Investigating the Effect of *PCSK9* Variants on Plasma Lipid Profile in African Blacks and U.S. Whites

by

Courtney Minoli Kasturiarachi

B.S. in Anthropology, Kent State University, 2020

Submitted to the Graduate Faculty of the

Department of Human Genetics

School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Public Health

University of Pittsburgh

2022

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This essay is submitted

by

Courtney Minoli Kasturiarachi

on

April 29, 2022

and approved by

Essay Advisor Essay Advisor: M. Ilyas Kamboh, PhD, Professor, Department of Human Genetics, School of Public Health, University of Pittsburgh

Kang-Hsien (Frank) Fan, PhD, Research Assistant Professor, Department of Human Genetics, School of Public Health, University of Pittsburgh

Allison L. Kuiper, PhD, Associate Professor, Department of Epidemiology, School of Public Health, University of Pittsburgh

Investigating the Effect of *PCKS9* Variants on Plasma Lipid Profile in African Blacks and U.S. Whites

Courtney Kasturiarachi, MPH

University of Pittsburgh, 2022

Abstract

In the United States, cardiovascular disease (CVD) is the leading cause of death and is a major public health issue. A combination of genetic and environmental factors leads to an increased risk of CVD—specifically, abnormal plasma lipid levels. Familial Hypercholesterolemia (FH) is a genetic condition that greatly increases the risk of CVD due to excess low-density lipoprotein cholesterol (LDL-C) in blood. Autosomal dominant mutations in the *LDLR, APOB,* and *PCSK9* genes cause FH. *PCSK9* mutations are the most rare and unique. Studies have found that *PCSK9* gain-of-function mutations result in FH, while loss-of-function mutations result in low plasma LDL-C levels. Results from *PCSK9* research aided in the creation of *PCSK9* inhibitors as a new therapy for FH. In this study, we sought to confirm the relationship between *PCSK9* variants and plasma lipid levels in two population-based samples not previously investigated for *PCSK9* variants. Study samples are composed of 788 African Blacks from Nigeria and 623 non-Hispanic Whites (NHWs) from the United States. *PCSK9* genotyping was completed using genome-wide chip followed by genotype-phenotype association analyses for total cholesterol (TC), LDL-C, high-density lipoprotein cholesterol (HDL-C) and triglycerides (TG). Of the 93 genotyped SNPs in the *PCSK9* region, 60 SNPs showed variations in either African Blacks (37), NHWs (4), or both (19). Seven SNPs in African Blacks and two SNPs in NHWs showed significant genotype associations with a least one lipid trait. In African Blacks, two SNPs were associated with TC, two with TG, two with HDL-C, and four with LDL-C. In NHWs, one SNP showed association with

HDL-C, and one with TC. One previously identified gain-of-function variant was present in three African Black subjects (Phe216Leu) and one of them had the lipid profile of FH.

In conclusion, this study successfully investigated the distribution pattern of *PCSK9* allele frequencies and their associations with plasma lipid profile in two previously uncharacterized populations. The relationship between Phe216Leu and abnormal lipid levels in one African Black subject appears to confirm its association with FH.

Table of Contents

List of Tables

List of Figures

Preface

I would first like to thank my essay advisor, Dr. M. Ilyas Kamboh for his invaluable guidance and patience in this essay project. It has been a true privilege to be a part of an incredible research team. I greatly appreciate the mentorship and support in the completion of this essay project. I would also like to thank all the members of the Kamboh lab for their support, assistance, and diligence to see this project through.

I also want to thank the following faculty and staff members for their support and help: Dr. Frank (Kang-Hsien) Fan, for his assistance with the statistical analysis and results interpretations, Dr. M. Muaaz Aslam, who patiently taught me how to properly analyze my data, and Dr. Allison L. Kuipers, for graciously taking the time to be a secondary reader for this essay. I sincerely appreciate you all for taking the time necessary to guide this project.

I would also like to thank the School of Public Health Faculty for their insight throughout my graduate courses. Specifically, I would like to thank Dr. Andrea Durst for her advising expertise and Dr. Cynthia Salter for radiating positivity and always making class an absolute joy. The knowledge I have gained will propel me to new heights and dreams.

