs15285** | 0.200 | T | 0.277 | 0.125 | **0.009** |

**Table 10 (Continued)**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 29088 | C>A | **Exon 10-3' UTR** | **rs3866471** | 0.105 | A | 0.117 | 0.094 | 0.601 |
| 29287 | G>A | 3' flanking | rs3916027 | 0.184 | A | 0.255 | 0.115 | **0.012** |
| 29315 | T>G | 3' flanking | rs9644636 | 0.339 | G | 0.348 | 0.330 | 0.795 |
| 29487 | T>A | 3' flanking | rs4921683 | 0.105 | A | 0.117 | 0.094 | 0.601 |
| 29547 | C>T | 3' flanking | rs4921684 | 0.105 | T | 0.117 | 0.094 | 0.601 |

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

| ***LPL* variant (Nucleotide #)** | **Alleles** | **Location** | **refSNP ID** | **All MAF (n=95)** | **MAF(%) assoc. allele High HDL-C** | **MAF(%) assoc. allele Low HDL-C** |
|---|---|---|---|---|---|---|
| 208 | T>C | 5' flanking | rs1470186 | 0.016 | 0.021 | 0.011 |
| 351 | C>A | 5' flanking | | 0.005 | 0.011 | 0.000 |
| 428 | G>A | 5' flanking | rs73667465 | 0.016 | 0.021 | 0.011 |
| 549 | C>T | 5' flanking | rs17091742 | 0.016 | 0.021 | 0.011 |
| 958 | G>A | 5' flanking | | 0.005 | 0.000 | 0.011 |
| 1088 | G>T | 5' flanking | | 0.005 | 0.011 | 0.000 |
| 1090 | T>G | **5' flanking** | **rs1800590 (-T93G)** | 0.011 | 0.011 | 0.010 |
| 1130 | G>C | 5' flanking | | 0.005 | 0.000 | 0.010 |
| 2913 | T>C | Intron 1 | | 0.005 | 0.000 | 0.011 |
| 4621 | C>G | Intron 1 | | 0.021 | 0.021 | 0.021 |
| 4948 | C>G | Intron 1 | rs6997330 | 0.022 | 0.033 | 0.010 |
| 5094 | C>G | Intron 1 | | 0.005 | 0.000 | 0.011 |
| 5107 | C>T | Intron 1 | | 0.005 | 0.011 | 0.000 |
| 5118 | A>T | Intron 1 | | 0.022 | 0.043 | 0.000 |
| 5200 | C>T | Intron 1 | | 0.006 | 0.000 | 0.011 |
| 5531 | G>A | Intron 1 | rs1031045 | 0.016 | 0.022 | 0.011 |
| 5772 | A>G | Intron 1 | rs60633545 | 0.016 | 0.022 | 0.010 |
| 6383 | G>T | Intron 1 | | 0.005 | 0.011 | 0.000 |
| 6435 | G>C | Intron 1 | | 0.005 | 0.011 | 0.000 |
| 6477 | T>C | Intron 1 | | 0.005 | 0.011 | 0.000 |
| 6553 | C>T | Intron 1 | rs59254395 | 0.016 | 0.021 | 0.010 |
| 6554 | A>G | Intron 1 | rs56043715 | 0.016 | 0.021 | 0.010 |
| 7130 | T>C | Intron 1 | rs28615996 | 0.021 | 0.029 | 0.014 |
| 7131 | T>G | Intron 1 | | 0.007 | 0.014 | 0.000 |
| 7313 | G>A | Intron 1 | rs28645722 | 0.016 | 0.022 | 0.010 |
| 7388 | C>G | Intron 1 | rs28575919 | 0.016 | 0.021 | 0.010 |
| 7503 | T>C | Intron 1 | rs6999612 | 0.016 | 0.021 | 0.010 |
| 7556 | T>C | Intron 1 | rs59811201 | 0.016 | 0.021 | 0.010 |
| 8221 | A>C | Intron 1 | rs7000460 | 0.018 | 0.026 | 0.011 |
| 8250 | G>A | Intron 1 | rs59630933 | 0.019 | 0.026 | 0.012 |
| 8467 | C>T | Intron 1 | | 0.005 | 0.000 | 0.010 |
| 8516 | delG | Intron 1 | | 0.011 | 0.011 | 0.010 |
| 9015 | A>G | Intron 1 | rs28445964 | 0.021 | 0.022 | 0.020 |
| 9024 | T>C | Intron 1 | | 0.008 | 0.000 | 0.014 |
| 9130 | T>A | Intron 1 | rs13252357 | 0.005 | 0.011 | 0.000 |
| 9411 | A>C | Intron 1 | rs28689946 | 0.016 | 0.021 | 0.011 |

**Table 11 (Continued)**

| | | | | | | |
|---|---|---|---|---|---|---|
| 9418 | G>A | Intron 1 | rs28582042 | 0.016 | 0.021 | 0.011 |
| 9589 | C>T | Intron 1 | | 0.016 | 0.021 | 0.011 |
| 9696 | G>T | Intron 1 | rs73667468 | 0.016 | 0.022 | 0.010 |
| 9914 | T>G | Intron 1 | rs73667469 | 0.016 | 0.022 | 0.011 |
| 10127 | G>A | **Exon 2** | **rs1801177 (D9N)** | 0.016 | 0.021 | 0.010 |
| 10632 | C>T | Intron 2 | | 0.005 | 0.000 | 0.001 |
| 10912 | A>G | Intron 2 | | 0.022 | 0.000 | 0.043 |
| 11050 | T>C | Intron 2 | rs7016529 | 0.016 | 0.023 | 0.010 |
| 11228 | T>C | Intron 2 | | 0.005 | 0.011 | 0.000 |
| 11574 | G>A | Intron 2 | rs34123038 | 0.049 | 0.056 | 0.043 |
| 11600 | G>C | Intron 2 | | 0.011 | 0.000 | 0.021 |
| 11760 | A>C | Intron 2 | rs73667470 | 0.016 | 0.021 | 0.010 |
| 11888_11889 | insA | Intron 2 | | 0.011 | 0.021 | 0.000 |
| 12224_12920 | del697 | Intron 2 | | 0.005 | 0.011 | 0.000 |
| 12484 | C>A | Intron 2 | | 0.048 | 0.044 | 0.052 |
| 12878_12889 | del12 | Intron 2 | | 0.005 | 0.011 | 0.000 |
| 12884_12887 | del4 | Intron 2 | | 0.016 | 0.033 | 0.000 |
| 13639 | G>A | Intron 2 | | 0.011 | 0.021 | 0.000 |
| 13854 | G>A | **Exon 3** | **rs1121923 (V108V)** | 0.016 | 0.021 | 0.010 |
| 15206 | C>A | Intron 3 | rs343 | 0.042 | 0.064 | 0.021 |
| 15245 | G>A | **Exon 4** | **rs248 (E118E)** | 0.047 | 0.053 | 0.042 |
| 16442 | G>C | Intron 5 | | 0.005 | 0.000 | 0.010 |
| 16563 | T>A | Intron 5 | | 0.016 | 0.011 | 0.021 |
| 17476 | A>C | Intron 5 | | 0.005 | 0.000 | 0.010 |
| 17948 | A>G | **Exon 6** | **rs268 (N291S)** | 0.033 | 0.022 | 0.043 |
| 18065 | T>G | Intron 6 | | 0.009 | 0.000 | 0.017 |
| 18297 | A>C | Intron 6 | | 0.005 | 0.011 | 0.000 |
| 18462 | T>G | Intron 6 | | 0.005 | 0.000 | 0.010 |
| 18621 | C>T | Intron 6 | | 0.005 | 0.011 | 0.000 |
| 18708 | T>C | Intron 6 | rs276 | 0.021 | 0.021 | 0.021 |
| 19815 | G>A | Intron 6 | | 0.005 | 0.011 | 0.000 |
| 20080 | C>T | Intron 6 | | 0.005 | 0.011 | 0.000 |
| 20663 | G>A | Intron 6 | rs296 | 0.005 | 0.000 | 0.010 |
| 21125_21128 | del4 | Intron 6 | | 0.005 | 0.000 | 0.011 |
| 21895 | T>G | Intron 7 | rs308 | 0.012 | 0.013 | 0.011 |
| 21965 | C>T | Intron 7 | rs310 | 0.006 | 0.013 | 0.000 |
| 22514 | T>C | Intron 7 | rs315 | 0.011 | 0.000 | 0.022 |

**Table 11 (Continued)**

| | | | | | | |
|---|---|---|---|---|---|---|
| 23388 | C>G | Intron 8 | rs318 | 0.018 | 0.024 | 0.012 |
| 23573 | T>C | Intron 8 | | 0.017 | 0.024 | 0.011 |
| 24143 | C>G | **Exon 9** | **rs328 (Ser447X)** | 0.005 | 0.011 | 0.000 |
| 24505 | A>G | Intron 9 | **rs329** | 0.016 | 0.021 | 0.011 |
| 24573 | T>C | Intron 9 | | 0.005 | 0.000 | 0.011 |
| 24899 | C>T | Intron 9 | | 0.016 | 0.021 | 0.010 |
| 25005 | A>G | Intron 9 | | 0.005 | 0.011 | 0.000 |
| 25320 | C>T | Intron 9 | | 0.005 | 0.011 | 0.000 |
| 27000 | C>T | Intron 9 | | 0.029 | 0.044 | 0.012 |
| 27160 | T>A | Intron 9 | | 0.027 | 0.058 | 0.000 |
| 27611 | T>C | **Exon 10-3' UTR** | rs3289 | 0.016 | 0.022 | 0.010 |
| 27688 | C>T | **Exon 10-3' UTR** | | 0.005 | 0.000 | 0.010 |
| 27783 | A>T | **Exon 10-3' UTR** | | 0.011 | 0.022 | 0.000 |
| 28407 | C>A | **Exon 10-3' UTR** | | 0.026 | 0.021 | 0.031 |
| 28524 | C>T | **Exon 10-3' UTR** | | 0.005 | 0.000 | 0.010 |
| 29474 | C>T | 3' flanking | | 0.026 | 0.043 | 0.010 |
| 29557_29558 | InsA | 3' flanking | | 0.005 | 0.000 | 0.010 |
| 29716 | T>C | 3' flanking | | 0.026 | 0.043 | 0.010 |