Finally, I would like to thank my friends and family for the infinite support over the last two years. Without you, I would not have dared to dream that I could achieve anything.

1.0 Introduction

In the United States, cardiovascular disease (CVD) is the leading cause of death for men and women (Virani et al., 2021). A combination of genetic and environmental factors leads to an increased risk of CVD. Factors that affect the risk of CVD include diet, smoking, obesity, diabetes, hypertension, age, genetics, and plasma lipid levels, including low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C) and triglycerides (TG). HDL is the "good" cholesterol—normal levels are greater than 60 mg/dL. LDL is the "bad" cholesterol—normal levels are considered less than 100 mg/dL. Lower levels of TG are considered better—ideal levels are less than 150 mg/dL. Normal total cholesterol levels should not exceed 200 mg/dL (Expert Panel 2001).

The most common type of CVD is coronary artery disease (CAD)—about 6.72% (18.2 million people) of the United States population have CAD (Virani et al., 2021). Heart disease causes mortality for all ethnicities. In 2015, the percentage of all deaths caused by heart disease was 18.3% in American Indian/Alaska Natives, 21.4% in Asian American/Pacific Islanders, 23.5% in non-Hispanic Blacks, 23.7% in non-Hispanic Whites, and 20.3 % in Hispanics compared to all deaths (Heron, 2019). Men have a higher mortality rate from CAD than women in all ethnicities—24.4% versus 22.3%, respectively. The lifetime risk for developing CAD is 49% in men aged 40 and 32% in women aged 40 (Sanchis-Gomar et al., 2016). CAD occurs when increased plasma LDL-C levels are associated with atherosclerosis in the coronary arteries. Atherosclerosis and plaque build-up, commonly lead to CVD that presents as angina, myocardial infarctions, and strokes (Youngblom et al. 2016).

Dysfunction of endothelial cells of the coronary artery causes the enhanced entry and accumulation of LDL particles in the intima, which leads to the oxidation of LDL, eliciting an immune response. The endothelial cells respond by releasing chemokines and adhesion molecules, which attract monocytes to the intima. The monocytes mature into macrophages (immune cells) and ingest the oxidized LDL creating foam cells. The fatty streak formed by foam cells indicates the early stages of atherosclerosis (Rafieian-Kopaei et al. 2014).

LDL is the most cholesterol rich lipid in the body—normal lipid metabolism is critical for cholesterol clearance. Apolipoprotein B (ApoB) molecules form an envelope around LDL. This allows for LDL to attach to LDL-receptors on the surface of the liver. In the endosome, the LDL particle-receptor complex is broken down into small lipid components. The LDL-receptor gets recycled back to the surface of the liver or destroyed by lysosomes. The lipid particle degrades via lysosomal enzymes. Several genes involved in lipid metabolism directly influences plasma lipid levels. Any dysfunction in these genes can lead to an increased risk of CAD development (~20 fold) (Chemello et al. 2021, Youngblom et al. 2016).

Familial hypercholesterolemia (FH) is a genetic disorder that often leads to earlier onset of CAD. Signs of FH include high levels of LDL-C (>190 mg/dL) and total cholesterol (> 310 mg/dL), family history of heart disease or heart attacks, swollen ankles, and xanthomas (Virani et al. 2021). The average age of onset for FH is 50 years of age (Youngblom et al. 2016). Mutations in five genes are associated with FH, of which three show autosomal dominant inheritance (*LDLR*, *PCSK9*, and *APOB*) and two show autosomal recessive inheritance (*LDLRAPI* and *ABCG5/ABCG8*). About 60-80% of FH cases are due to autosomal dominant mutations. *PCKS9* mutations are unique and rare—only 2-4% of FH cases are associated with the *PCSK9* mutations (Pham et al. 2021). Gain-of-function *PCSK9* mutations result in the destruction LDL receptors

resulting in FH. Loss-of-function *PSCK9* mutations increase the number of LDL receptors, resulting in the clearance of more LDL particles from the bloodstream and, thus, lower levels of circulating cholesterol. These loss-of-function mutations are often nonsense mutations with a prevalence of about 2.6% in African populations, and about three percent in European populations (Khoury et al. 2017). Results from *PCSK9* research aided in the creation of *PCSK9* inhibitors as a new therapy for FH (Rosenson et al. 2018). However, the general public is not able to obtain ideal cholesterol levels on standard lipid lowering medication alone (Sabatine et al. 2017; Schwartz et al. 2018).