## 3.3 LINKAGE DISEQUILIBRIUM (LD) AND TAGGER ANALYSES OF *LPL* VARIANTS

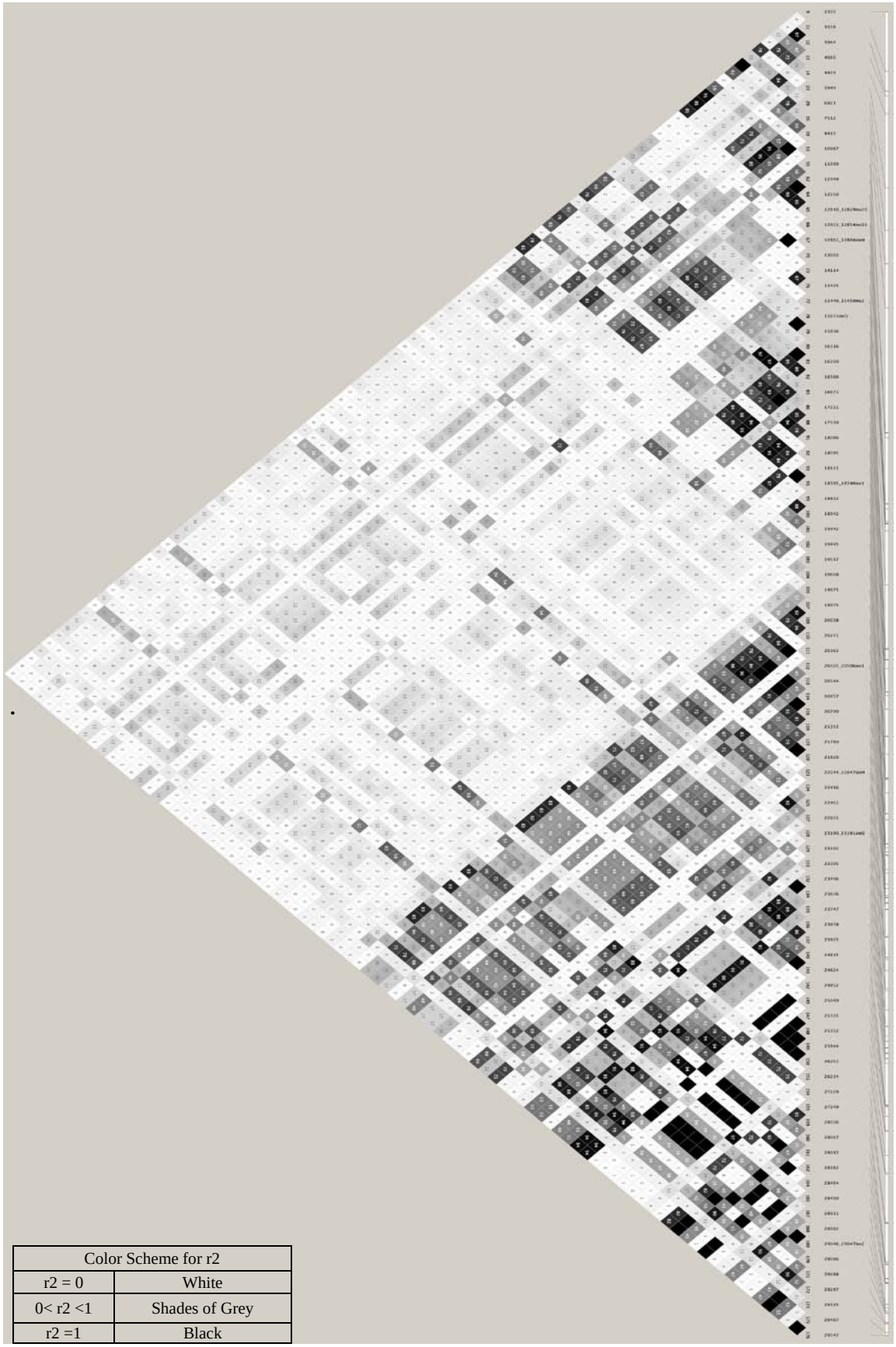
We used LD and Tagger analysis to identify tagSNPs for the common variants (MAF≥0.05) that we identified by resequencing. By using tagger analysis with an $r^2$ cutoff of 0.9, 43 tagSNP bins were identified. Table 12 shows the bins identified by Tagger analysis and Figure 5 demonstrates the LD plot. Twelve variants that were genotyped in our entire NHW sample are underlined in Table 12. We also used LD analysis in Haploview to see the LD pattern of the 32 common variants that had p-value≤0.05 when comparing the allele frequencies between the high HDL-C and low HDL-C groups (Figure 6).

**Table 12.** Tagger results using Haploview of *LPL* common variants in NHWs

| Bin | Variants captured |
|---|---|
| 1 | 28093, 23747, 26201, 25352, 25844, 25335, 24852, 29046_29047ins2, 28464, 28982 |
| 2 | 29547, 29487, 24824, 24815, 28036, 28382, 29088 |
| 3 | 28911, 29086, 23858, 29287, 28067 |
| 4 | 7512, 12810_12829ins20, 12550, 12449, 2335 |
| 5 | 16386, 18121, 17599, 18395_18396ins1 |
| 6 | 20790, 20505_20506ins1, 20657, 20271 |
| 7 | 16671, 15653del1, 15836 |
| 8 | 23955, 23636, 23496(*Hind*III) |
| 9 | 26234, 23395, 28490 |
| 10 | 22855, 22416, 27249 |
| 11 | 20038, 19975, 21353 |
| 12 | 3558, 3964 |
| 13 | 16320, 16316 |
| 14 | 21780, 21820 |
| 15 | 8415, 11090 |
| 16 | 13003, 12853_12854ins16 |
| 17 | 18086, 1723 |
| 18 | 6821 |

**Table 12 (Continued)**

| | |
|---|---|
| 19 | 19608 |
| 20 | 15425 |
| 21 | 15449_15450ins2 |
| 22 | 25049 |
| 23 | 12861_12864del4 |
| 24 | 18095 |
| 25 | 14114 |
| 26 | 19442 |
| 27 | 19445 |
| 28 | 4424 |
| 29 | 23192 |
| 30 | 20544 |
| 31 | 19675 |
| 32 | 27229 |
| 33 | 20363 |
| 34 | 4060 |
| 35 | 18822 |
| 36 | 19517 |
| 37 | 10987 |
| 38 | 29315 |
| 39 | 22461 |
| 40 | 18942 |
| 41 | 22044_22047del2 |
| 42 | 23190_23191del2 |
| 43 | 5949 |



| Color Scheme for r2 | |
|---|---|
| r2 = 0 | White |
| 0< r2 <1 | Shades of Grey |
| r2 =1 | Black |

**Figure 5.** LD analysis of identified common variants in NHWs

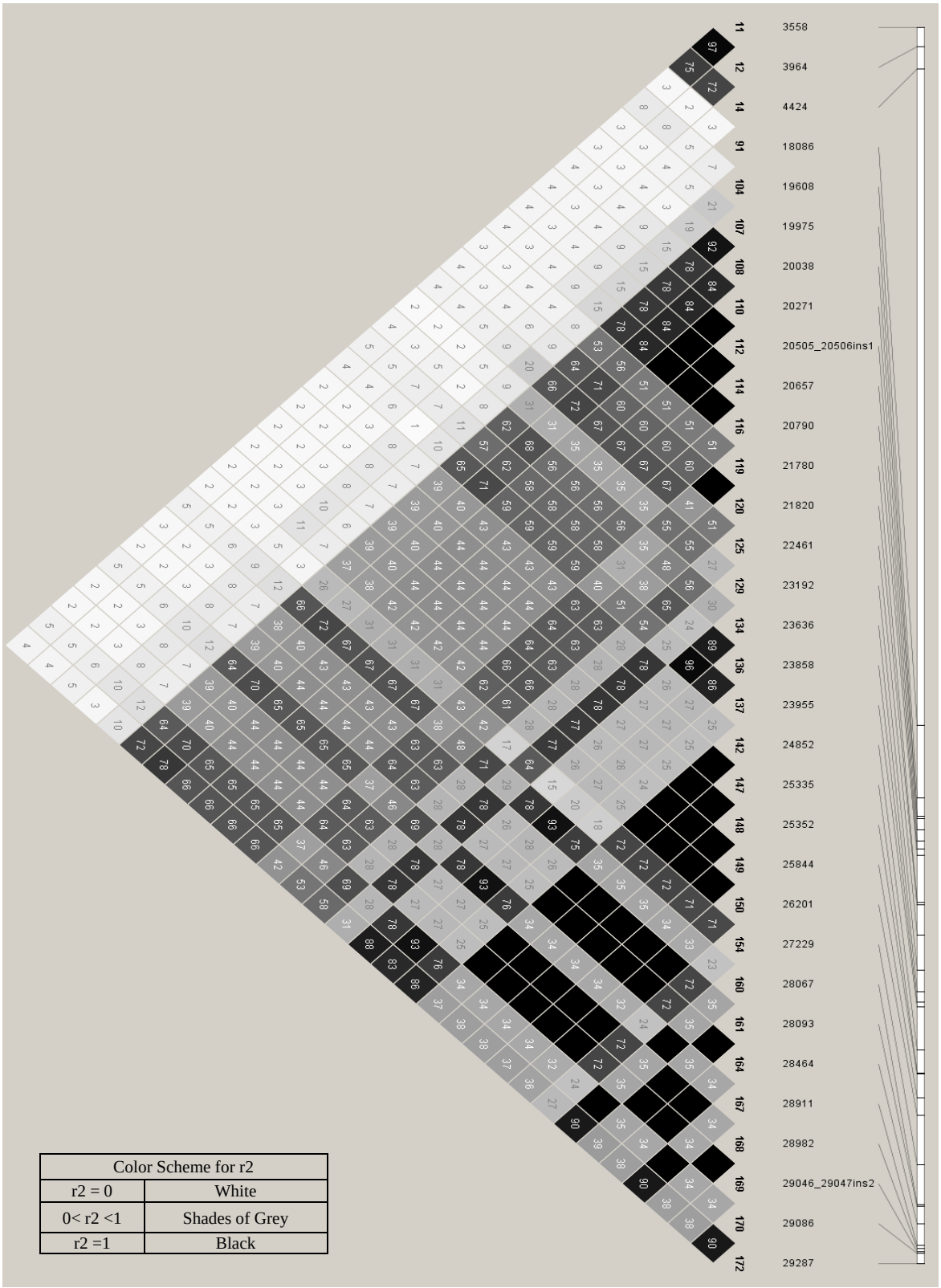



**Figure 6.** LD analysis of the 32 identified variants which had a statistically significant p-value <0.05 when comparing the allele frequencies between the high HDL-C and low HDL-C groups

## 3.4 GENOTYPING OF IDENTIFIED VARIANTS IN THE ENTIRE NHW SAMPLE AND THEIR ASSOCIATION ANALYSIS WITH PLASMA LIPOPROTEIN LEVELS

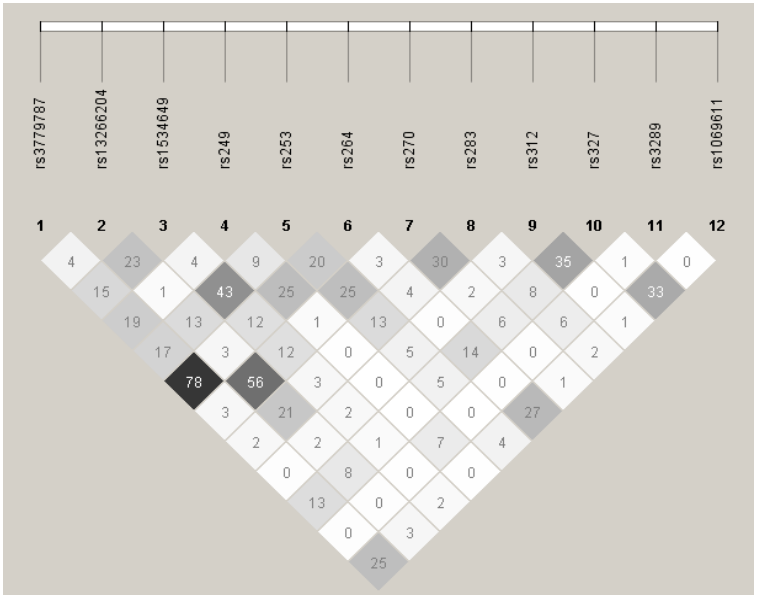
To date, we have genotyped 12 variants in our NHW sample (n=623) using pre-made TaqMan SNP Genoytping Assays. Genotyping results and features of these tagSNPs are shown in the Table 13.

**Table 13.** Genotyping results of 12 tagSNPs screened in the entire NHW sample

| ***LPL*** **refSNP ID** | **Location** | **Position (Nucleotide #)** | **Number of individuals genotyped in NHWs** | **MAF % in *CEU (HAPMAP)** | **MAF % in NHWs** | **Call rates (%)** | **P-value** |
|---|---|---|---|---|---|---|---|
| rs3779787 | Intron 1 | 2335 | 619 | 0.175 | 0.153 | 99.40 | 0.376 |
| rs13266204 | Intron 1 | 4424 | 620 | 0.259 | 0.215 | 99.50 | 0.616 |
| rs1534649 | Intron 1 | 4060 | 618 | 0.472 | 0.453 | 99.20 | 0.599 |
| rs249 | Intron 4 | 15425 | 614 | 0.075 | 0.074 | 98.60 | 0.211 |
| rs253 | Intron 4 | 15836 | 623 | 0.450 | 0.451 | 100.00 | 0.505 |
| rs264 | Intron 5 | 17599 | 605 | 0.142 | 0.140 | 97.10 | 0.428 |
| rs270 | Intron 6 | 18095 | 618 | 0.208 | 0.172 | 99.20 | 0.896 |
| rs283 | Intron 6 | 19517 | 618 | 0.188 | 0.201 | 99.20 | 0.352 |
| rs312 | Intron 7 | 22416 | 617 | 0.092 | 0.112 | 99.00 | 0.976 |
| rs327 | Intron 8 | 23955 | 619 | 0.250 | 0.264 | 99.40 | 0.478 |
| rs3289 | Exon 10 3'UTR | 27611 | 622 | 0.042 | 0.026 | 99.80 | 1.000 |
| rs1059611 | Exon 10 3'UTR | 28982 | 618 | 0.133 | 0.106 | 99.20 | 1.000 |

* CEU: U.S. residents with northern and western European

None of the variants showed deviation from HWE (P>0.05) in the total sample. Figure 7 shows the LD analysis that was repeated for the variants screened using TaqMan assays. The LD patterns between the 12 variants were similar to those observed in the subset (high and low HDL groups) of our sample that we used for sequencing.




| Color Scheme for r2 | |
|---|---|
| r2 = 0 | White |
| 0< r2 <1 | Shades of Grey |
| r2 =1 | Black |

**Figure 7.** LD analysis of the variants screened in the entire NHW sample

These 12 variants screened in total NWH sample were analyzed by using both additive and dominant models for their relation to plasma total cholesterol, HDL-C, LDL-C and TG levels. Table 14 shows the genotype counts, covariate adjusted mean for each lipid level and adjusted p-values for additive and dominant models, respectively. The remaining common tagSNPs that are not included in these analyzes are currently being genotyped by Sequenome IPLEX genotyping platform so they are not included in this study.

Of the 12 variants, 4 (rs1059611, rs270, rs3779787, rs327) had statistically significant p-values associated with HDL-C level in both additive and dominant models, but rs13266204 had a significant P-value (0.013) only in additive model and had a marginally significant p-value (0.080) in the dominant model.

Two of the 12 variants (rs327, rs270) showed statistically significant association with triglycerides levels in both additive and dominant models. rs312 showed statistically significant

association with triglyceride levels in additive model (P=0.0091) and it showed marginally significant association in the dominant (P=0.050) model. On the other hand, rs283 had a statistically significant P-value (0.024) associated with TG levels only in dominant model but it had a marginally significant P-value (0.068) in additive model.

There were two variants (rs270 and rs312) that had marginally statistically significant P-values associated with LDL-C levels; rs270 showed marginally significant association in only additive model (P=0.056), however, rs312 showed marginally significant association in only dominant model (P=0.087).

**Table 14.**Genotype distributions and adjusted p-values for 12 *LPL* variants in NHWs

| | **rs13266204_Add** | | | **rs13266204_Dom** | |
|---|---|---|---|---|---|
| | AA[380] | GA[203] | GG[31] | AA[380] | GA/GG[234] |
| **Total Cholesterol Adjusted mean± SD** | 216.03±2.13 | 219.34±2.90 | 218.32±7.21 | 216.02±2.13 | 219.2±2.69 |
| **P-value** | P= 0.400 | | | P=0.342 | |
| | AA[383] | GA[203] | GG[31] | AA[383] | GA/GG[234] |
| **HDL-C Adjusted mean± SD** | 50.95±0.64 | 49.95±0.88 | 45.26±2.19 | 50.93±0.65 | 49.3±0.82 |
| **P-value** | **P=0.013** | | | **P=0.080** | |
| | AA[378] | GA[203] | GG[29] | AA[378] | GA/GG[232] |
| **LDL-C Adjusted mean± SD** | 137.40±1.97 | 141.56±2.68 | 143.13±6.89 | 137.41±1.97 | 141.76±2.50 |
| **P-value** | P=0.164 | | | P=0.160 | |
| | AA[380] | GA[203] | GG[31] | AA[380] | GA/GG[234] |
| **Triglycerides Adjusted mean± SD** | 140.70±3.52 | 139.13±4.79 | 180.90±11.92 | 140.9±3.54 | 144.89±4.49 |
| **P-value** | P=0.082 | | | P=0.329 | |
| | **rs312_Add** | | | **rs312_Dom** | |
| | GG[482 | GC[122] | CC[7] | GG[482] | GC/CC[129] |
| **Total Cholesterol Adjusted mean± SD** | 215.70±1.8 | 222.44±3.67 | 211.87±1.15 | 215.71±1.89 | 221.89±3.57 |
| **P-value** | P=0.174 | | | P=0.118 | |
| | GG[484] | GC[123] | CC[7] | GG[482] | GC/CC[129] |
| **HDL-C Adjusted mean± SD** | 50.21±0.58 | 51.06±1.12 | 57.26±4.66 | 50.21±0.58 | 51.39±1.1 |
| **P-value** | P=0.118 | | | P=0.201 | |
| | GG[480] | GC[120] | CC[7] | GG[480] | GC/CC[127] |
| **LDL-C Adjusted mean± SD** | 137.47±1.75 | 144.04±3.42 | 138.97±13.99 | 137.47±1.75 | 143.77±3.33 |
| **P-value** | P=0.114 | | | **P=0.087** | |
| | GG[482] | GC[122] | CC[7] | GG[482] | GC/CC[129] |
| **Triglycerides Adjusted mean± SD** | 144.13±3.1 | 136.49±6.14 82 | 85.52±25.36 | 144.19±3.16 | 133.65±6.00 |
| **P-value** | **P=0.009** | | | **P=0.050** | |
| | **rs1059611_Add** | | | **rs1059611_Dom** | |
| | TT[490] | TC[116] | CC[7] | TT[490] | TC/CC[123] |
| **Total Cholesterol Adjusted mean± SD** | 216.74±1.8 | 219.38±3.73 | 210.19±15.10 | 216.76±1.89 | 218.88±3.63 |
| **P-value** | P=0.713 | | | P=0.599 | |
| | TT[491] | TC[117] | CC[7] | TT[491] | TC/CC[124] |
| **HDL-C Adjusted mean± SD** | 49.71±0.58 | 52.78±1.13 | 49.36±4.61 | 49.72±0.58 | 52.6±1.1 |
| **P-value** | **P=0.036** | | | **P=0.018** | |
| | TT[486] | TC[117] | CC[7] | TT[486] | TC/CC[123] |
| **LDL-C Adjusted mean± SD** | 139.10±1.75 | 138.78±3.45 | 138.41±13.96 | 139.10±1.75 | 138.76±3.36 |
| **P-value** | P=0.925 | | | P=0.928 | |
| | TT[490] | TC[115] | CC[7] | TT[490] | TC/CC[122] |
| **Triglycerides Adjusted mean± SD** | 143.66±3.117 | 139.91±6.29 | 113.17±25.35 | 143.72±3.17 | 138.44±6.13 |
| **P-value** | P=0.346 | | | P=0.481 | |

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

Natural log transformed data were used for association analysis of triglycerides and HDL-C levels