1.1 Proprotein Convertase Subtilisin/Kexin Type 9 (PCSK9)

The *PCSK9* gene is located on chromosome 1p32 and consists of 12 exons with high expression in the liver, small intestine, and kidney (Abifadel et al. 2003). The PCSK9 protein consists of 692 amino acids containing a signal peptide, prodomain, catalytic domain, and C-terminal region (**Figure 1**). Initially this was named as neural apoptosis-regulated convertase type 1 (NARC1). In the endoplasmic reticulum (ER), PCSK9 is autocatalytically cleaved after synthetization as a proenzyme (inactive). PCSK9 is then ready for secretion and transportation. PCSK9 is responsible for regulating the degradation of the LDL receptor—secreted PCSK9 binds to the epidermal growth factor A (EGF-A) of the LDL receptor and prevents the receptor from being recycled. The complex is degraded by endosomes or lysosomal compartments. Dysfunction of PCSK9 results in improper receptor recycling and reduction of LDL receptors on the surface of cells (El Khoury et al. 2017). Several functional *PCSK9* allelic variants have been identified to be associated with either FH due to high LDL-C or protection against CAD due to low LDL-C levels.

Table 1 depicts the seven known protein coding variants —four associated with low LDL-C levels and three with FH.

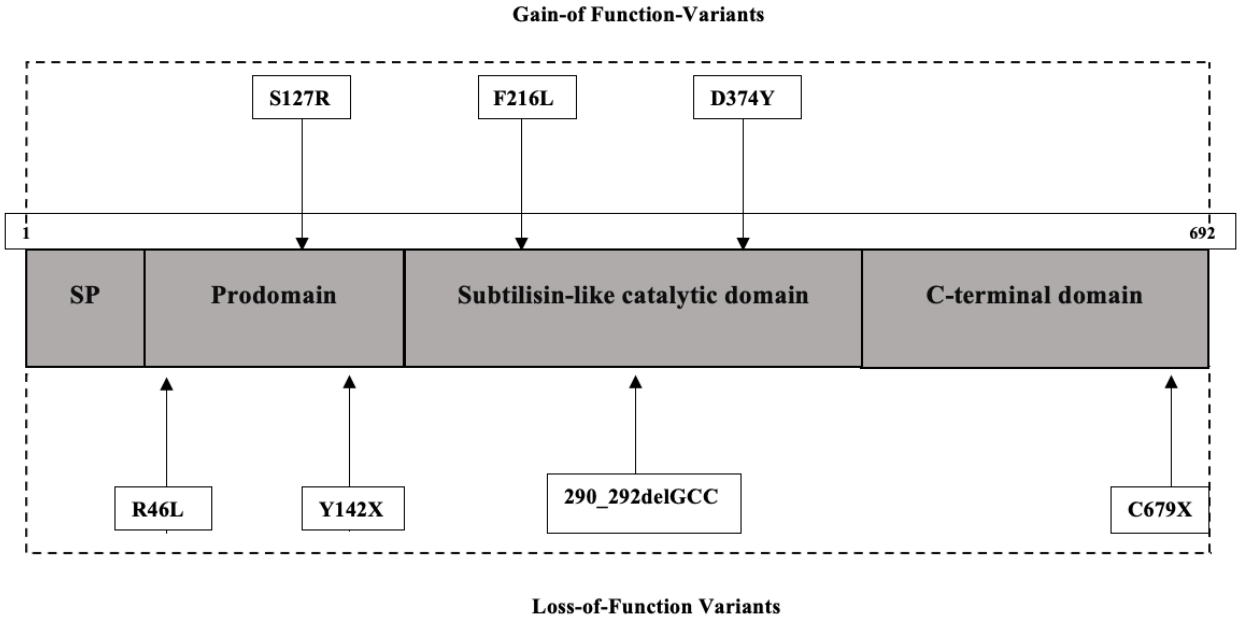

Figure 1. Structure of *PCSK9* and Coding Variant Locations

-The structure of *PCSK9* with selected variant locations. SP, signal peptide; R46L, Arg46Leu; Y142X, Tyr142Ter; 290_292delGCC, 3-BP DEL 290GCC; C679X, Cys679Ter; S127R, Ser127Arg; F216L, Phe216Leu; D374Y, Asp374Tyr
(Adopted and redrawn from Maxwell and Breslow 2005; Soutar and Naomova 2007; Guo et al. 2020)