**Table 14 (Continued)**

| | rs253_Add | | | rs253_Dom | |
|---|---|---|---|---|---|
| | CC[181] | TC[315] | TT[121] | CC[181] | TC/TT[436] |
| **Total Cholesterol Adjusted mean± SD** | 213.91±3.01 | 218.61±2.32 | 218.93±3.71 | 213.82±3 | 218.69±2 |
| **P-value** | P=0.231 | | | P=0.168 | |
| | CC[182] | TC[316] | TT[122] | CC[182] | TC/TT[438] |
| **HDL-C Adjusted mean± SD** | 51.19±0.92 | 49.94±0.71 | 50.38±1.13 | 51.19±0.92 | 50.06±0.61 |
| **P-value** | P=0.595 | | | P=-0.290 | |
| | CC[181] | TC[312] | TT[120] | CC[181] | TC/TT[432] |
| **LDL-C Adjusted mean± SD** | 136.34±2.78 | 140.15±2.16 | 140.34±3.45 | 136.34±2.78 | 140.20±1.86 |
| **P-value** | P=0.309 | | | P=0.239 | |
| | CC[182] | TC[314] | TT[121] | CC[182] | TC/TT[435] |
| **Triglycerides Adjusted mean± SD** | 139.61±5.01 | 143.13±3.87 | 143.24±6.20 | 139.61±5.00 | 143.16±3.35 |
| **P-value** | P=0.686 | | | P=0.683 | |
| | **rs1534649_Add** | | | **rs1534649_Dom** | |
| | GG[177] | GT[314] | TT [122] | GG[177] | GT/TT [436] |
| **Total Cholesterol Adjusted mean± SD** | 216.50±3.022 | 217.63±2.32 | 215.99±3.65 | 216.51±3.02 | 217.17±2 |
| **P-value** | P=0.956 | | | P=0.850 | |
| | GG[179] | GT[314] | TT [123] | GG[179] | GT/TT [437] |
| **HDL-C Adjusted mean± SD** | 50.47±0/92 | 50.52±0.71 | 49.82±1.12 | 50.48±0.92 | 50.32±0.61 |
| **P-value** | P=0.730 | | | P=0.815 | |
| | GG[177] | GT[311] | TT [121] | GG[177] | GT/TT [432] |
| **LDL-C Adjusted mean± SD** | 138.43 | 138.81±2.16 | 139.61±3.39 | 138.43±2.78 | 140.20±1.86 |
| **P-value** | P=0.792 | | | P=0.855 | |
| | GG[179] | GT[312] | TT [122] | GG[179] | GT/TT [434] |
| **Triglycerides Adjusted mean± SD** | 145.73±5.04 | 140.98±3.91 | 138.33±6.13 | 145.74±5.04 | 140.24±3.36 |
| **P-value** | P=0.175 | | | P=0.219 | |
| | **rs327_Add** | | | **rs327_Dom** | |
| | TT[337] | TG[229] | GG[47] | TT[337] | TG/GG[276] |
| **Total Cholesterol Adjusted mean± SD** | 217.07±2.23 | 216.02±2.68 | 222.89±5.84 | 217.07±2.23 | 217.18±2.47 |
| **P-value** | P=0.645 | | | P=0.971 | |
| | TT[338] | TG[231] | GG[47] | TT[338] | TG/GG[278] |
| **HDL-C Adjusted mean± SD** | 49.64±0.68 | 50.97±0.82 | 52.52±1.80 | 49.64±0.68 | 51.23±0.76 |
| **P-value** | **P=0.030** | | | **P=0.038** | |
| **LDL-C Adjusted mean± SD** | TT[335] | TG[227] | GG[47] | TT[335] | TG/GG[274] |
| **P-value** | 139.14±2.07 | 137.65±2.50 | 144.17±5.40 | 139.14±2.07 | 138.76±2.29 |
| | P=0.746 | | | P=0.900 | |
| | TT[337] | TG[229] | GG[47] | TT[337] | TG/GG[276] |
| **Triglycerides Adjusted mean± SD** | 146.94±3.71 | 136.95±4.48 | 132.08±9.72 | 146.94±3.70 | 136.12±4.10 |
| **P-value** | **P=0.014** | | | P=**0.015** | |

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

Natural log transformed data were used for association analysis of triglycerides and HDL-C levels

**Table 14 (Continued)**

| | rs264_Add | | | rs264_Dom | |
|---|---|---|---|---|---|
| | GG[440] | GA[150] | AA[9] | GG[440] | GA/AA[159] |
| **Total Cholesterol Adjusted mean± SD** | 216.77±1.96 | 217.53±3.31 | 222.74±13.31 | 216.76±1.96 | 217.81±3.22 |
| **P-value** | P=0.709 | | | P=0.774 | |
| | GG[442] | GA[151] | AA[9] | GA[442] | GA/AA[160] |
| **HDL-C Adjusted mean± SD** | 50.20±0.60 | 51.03±1.02 | 50.24±4.12 | 50.2±0.6 | 50.99±0.99 |
| **P-value** | P=0.380 | | | P=0.370 | |
| | GG[438] | GA[148] | AA[9] | G[438] | GA/AA[157] |
| **LDL-C Adjusted mean± SD** | 138.77±1.82 | 138.18±3.08 | 147.98±12.31 | 138.75±1.82 | 138.72±3.00 |
| **P-value** | P=0.850 | | | P=0.993 | |
| | GG[442] | GA[148] | AA[9] | G[442] | GA/AA[157] |
| **Triglycerides Adjusted mean± SD** | 143.18±3.30 | 141.75±5.62 | 124.00±22.50 | 143.21±3.30 | 140.77±5.47 |
| **P-value** | P=0.660 | | | P=0.783 | |
| | **rs270_Add** | | | **rs270_Dom** | |
| | CC[419] | CA[174] | AA[19] | CC[419] | CA/AA[193] |
| **Total Cholesterol Adjusted mean± SD** | 215.90±2.02 | 219.26±3.04 | 231.19±9.08 | 215.91±2.02 | 220.45±2.89 |
| **P-value** | P=0.100 | | | P=0.187 | |
| | CC[422] | CA[174] | AA[19] | CC[422] | CA/AA[193] |
| **HDL-C Adjusted mean± SD** | 51.22±0.62 | 48.58±0.94 | 49.91±2.80 | 51.23±0.62 | 48.71±0.89 |
| **P-value** | **P=0.020** | | | **P=0.008** | |
| | CC[417] | CA[172] | AA[19] | CC[417] | CA/AA[191] |
| **LDL-C Adjusted mean± SD** | 137.53±1.87 | 141.17±2.83 | 153.95±8.49 | 137.54±1.87 | 142.26±2.69 |
| **P-value** | **P=0.056** | | | P=0.123 | |
| | CC[419] | CA[174] | AA[19] | CC[419] | CA/AA[193] |
| **Triglycerides Adjusted mean± SD** | 137.93±3.39 | 153.27±5.13 | 136.79±15.31 | 137.93±3.39 | 151.62±4.87 |
| **P-value** | **P=0.031** | | | **P=0.012** | |
| | **rs249_Add** | | | **rs249_Dom** | |
| | TT[523] | TC[79] | CC[6] | TT[523] | TC/CC[85] |
| **Total Cholesterol Adjusted mean± SD** | 217.39±1.83 | 214.25±4.55 | 247.51±16..36 | 217.34±1.83 | 216.54±4.41 |
| **P-value** | P=0.745 | | | P=0.864 | |
| | TT[526] | TC[79] | CC[6] | TT[526] | TC/CC[85] |
| **HDL-C Adjusted mean± SD** | 50.62±0.56 | 49.41±1.39 | 46.93±5.00 | 50.63±0.45 | 49.24±1.34 |
| **P-value** | P=0.343 | | | P=0.355 | |
| | TT[521] | TC[77] | CC[6] | TT[521] | TC/CC[83] |
| **LDL-C Adjusted mean± SD** | 139.39±1.70 | 134.85±4.26 | 170.56±15.15 | 139.34±1.70 | 137.36±4.13 |
| **P-value** | P=0.890 | | | P=0.653 | |
| | TT[524] | TC[78] | CC[6] | TT[524] | TC/CC[84] |
| **Triglycerides Adjusted mean± SD** | 140.82±3.03 | 147.94±7.61 | 149.88±27.24 | 140.82±3.03 | 148.08±7.35 |
| **P-value** | P=0.380 | | | P=0.4213 | |

Mean and *p*-values adjusted for BMI, age, smoking, and sex
Natural log transformed data were used for association analysis of triglycerides and HDL-C levels

**Table 14 (Continued)**

| | rs283_Add | | | rs283_Dom | |
|---|---|---|---|---|---|
| | AA[394] | AT[189] | TT[29] | AA[394] | AT/TT[218] |
| **Total Cholesterol Adjusted mean± SD** | 217.60±2.10 | 216.97±2.99 | 214.60±7.52 | 217.59±2.1 | 216.66±2.79 |
| **P-value** | P=0.717 | | | P=0.782 | |
| | AA[397] | AT[189] | TT[29] | AA[397] | AT/TT[218] |
| **HDL-C Adjusted mean± SD** | 50.68±0.64 | 50.09±0.91 | 48.17±2.29 | 50.68±064 | 49.83±0.85 |
| **P-value** | P=0.350 | | | P=0.402 | |
| | AA[392] | AT[187] | TT[29] | AA[392] | AT/TT[216] |
| **LDL-C Adjusted mean± SD** | 139.80±1.95 | 137.57±2.78 | 138.53±6.96 | 139.80±1.95 | 137.70±2.59 |
| **P-value** | P=0.560 | | | P=0.504 | |
| | AA[394] | AT[189] | TT[29] | AA[394] | AT/TT[218] |
| **Triglycerides Adjusted mean± SD** | 138.20±3.48 | 151.36±4.96 | 138.14±12.49 | 138.17±3.48 | 149.59±4.63 |
| **P-value** | **P=0.068** | | | **P=0.024** | |
| | **rs3289_Add** | | | **rs3289_Dom** | |
| | TT[584] | TC[32] | CC[0] | TT[584] | TC/CC[32] |
| **Total Cholesterol Adjusted mean± SD** | 217.66±1.75 | 209.97±7.09 | | 217.66±1.75 | 209.97±7.09 |
| **P-value** | P=0.289 | | | P=0.289 | |
| | TT[587] | TC[32] | CC[0] | TT[587] | TC/CC[32] |
| **HDL-C Adjusted mean± SD** | 50.44±0.53 | 50.42±2.17 | | 50.44±0.53 | 50.42±2.17 |
| **P-value** | P=0.876 | | | P=0.876 | |
| | TT[580] | TC[32] | CC[0] | TT[580] | TC/CC[32] |
| **LDL-C Adjusted mean± SD** | 139.62±1.62 | 129.13±6.55 | | 139.62±1.62 | 129.13±6.55 |
| **P-value** | P=0.118 | | | P=0.118 | |
| | TT[584] | TC[32] | CC[0] | TT[584] | TC/CC[32] |
| **Triglycerides Adjusted mean± SD** | 141.28±2.91 | 152.51±11.81 | | 141/28±2.91 | 152.51±11.81 |
| **P-value** | P=0.378 | | | P=0.378 | |
| | **rs3779787_Add** | | | **rs3779787_Dom** | |
| | GG[436] | GT[166] | TT[11] | GG[436] | GT/TT[177] |
| **Total Cholesterol Adjusted mean± SD** | 217.06±1.98 | 216.70±3.19 | 234±12.15 | 217.04±1.98 | 217.74±3.1 |
| **P-value** | P=0.582 | | | P=0.844 | |
| | GG[436] | GT[167] | TT[11] | GG[438] | GT/TT[178] |
| **HDL-C Adjusted mean± SD** | 49.73±0.60 | 52.06±0.97 | 48.94±3.69 | 49.74±0.6 | 51.88±0.94 |
| **P-value** | **P=0.033** | | | **P=0.018** | |
| | GG[433] | GT[165] | TT[11] | GG[433] | GT/TT[176] |
| **LDL-C Adjusted mean± SD** | 139.35 ±1.84 | 136.99±2.95 | 159.09±11.22 | 139.31±1.84 | 138.33±2.88 |
| **P-value** | P=0.830 | | | P=0.770 | |
| | GG[438] | GT[164] | TT[11] | GG[438] | GT/TT[175] |
| **Triglycerides Adjusted mean± SD** | 144.66±3.29 | 137.02±5.33 | 129.24±20.22 | 144.68±3.28 | 136.55±5.18 |
| **P-value** | P=0.225 | | | P=0.222 | |

Mean and *p*-values adjusted for BMI, age, smoking, and sex
Natural log transformed data were used for association analysis of triglycerides and HDL-C levels

# 4.0 DISCUSSION

Lipoprotein lipase (*LPL*) is a biological candidate gene for cardiovascular disease due to its major role in the catabolism of TG rich lipoproteins. Clinically abnormal lipid levels have been associated with CHD, atherosclerosis and obesity in individuals who have functional DNA sequence variation in their *LPL* gene (Murthy et al. 1996; Reymer et al. 1995; Brunzell et al. 1995; Wiebusch et al. 1992). The sequence variation of *LPL* is also associated with Mendelian disorders. *LPL* deficiency, known as Type 1 hyperlipoproteinemia or familial chylomicronemia (MIM 238600), is a rare autosomal recessive disorder (1/1,000,000) caused by defects in the *LPL* gene and it is correlated with severe hypertriglyceridemia due to chylomicronemia and VLDL accumulation with very low levels of LDL-C and HDL-C levels (<20mg/dL).

Several candidate gene studies and genome wide association studies (GWAS) have identified many SNPs in the *LPL* gene associated with lipid levels (Aulchenko et al. 2009; Kooner et al. 2008; Klos et al. 2006; Chasman et al. 2008; Ahn et al. 1993; Razzaghi et al. 2000; Willer et al. 2008; Chamberlein et al. 1999; Wallacee et al. 2008; Humphries et al. 1998). In this study, we resequenced the entire *LPL* gene (~30kb) in 95 healthy NHW individuals having extreme HDL-C to evaluate the impact of *LPL* genetic variation on HDL-C and correlated lipid levels. Our comprehensive resequencing effort of this gene, the first to our knowledge, revealed a total of 179 variants; 162 single nucleotide substitutions, 17 indels and a microsatellite (tetranucleotide repeat marker). Of 178 variants, 88 had a MAF≥0.05 and 91 had a MAF <0.05.

## 4.1 COMPARISON OF OUR RESEQUENCING RESULTS WITH TWO PUBLISHED STUDIES

We compared our sequencing results with two previous published sequencing reports (Nickerson et al. 1998; Wright et al. 2008). Nickerson et al. (1998) examined a portion of the *LPL* gene (9.7kb), from 3' end of intron 3 to 5' end of intron 9, among 71 individuals from three racial groups including European-Americans from Rochester, Minnesota (n=23), African-Americans from Jackson, Mississippi (n=24) and Europeans from North Karelia, Finland (n=24). They identified 88 variants, of which 79 were single nucleotide substitutions, 8 were indels and one was a microsatellite. Of 88 variant, 56 and 59 were identified in their white populations, North Karelia, Finland and Rochester, Minnesota, respectively. We compared our sequencing results with the variants they identified in Rochester samples. Fifty three of the 59 reported variants were founded in our NHW sample in this 9.7 kb region, including the reported microsatellite $(ATTT)_n$. Of the 6 reported variants, which we did not find in this study, 5 had MAF$\leq$0.05 and one was 8538delA with a MAF of 0.20 which was observed as a sequencing artifact in our sample. On the other hand, we identified 17 variants (0.005$\leq$MAF$\leq$0.053) that were not reported in their study. This inconsistency might be due to the sample size differences or due to the different software that were used in identifying variants (Nickerson et al. used PolyPhred Program, but we used Variant Reporter and Sequencher).

Wright et al. (2008) sequenced the 10 exons and intron-exon boundaries of the *LPL* gene, plus 1kb in the promoter region, and about 300 bases in the 3'flanking region in 19 Northern Irish individuals with extreme hypertriglyceridemia (HTG). They identified a total of 42 variants and found rs268 (N291S) to be a major predisposing factor for HTG. We identified 39 of the reported

42 variants, including N291S (discussed later on). The three variants they identified (590G>A, 1018G>A and 345A>C) but we did not, had MAF <0.05 in their Northern Irish sample. Two of these 3 variants were exonic (590(G>A) in exon 5 and 1018G>A in exon 6). Interestingly, these two variants were not identified by Nickerson et al. (1998), either. Since Wright et al. (2008) used a selected sample with HTG, it is not surprising that these variants were not identified in ours and Nickerson et al.'s population-based samples. On the other hand, although we identified 23 exonic variants in our sample, they identified only 15 exonic variants including the 2 we did not identify. Of 11 exonic variants that they did not identify, ten were located in the UTR of exon 10, and one was [rs328 (Ser447X)] located in exon 9.

## 4.2 DISTRIBUTION OF *LPL* VARIANTS AMONG HIGH AND LOW HDL-C GROUPS

### 4.2.1 Distribution of common *LPL* variants among high and low HDL-C groups in NHWs

Of the 88 common variants (MAF ≥0.05) observed in this study 32 showed significantly different distribution (P ≤0.05) between high and low HDL-C groups. Of these 32 variants, 25 have been previously reported in Chip Bioinformatics, and remaining 7 are the new variants.

Two of the common variants, rs326 (23858) and rs13702 (28911), that revealed significant difference between the high and low HDL-C groups in our study, have also been reported previously to be associated with HDL-C levels (Klos et al. 2006; Kooner et al. 2008). Klos et al. (2006) genotyped 3,993 individuals (1,132 black females, 807 black males, 1,101 white females and 953 white males) for rs326 and rs13702 variants and found significant association with HDL-C in black males (P=0.013) and black females (P=0.004). The association

between HDL-C levels and rs326 was also confirmed in a GWAS ($P=1.4\times10^{-7}$) comprising 5,968 individuals (Kooner et al. 2008). In both studies, the rs326/G allele was found to be associated with high HDL-C levels than the other allele. Likewise, in our study the frequency of the G allele for rs326 was higher in the high HDL-C group than the low HDL-group (0.278 vs. 0.133; P=0.017). As reported previously, the rs326 and rs13702 variants were in strong LD in our sample ($r^2=0.933$). Our tagger analysis showed that rs13702 and rs326 were in the same bin along with rs15285 (29086C>T) and rs3916027 (29287G>A), which also demonstrated significant difference between the high and low HDL-C groups (P=0.009 and P=0.012, respectively).

A common *LPL* variant in intron 8, *Hind*III (23496T>G, rs320), has been reported to be associated with plasma TG and HDL-C in several, but not all studies (Ahn et al. 1993; Razzaghi et al. 2000; Chamberlain et al. 1989) As expected, the frequency of the *Hind*III/G allele in the high HDL-C group was almost twice to that observed in the low HDL-C group (0.244 vs 0.128; P=0.053) in our sequencing sample, The *Hind*III/G allele is in strong LD with the 447X allele of the functional Ser447X polymorphism (rs328) in exon 9 (Razzaghi et al. 2000; Humphries et al. 1998;). The S447X variant, which has been reported as a gain-of-function mutation, leads to a premature stop codon, resulting a truncated protein which is two amino acid shorter than the full-length product that increases LPL protein expression without changing specific activity (Merkel et al. 2002; Rip et al. 2006). Wittrup et al. (1999) found that the carriers of Ser447X had reduced TGs, increased HDL-C and 0.8-fold reduced risk of ischemic heart disease. Since these two variants are in strong LD, there has been a debate in the literature whether *Hind*III is a genetic marker for the functional Ser447X variant or this is functional by itself (Razzaghi et al. 2000; Humphries et al. 1998; Merkel et al. 2002). However, Chen et al (2008) have recently demonstrated that the *Hind*III site binds to a transcription factor and affects LPL expression and

thus, this is a functional variant by itself. Although we identified the *Hind*III polymorphism in our sample, we found only one example of S447X (24143C>G) in the high HDL group using Variant Reporter. We further looked for this variant in Sequencher. However, we did not see any clear peaks as an evidence of variation at this site in Sequencher.

Another common variant *Pvu*II (rs285) that we identified in our sequencing sample has been previously reported to be associated with TG levels and hypertriglyceridemia in some but not all studies (Chamberlein et al. 1999; Ahn et al. 1993). However, in our sequencing sample we found a comparable frequency of the T allele between the high and low HDL-C groups (0.444 vs 0.370; P =0.304).

To our knowledge, seven GWA studies have found significant association signals within the *LPL* gene and flanking regions associated with HDL-C levels (Boes et al. 2010). The reported SNPs with the lowest p-values are: rs2083637 ($P=5.5\times10^{-18}$) (Aulchenko et al. 2009), rs10503669 ($P =4.1\times10^{-19}$) (Willer et al. 2008), rs331 ($P=9.1\times10^{-7}$) (Chasman et al. 2008), rs328 ($P =9\times10^{-23}$) (Kathiresan et al. 2008b), rs17482753 ($P =2.71\times10^{-5}$) (Heid et al. 2008), rs17411031 ($P =1.28\times10^{-10}$) (Wallacee et al. 2008), and rs326 ($P =1.4\times10^{-7}$) (Kooner et al. 2008). Of these 7 SNPs, we identified rs326 and rs331 SNPs in our study. As mentioned above, our result for rs326 was consistent with reported GWAS, but we did not find a significant difference between high and low HDL-C groups for rs331 (0.154 vs. 0.096; P =0.247), although the frequency of the A allele was higher in high HDL-C group which is consistent with the previous reports. The other 5 GWAS significant SNPs are located in the distant 5'flanking region of the *LPL* that was outside the region we sequenced.

### 4.2.2 Distribution of relatively uncommon or rare *LPL* variants among high and low HDL-C groups in NHWs

In addition to the 88 common variants identified in sequencing, we identified 91 uncommon or rare variants (MAF< 0.05). Of the 91 relatively uncommon or rare variants, 21 were present only in the low HDL-C group and 25 were present only in the high HDL-C group; 45 exist in both groups. Since the P-values for the uncommon or rare variants (MAF< 0.05) between the two HDL groups were unreliable due to the small resequencing sample size (47 in high HDL and 48 in low HDL), we performed a preliminary analysis for 91 variants to determine the accumulation of rare variants in the high and low HDL-C group. Forty one of 47 (87.2%) individuals with high HDL-C had minimum one rare variant versus 35 of 48 (72.91%) individuals with low HDL-C; 24 of 47 (51.1%) individuals high HDL-C had minimum two rare variants versus 16 of 48 (33.3%) individuals with low HDL-C; 18 of 47 (38.2%) individuals with high HDL-C had minimum three rare variants versus 8 out of 48 (16.6%) individuals with low HDL-C. Overall, the prevalence of uncommon or rare variants was higher in the high HDL-C group than the low HDL-C group, which contradicts observation made by Cohen et al. (2004) by sequencing three other lipid genes. Cohen et al (2004) resequenced the *ABCA1*, *APOA1* and *LCAT* genes in individuals having extremely low ($5^{th}$ percentile) and high ($95^{th}$ percentile) HDL-C levels from a population-based sample to test whether the accumulation of rare variants collectively contribute to variation in HDL-C levels. They observed accumulation of rare sequence variants in subjects with low HDL-C levels and concluded that multiple rare variants were significant contributor to low HDL-C levels.

To date, in Online Mendelian Inheritance in Man database (OMIM), (http://www.ncbi.nlm.nih.gov/omim) a list of 41 causative mutations or rare variants have been reported in *LPL*; 35 mutations were reported causing LPL deficiency; 4 were reported causing Combined familial hyperlipidemia (CFH); one single nucleotide polymorphism (SNP) (Ser447X) reported as contributing a defect in lipid interface recognition in a Type 1 hyperlipidemia patient and another SNP (rs326), a HDL-C level quantitative trait locus 11 (HDLCQ11) was reported for being associated with high HDL-C levels. Not surprisingly, we did not identify any example of the reported 35 rare mutation associated with LPL deficiency because our sample was population-based from apparently healthy individuals. However, of the 4 reported variants associated with CFH, three were identified in our resequencing sample; D9N (rs1801177), -T93G (rs1800590) and N291S (rs268). Both coding variants, D9N and N291S, are associated with reduced enzymatic activty. –T93G is a functional promoter variant located in the 5' flanking region of the *LPL* gene and is in LD with D9N among Caucasians (Zhang et al. 1996; Mailly et al. 1995; Merkel et al. 2002). The minor allele frequency of these variants in our resequencing sample was ≤0.05 (Table 15) and so there were not enough individuals in the high and low HDL groups to make a conclusive statement about their role in affecting HDL-C levels. However, these variants will be genotyped in our entire sample in addition to other rare variants that we identified in the current study to evaluate their association with lipid levels.

**Table 15. Distribution of identified functional variants in high and low HDL-C groups**

| Nucleotide # | Alleles | Location | refSNP ID | All MAF (n=95) | MAF(%) assoc. allele High HDL-C | MAF(%) assoc.allele Low HDL-C |
|---|---|---|---|---|---|---|
| **1090** | T>G | 5' flanking | rs1800590 (-T93G) | **0.011** | 0.011 | 0.010 |
| **10127** | G>A | **Exon 2** | rs1801177 (D9N) | **0.016** | 0.021 | 0.010 |
| **17948** | A>G | **Exon 6** | rs268 (N291S) | **0.033** | 0.022 | 0.043 |
| **24143** | C>G | **Exon 9** | rs328(Ser447X) | **0.005** | 0.011 | 0.000 |

## 4.3 DISTRIBUTION OF SELECTED COMMON VARIANTS IN THE TOTAL NHW SAMPLE

To date, we have screened a total of 12 common variants in the entire NHW sample (n=623) by using TaqMan Genotyping assays (Table 6 and Table 7 in Section 2.3). Of the 12 variants, rs1059611, rs270, rs3779787 and rs327 showed statistically significant association with HDL-C level ($P<0.05$) in both additive and dominant models, but rs13266204 had a significant P-value (0.013) only in the additive model, but showed a marginally significant association (P=0.080) in the dominant model. To our knowledge, none of these SNPs were previously reported as being significantly associated with HDL-C levels. In our resequencing samples (n=95), we also observed significantly different distribution of rs1059611 (28982T>C), rs270 (18095C>A), rs327 (23955T>G) and rs13266204 (4424A>G) between the high and low HDL-C groups (P=0.003, 0.045, 0.018, and 0.038, respectively). Although we observed significant association of rs3779787 in the total sample ($P_{\text{additive model}}=0.033$ and $P_{\text{dominant model}}=0.018$) the difference of minor allele frequencies between the high and low HDL-C groups was marginally significant

(P=0.096). This can be explained by the sample size differences in genotyping and resequencing. For rs3779787, T allele was associated with high HDL-C levels in our genotyping results ($P_{\text{additive model}}$=0.033 and $P_{\text{dominant model}}$=0.018) and T allele frequency was also higher in high HDL-C group than low HDL-C group, although it did not achieve statistical significance at 5% level (P=0.096). Noteworthy, although we identify only one example of the Ser447X variant in our resequencing sample, it was in LD with rs1059611 which was genotyped in our entire NHW sample and showed significant association with HDL-C levels, as discussed above.

Two of the 12 variants, rs327 (T>G), rs270 (C>A), showed statistically significant association with TGs levels in both additive (P=0.014 and P=0.031, respectively) and dominant models (P=0.015 and P=0.012, respectively). rs312 (G>C) also showed statistically significant association with TG levels in additive model (P=0.009), but it showed marginally significant association in dominant model (P=0.050).The rs327/T and rs270/A alleles were associated with high TG levels as compared to other corresponding alleles of these two SNPs. Both of these variants have also demonstrated significant differences between the high and low HDL-C groups in the subset sequencing sample, as discussed above. Recently, the association between TG levels and rs327 was also reported by Smith et al. (2010), but they did not observe significant association between rs270 and TG levels in any of the statistical models that they used. In their study, rs327 was shown to affect regulation of *LPL* in vitro and they reported this as a new potential functional variant of *LPL* (Smith et al. 2010). We also observed association of rs283 with TG levels (P=0.024); however this association was not identified by Smith et al. (2010).

There was one variant, rs270 (C>A), that showed statistically significant association (P=0.020) with LDL-C levels; rs270/A allele was associated with high LDL-C levels. The same allele was also associated with low HDL-C (P=0.020 and 0.008, for additive and dominant

models, respectively) and high TG levels (P=0.031 and 0.012 for additive and dominant models, respectively). Additionally, rs312 (G>C) showed marginally significant association with TG levels in only dominant model (P=0.087); the G allele was associated with low LDL-C levels for this variant.

Previously, rs3289 located in UTR of exon 10, has shown to affect LPL regulation in vitro and also revealed significant association with TG levels (Smith et al. 2010). However, in our study we found no such association with TG or other lipid traits. This might be due to different sample sizes and sample selection criteria that were used in genotyping.

# 5.0 CONCLUSIONS

*LPL* plays a vital role in lipid metabolism and genetic variation in this gene affects monogenic lipid traits and also HDL-C and TG variation in the general population. To further investigate the role of the rare and common variation in the *LPL* gene, we resequenced the entire gene in selected individuals falling in the upper and lower $5^{th}$ percentile of HDL-C in a NHW sample. We identified 179 variants, of which 88 had MAF≥0.05 and 91 had MAF<0.05. Of the 88 common variants, 32 demonstrated significant association with HDL-C levels. Overall, the prevalence of uncommon or rare variants was higher in the high HDL-C group than the low HDL-C group. Our comprehensive resequencing study reconfirms the functional significance of the *LPL* gene in lipid metabolism.

## BIBLIOGRAPHY


Acton S, Rigotti A, Landschulz KT, Xu S, Hobbs HH, Krieger M. Identification of scavenger receptor SR-BI as a high density lipoprotein receptor. *Science* 1996; 271:518–520.

Ahn YI, Kamboh MI, Hamman RF, Cole SA, Ferrell RE. Two DNA polymorphisms in the lipoprotein lipase gene and their associations with factors related to cardiovascular disease. *J Lipid Res* 1993; 34:421-428.

American Heart Association. Heart and Stroke Statistics 2010 Update.

Austin MA, King MC, Bawol RD, Hulley SB, Friedman GD. Risk factors for coronary heart disease in adult female twins. Genetic heritability and shared environmental influences. *Am J Epidemiol* 1987; 125:308-318.

Attie AD. High-maintenance proteins and hypertriglyceridemia *Nat Genet* 2007; 39;1424-1425

Aulchenko YS, Ripatti S, Lindqvist 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; 41:5-6.

Barr DP, Russ EM, Eder HA. Protein-lipid relationships in human plasma. II. In atherosclerosis and related conditions. *Am J Med* 1951; 11:480-93.

Barter P, Kastelein J, Nunn A, Hobbs R High densitylipoproteins (HDLs) and atherosclerosis; the unanswered questions. *Atherosclerosis* 2003; 168:195-211

Beisiegel U, Weber W, Bengtsson-Olivecrona G. Lipoprotein lipase enhances the binding of chylomicrons to low density lipoprotein receptor-related protein *Proc Natl Acad Sci USA* 1991; 88:8342-8346.

Bengtsson-Olivecrona G, Olivecrona T, Jörnvall H. Lipoprotein lipases from cow, guinea-pig and man. Structural characterization and identification of protease-sensitive internal regions. *Eur J Biochem* 1986; 161:281-8.

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

Breslow JL. Genetic markers for coronary theart disease. *Clin Cardiol* 2001; 24: 14-17.

Brunzell JD. Familial lipoprotein lipase deficiency and other causes of chylomicronemia syndrome. in *The Metabolic and Molecular Basis of Inherited Diseases* (eds Scriver, C.R., Beaudet, A.L., Sly, W.S. & Valle, D.) 1913–1932 (McGraw-Hill Inc., New York, 1995).

Chasman DI Paré G, Zee RY, Parker AN, Cook NR, Buring JE, Kwiatkowski DJ, Rose LM, Smith JD, Williams PT, Rieder MJ, Rotter JI, Nickerson DA, Krauss RM, Miletich JP, Ridker PM. 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.

Chamberlain JC, Thorn JA, Oka K, Galton DJ, Stocks J. DNA polymorphisms at the lipoprotein lipase gene: associations in normal and hypertriglyceridaemic subjects. *Atherosclerosis* 1999; 79:85-91.

Chen TJ, Ji CY, Hu YH. Genetic and environmental influences on serum lipids and the effects of puberty: a Chinese twin study. Acta Paediatr 2009; 98:1029-1036.

Chen Q, Razzaghi H, Demirci FY, Kamboh MI. Functional significance of lipoprotein lipase HindIII polymorphism associated with the risk of coronary artery disease. *Atherosclerosis* 2008; 200:102-108.

Costanza MC, Cayanis E, Ross BM, Flaherty MS, Alvin GB, Das K, Morabia A. Relative contributions of genes, environment, and interactions to blood lipid concentrations in a general adult population. *Am J Epidemiol* 2005; 161:714-724.

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.

Cooper RS, Ford E. Comparability of risk factors for coronary heart disease among blacks and whites in the NHANES I Epidemiologic Follow-up Study. *Ann Epidemiol* 1992; 2:637.

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.

Deeb SS, Peng RL. Structure of the human lipoprotein lipase gene. *Biochemistry* 1989; 28:4131-4135.

Demirci FY, Dressen AS, Hamman RF, Bunker CH, Kammerer CM, Kamboh MI. Association of a common G6PC2 variant with fasting plasma glucose levels in non-diabetic individuals. *Ann Nutr Metab* 2010; 56:59-64.

Eckel RH. Obesity and heart disease: a statement for healthcare professionals from the Nutrition Committee, American Heart Association. *Circulation* 1997; 96:3248-3250.

Ehrenborg E et al. Ethnic variation and in vivo effects of the -93t-->g promoter variant in the lipoprotein lipase gene. *Arterioscler Thromb Vasc Biol* 1997; 17:2672-2678.

Florentin M, Liberopoulos EN, Wierzbicki AS, Mikhailidis DP. Multiple actions of high-density lipoprotein. *Curr Opin Cardiol* 2008; 23:370-378.

Genest J, Marcil M, Denis M, Yu L. High density lipoproteins in health and disease. *J Invest Med* 1999; 47:31– 42.

Gimble JM, Hua X, Wanker F, Morgan C, Robinson C, Hill MR, Nadon N. In vitro and in vivo analysis of murine lipoprotein lipase gene promoter: tissue-specific expression. *Am J Physiol Endocrinol Metab* 1995; 268:213–218

Gordon T, Castelli WP, Hjortland MC, Kannel WB, Dawber TR. High density lipoprotein as a protective factor against coronary heart disease. The Framingham Study. *Am J Med* 1977; 62:707-714.

Goldberg IJ. Lipoprotein lipase and lipolysis: central roles in lipoprotein metabolism and atherogenesis. *J Lipid Res* 1996; 37:693-707.

Gordon D, Rifkind BM. Current concepts: high density lipoproteins–the clinical implications of recent studies. *N Engl J Med* 1989; 321

Goode EL, Cherny SS. Christian JC, Heritability of longitudinal measures of body mass index and lipid and lipoprotein levels in aging twins. *Twin Res Hum Genet* 2007; 10:703-711.

Grundy SM, Pasternak R, Greenland P, Smith S Jr, Fuster V. AHA/ACC scientific statement: Assessment of cardiovascular risk by use of multiple-risk-factor assessment equations: a statement for healthcare professionals from the American Heart Association and the American College of Cardiology. *J Am Coll Cardiol* 1999; 34:1348-1359.

Hahn PF. Abolishment of alimentary lipemia following injection of heparin. *Science* 1943; 98:19-20.

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.

Havel RJ, Gordon JR. Idiopathic hyperlipemia: metabolic studies in an affected family. *J Clin Invest* 1960; 39:1777-1790.

Henderson HE, Ma Y, Hassan MF, Monsalve MV. Amino acid substitution (Ile194----Thr) in exon 5 of the lipoprotein lipase gene causes lipoprotein lipase deficiency in three unrelated probands. Support for a multicentric origin. *J Clin Invest* 1991; 87:2005-2011.

Heid IM, et al. Genome-wide association analysis of high-density lipoprotein cholesterol in the population-based KORA study sheds new light on intergenic regions. *Circ Cardiovasc Genet* 2008; 1:10-20.

Heron M, Hoyert DL, Murphy SL, Xu J, Kochanek KD, Tejada-Vera B. Deaths: final data for 2006. *Natl Vital Stat Rep* 2009; 57:1-134.

Humphries SE, Nicaud V, Margalef J, Tiret L, Talmud PJ. Lipoprotein lipase gene variation is associated with a paternal history of premature coronary artery disease and fasting and postprandial plasma triglycerides: the European Atherosclerosis Research Study (EARS). *Arterioscler Thromb Vasc Biol* 1998; 18:526-534.

Iannotti RJ, Zuckerman AE, Rifai N. Correlations of cardiovascular disease risk factors between African-American siblings. *J Pediatr* 2000; 136:111–119.

Johnson JL, Heineman EF, Heiss G, Hames CG, Tyroler HA. Cardiovascular disease risk factors and mortality among black women and white women aged 40–64 years in Evans County, Georgia. *Am J Epidemio*l 1986; 123:209.

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* 2008a; 40:189–197.

Kathiresan S, Melander O, Anevski D, Guiducci C, Burtt NP, Roos C, et al., Polymorphisms associated with cholesterol and risk of cardiovascular events. *N Engl J Med* 2008b; 1240–1249.

Kathiresan S, Manning AK, Demissie S, D'Agostino RB, Surti A, Guiducci C, Gianniny L, Burtt NP, Melander O, Orho-Melander M, Arnett DK, Peloso GM, Ordovas JM, Cupples LA. A genome-wide association study for blood lipid phenotypes in the Framingham Heart Study. BMC Med Genet 2007; (8 Suppl) 1:S17.

Keil JE, Loadholt CB, Weinrich MC, Sandifer SH,Boyle E Jr. Incidence of coronary heart disease in blacks in Charleston, South Carolina. *Am Heart J* 1984; 108:779.

Kirchgessner TG, Chuat JC, Heinzmann C, Etienne J, Guilhot S, Svenson K, Ameis D, Pilon C, d'Auriol L, Andalibi A. Organization ofthe human lipoprotein lipase gene and evolution of the lipase gene family. *Proc Natl Acad Sci USA* 1989; 86:9647–9651.

Kooner JS, Chambers JC, Aguilar-Salinas CA, et al. Genome-wide scan identifies variation in MLXIPL associated with plasma triglycerides. *Nat Genet* 2008; 40:149-151.

Klos KL, Sing CF, Boerwinkle E, Hamon SC, Rea TJ, Clark A, Fornage M, Hixson JE. Consistent effects of genes involved in reverse cholesterol transport on plasma lipid and apolipoprotein levels in CARDIA participants. *Arterioscler Thromb Vasc Biol* 2006; 26:1828-1836.

Klos KL, Kullo IJ. Genetic determinants of HDL: monogenic disorders and contributions to variation. *Curr Opin Cardiol* 2007; 22:344-51.

Kronenberg F, Coon H, Ellison RC, et al. Segregation analysis of HDL cholesterol in the NHLBI Family Heart Study and in Utah pedigrees. *Eur J Hum Genet* 2002; 10:367-374.

Kounnas MZ, Chappell DA, Strickland DK, Argraves WS. Glycoprotein 330, a member of the low density lipoprotein receptor family, binds lipoprotein lipase in vitro. *J Biol Chem* 1993; 268:14176-14181.

Komurcu-Bayrak E, Onat A, Poda M, Humphries SE, et al. The S447X variant of lipoprotein lipase gene is associated with metabolic syndrome and lipid levels among Turks. *Clin Chim Acta* 2007; 383:110-115.

Knoblauch H, Busjahn A, Münter S, Nagy Z, Faulhaber HD, Schuster H, Luft FC. Heritability analysis of lipids and three gene loci in twins link the macrophage scavenger receptor to HDL cholesterol concentrations. *Arterioscler Thromb Vasc Biol* 1997; 17:2054-2060.

Lookene A, Bengtsson-Olivecrona G. Chymotryptic cleavage of lipoprotein lipase. Identification of cleavage sites and functional studies of the truncated molecule. *Eur J Biochem* 1993; 213:185-194.

Lloyd-Jones D, Adams RJ, Brown TM, Carnethon M,. Heart disease and stroke statistics--2010 update: a report from the American Heart Association. *Circulation* 2010;121:e46-e215.

Marenberg ME, Risch N, Berkman LF, Floderus B, De Faire U. Genetic susceptibility to death from coronary heart disease in a study of twins. *N Engl J Med* 1994; 330:1041-1046.

Mailly F, Tugrul Y, Reymer PW, Bruin T, Seed M, Groenemeyer BF, Asplund-Carlson A, Vallance D, Winder AF, Miller GJ, et al. A common variant in the gene for lipoprotein lipase (Asp9-->Asn). Functional implications and prevalence in normal and hyperlipidemic subjects. *Arterioscler Thromb Vasc Biol* 1995; 15:468-478.

Medh JD, Bowen SL, Fry GL, Ruben S, Andracki M, Inoue I, Lalouel JM, Strickland DK, Chappell DA. Lipoprotein lipase binds to low density lipoprotein receptors and induces receptor-mediated catabolism of very low density lipoproteins in vitro. *J Biol Chem* 1996; 271:17073-17080.

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

Merkel M, Weinstock PH, Chajek-Shaul T, Radner H, Yin B, Breslow JL, Goldberg IJ. Lipoprotein lipase expression exclusively in liver. A mouse model for metabolism in the neonatal period and during cachexia. *J Clin Invest* 1998a; 102:893-901.

Merkel M, Kako Y, Radner H, Cho IS, Ramasamy R, Brunzell JD, Goldberg IJ, Breslow JL. Catalytically inactive lipoprotein lipase expression in muscle of transgenic mice increases very low density lipoprotein uptake: direct evidence that lipoprotein lipase bridging occurs in vivo. *Proc Natl Acad Sci USA* 1998b; 95:13841-13846.

Merkel M, Heeren J, Dudeck W, Rinninger F, Radner H,et al. Inactive lipoprotein lipase (LPL) alone increases selective cholesterol ester uptake in vivo, whereas in the presence of active LPL it also increases triglyceride hydrolysis and whole particle lipoprotein uptake. *J Biol Chem* 2002; 277:7405-7411.

Morabia A, Cayanis E, Costanza MC,et al. Association between lipoprotein lipase (*LPL*) gene and blood lipids: a common variant for a common trait? *Genet Epidemiol* 2003; 24:309-321.

Murthy V, Julien P, Gagne C. Molecular pathobiology of the human lipoprotein lipase gene. *Pharmacol Ther* 1996; 70:101–135.

Nettleton JA, Steffen LM, Ballantyne CM, Boerwinkle E, Folsom AR, Associations between HDL-cholesterol and polymorphisms in hepatic lipase and lipoprotein lipase genes are modified by dietary fat intake in African American and White adults. *Atherosclerosis* 2007;194, e131–e140.

Nickerson DA, Taylor SL, Weiss KM,et al. DNA sequence diversity in a 9.7-kb region of the human lipoprotein lipase gene. *Nat Genet* 1998; 19:233-240

Péterfy M, Ben-Zeev O, Mao HZ, Weissglas-Volkov D, Aouizerat BE, Mutations in LMF1 cause combined lipase deficiency and severe hypertriglyceridemia. *Nat Genet* 2007; 39:1483-1487.

Peter P. Toth. The "Good Cholesterol": High-Density Lipoprotein. *Circulation* 2005; 111:e89-e91

Pallaud C, Gueguen, Sass C, Grow M, Cheng S, Siest G, et al. Genetic influences on lipid metabolism trait variability within the Stanislas Cohort. *J.Lipid Res* 2001; 42:1879–1890.

Preiss-Landl K, Zimmermann R, Hämmerle G, Zechner R. Lipoprotein lipase: the regulation of tissue specific expression and its role in lipid and energy metabolism. *Curr Opin Lipidol* 2002; 13:471-481.

Pérusse L, Rice T, Després JP, Bergeron J, Province MA, Gagnon J, Leon AS, Rao DC, Skinner JS, Wilmore JH, Bouchard C. Familial resemblance of plasma lipids, lipoproteins and postheparin lipoprotein and hepatic lipases in the HERITAGE Family Study. *Arterioscler Thromb Vasc Biol* 1997; 17:3263-3269.

Razzaghi H, Aston CE, Hamman RF, Kamboh MI. Genetic screening of the lipoprotein lipase gene for mutations associated with high triglyceride/low hdl-cholesterol levels. *Hum Genet* 2000; 107:257-267.

Reymer PW. et al. A lipoprotein lipase mutation (Asn291Ser) is associated with reduced HDL cholesterol levels in premature atherosclerosis. *Nature Genet* 1995; 10:28–33.

Radha V, Mohan V, Vidya R, Ashok AK, Deepa R, Mathias RA.Association of lipoprotein lipase Hind III and Ser 447 Ter polymorphisms with dyslipidemia in Asian Indians. *Am J Cardiol* 2006; 97:1337–1342.

Ridker PM, Paré G, Parker AN, et al. Polymorphism in the CETP gene region, HDL cholesterol, and risk of future myocardial infarction: Genomewide analysis among 18 245 initially healthy women from the Women's Genome Health Study Circ *Cardiovasc Genet* 2009; 2:26-33.

Rip J, Nierman MC, Ross CJ, Jukema JW, Hayden MR, Kastelein JJ, Stroes ES, Kuivenhoven JALipoprotein lipase S447X: a naturally occurring gain-of-function mutation. *Arterioscler Thromb Vasc Biol* 2006; 26:1236-1245

.

Rye KA, Bursill CA, Lambert G, Tabet F, Barter PJ. The metabolism and anti-atherogenic properties of HDL. *J Lipid Res* 2009; 50 (Suppl):S195-S200.

Rust S, Walter M, Funke H, von Eckardstein A, Cullen P, Kroes HY, et al. Assignment of Tangier disease to chromosome 9q31 by a graphical linkage exclusion strategy. *Nat Genet* 1998; 20:312

Sabatti C, Service SK, Hartikainen AL, Pouta A, Ripatti S, et al. Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. *Nat Genet* 2008; 41:35-46.

Santamarina-Fojo S, Dugi KA. Structure, function and role of lipoprotein lipase in lipoprotein metabolism. *Curr Opin Lipidol* 1994; 5:117-125.

Sattler W, Levak-Frank S, Radner H, Kostner GM, Zechner R. Muscle-specific overexpression of lipoprotein lipase in transgenic mice results in increased alpha-tocopherol levels in skeletal muscle. *Biochem J* 1996; 318:15-19.

Seo T, Al-Haideri M, Treskova E, Worgall TS, Lipoprotein lipase-mediated selective uptake from low density lipoprotein requires cell surface proteoglycans and is independent of scavenger receptor class B type 1. *J Biol Chem* 2000; 275:30355-30362.

Semenkovich CF, Chen SH, Wims M, Luo CC, Li WH, Chan L. Lipoprotein lipase and hepatic lipase mRNA tissue specific expression, developmental regulation, and evolution. *J Lipid Res* 1989; 30:423-431.

Smith AJ, Palmen J, Putt W, Talmud PJ, Humphries SE, Drenos F. Application of statistical and functional methodologies for the investigation of genetic determinants of coronary heart disease biomarkers: lipoprotein lipase genotype and plasma triglycerides as an exemplar. *Hum Mol Genet* 2010; 19:3936-3947.

Sparkes RS, Zollman S, Klisak I, Kirchgessner TG, Komaromy MC, Mohandas T, Schotz MC, Lusis AJ. Human genes involved in lipolysis of plasma lipoproteins: mapping of loci for lipoprotein lipase to 8p22 andhepatic lipase to 15q21. *Genomics* 1987; 1:138-144.

Stein O, Stein Y. Atheroprotective mechanisms of HDL. *Atherosclerosis* 1999; 144:285-303.

Sviridov D, Nestel PJ. Genetic factors affecting HDL levels, structure, metabolism and function. *Curr Opin Lipidol* 2007; 18:157-163.

Takahashi S, Suzuki J, Kohno M, Oida K, Tamai T, Enhancement of the binding of triglyceride-rich lipoproteins to the very low density lipoprotein receptor by apolipoprotein E and lipoprotein lipase. *J Biol Chem* 1995; 270:15747-15754.

Talmud PJ, Hawe E, Robertson K, Miller GJ, Miller NE, Humphries SE, Genetic and environmental determinants of plasma high density lipoprotein cholesterol and apolipoprotein AI concentrations in healthy middle-aged men. *Ann Hum Genet* 2002; 66:111-124.

Talmud PJ, Stephens JW. Lipoprotein lipase gene variants and the effect of environmental factors on cardiovascular disease risk. *Diabetes Obesity Metab* 2004; 6:1-7.

Tiret L. Gene-environment interaction: a central concept in multifactorial diseases. *Proc Nutr Soc* 2002; 61:457–463.

Tsompanidi EM, Brinkmeier MS, Fotiadou EH, Giakoumi SM, Kypreos KE. HDL biogenesis and functions: role of HDL quality and quantity in atherosclerosis. *Atherosclerosis* 2010; 208:3-9.

Tsutsumi K. Lipoprotein Lipase and Atherosclerosis *Curr Vasc Pharmacol* 2003; 1:11-17.

Ukkola O, Garenc C, Pérusse L,et al. Genetic variation at the lipoprotein lipase locus and plasma lipoprotein and insulin levels in the Québec Family Study. *Atherosclerosis* 2001; 158:199-206.

Ulrich S, Hingorani AD, Martin J, Vallance P. Lifetime risk of developing coronary heart disease. *Lancet* 1999; 353:925.

Van Bennekum AM, Kako Y, Weinstock PH, Harrison EH. Lipoprotein lipase expression level influences tissue clearance of chylomicron retinyl ester. *J Lipid Res* 1999; 40:565-574.

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

Von Eckardstein A, Assmann G. Prevention of coronary heart disease by raising of high density lipoprotein cholesterol. *Curr Opin Lipidol* 2000; 11:627-637.

Vergeer M, Holleboom AG, Kastelein JJ, Kuivenhoven JA. The HDL hypothesis: does high-density lipoprotein protect from atherosclerosis? *J Lipid Res* 2010; 51:2058-2073.

Wallace C, Newhouse SJ, Braund P, Zhang F, Tobin M, Falchi M, Ahmadi K, Dobson RJ, Marçano AC, Hajat C, Burton P, Deloukas P, Brown M, Connell JM, Dominiczak A, Lathrop GM, Webster J, Farrall M, Spector T, Samani NJ, Caulfield MJ, Munroe PB. Genome-wide association study identifies genes for biomarkers of cardiovascular disease: serum urate and dyslipidemia. *Am J Hum Genet* 2008; 82:139–149.

Wang H, Eckel RH. Lipoprotein lipase: from gene to obesity. A*m J Physiol Endocrinol Metab* 2009; 297

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

Weissglas-Volkov D, Pajukanta P. Genetic causes of high and low serum HDL-cholesterol. *J Lipid Res* 2010; 51:2032-2057.

Wion KL, Kirchgessner TG, Lusis AJ. Human lipoprotein lipase complementary DNA sequence. *Science* 1987; 235:1638-1641.

Wiebusch H, Funke H, Kastelein JJP. Mutations in the lipoprotein lipase gene are not restricted to patients with type 1 hyperlipidemia. *Circulation* 1992; 86:609.

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.

Wittekoek ME, A common mutation in the lipoprotein lipase gene (N291S) alters the lipoprotein phenotype and risk for cardiovascular disease in patients with familial hypercholesterolemia. *Circulation* 1998; 97:729-735.

Wittrup HH, Tybjaerg-Hansen A, Steffensen R, Deeb SS, Brunzell JD, Jensen G, Nordestgaard BG. Mutations in the lipoprotein lipase gene associated with ischemic heart disease in men. The Copenhagen city heart study. *Arterioscler Thromb Vasc Biol* 1999; 19:1535-40.

Wright WT, Young IS, Nicholls DP, Graham CA. Genetic screening of the LPL gene in hypertriglyceridaemic patients. *Atherosclerosis* 2008; 199:187-192.

Xie C, Wang ZC, Liu XF, Yang MS. The common biological basis for common complex diseases: evidence from lipoprotein lipase gene. *Eur J Hum Genet* 2010; 18:3-7.

Yacoub LK, Vanni TM, Goldberg IJ. Lipoprotein lipase mRNA in neonatal and adult mouse tissues: comparison of normal and combined lipase deficiency (cld) mice assessed by in situ hybridization. *J Lipid Res* 1990; 31:1845-1852.

Yang CY, Gu ZW, Yang HX, Rohde MF, Gotto AM Jr, Pownall HJ. Structure of bovine milk lipoprotein lipase. J *Biol Chem* 1989; 264:16822-16827.

Zhang H, Henderson H, Gagne SE, Clee SM, Miao L, Liu G, Hayden MR. Common sequence variants of lipoprotein lipase: standardized studies of in vitro expression and catalytic function. *Biochim Biophys Acta* 1996; 1302:159-166.