Table 1. Summary of Known *PCSK9* Coding Variants

Phenotype	African MAF	U.S. Whites MAF	Mutation	Ref SNP ID #	Chromosome 1 position: base pair (BP)(GRCh37)
FH	0.0000	0.0000	Ser127Arg	rs28942111	1:55509689
FH	3.0E-5	1.0E-5	Phe216Leu	rs28942112	1:55518073
FH	0.0001	0.0000	Asp374Tyr	rs137852912	1:55523127
Low LDL-C	0.00298	0.00017	Tyr142Ter	rs67608943	1:55512222
Low LDL-C	0.00838	0.00023	Cys679Ter	rs28362286	1:55529215
Low LDL-C	0.0041	0.02222	Arg46Leu	rs11591147	1:55505647
Low LDL-C	NA	NA	3bp Del 290GCC	rs587776545	1:55509598

-modeled after OMIM (#607786). FH: Familial Hypercholesterolemia; LDL-C: Low-density lipoprotein cholesterol; MAF: minor allele frequency. MAFs taken from dbSNP (NCBI 2022).

1.1.1 Gain-of-Function Mutations Associated with Familial Hypercholesterolemia (FH)

Gain-of-function mutations were the first variants discovered in *PCSK9*. Abifadel et al. (2003) sequenced the *PCSK*9 gene in a French family with autosomal dominant hypercholesterolemia and identified an amino acid change at codon 127(Ser127Arg). This mutation creates an RGD attachment site (Arg-Gly-Asp site) that is thought to interact with the integrin binding in the mechanism (Abifadel et al. 2003). 19 members of the family were sampled and analyzed by haplotype. The same haplotype was observed in affected family members, between D1S722 and D1S2890 markers on chromosome 1. 41 genes are encoded in this region—upon sequencing, *PCSK9* was identified as the causal gene. The same mutation was later confirmed in another affected French family (Abifadel et al. 2003).

Abifadel et al. (2003) also identified a second mutation in a third French family—Phe216Leu. The proband family member died of a myocardial infarction at age 49. Upon sequence analysis, the proband showed a different amino acid substitution in exon 4. Kwon et al. (2008) investigated the structure of the *PCSK9* gene to understand the molecular mechanisms of mutations and confirmed that Phe216Leu is located in a disordered loop. Although Phe216Leu does not interact with the EGF-A complex structure, the amino acid substitution destroys a salt bridge downstream of a cleavage event that likely reduces the interaction resulting in the dysfunction of LDL-receptor regulation (Kwon et al. 2008).

Sun et al. (2005) identified the Asp374Tyr mutation in three families of European ancestry. Twelve of the affected family members had severe hypercholesterolemia compared to other heterozygous mutations and suggested that the Asp374Tyr mutation is associated with increased secretion of ApoB containing lipoproteins. This likely explained the increase levels of LDL-C in this mutation carriers. The replacement of Asp by Tyr places a hydroxyl group from His306 on

EGF-A, creating a hydrogen bond that increases PCSK9 activity leading to the enhanced degradation of LDL-receptors (Kwon et al. 2008).

1.1.2 Loss-of-Function Mutations Associated with Low LDL-Cholesterol

Cohen et al. (2006) conducted a study investigating the effects of *PCSK9* loss-of-function mutations on plasma lipid levels. They sequenced 128 subjects with low levels of LDL-C, including 50% US Blacks, and identified two nonsense mutations. The Tyr142Ter mutation was identified in 3 of the 64 Black subjects. This mutation is located in exon 3 of the gene resulting in the predicted deletion of the last four-fifths of the protein. This mutation was also found to be protective against CAD in US Black participants (Cohen et al. 2006).

In 4 of the 64 US Black subjects, the second nonsense mutation identified was Cys679Ter that resulted in the cleavage of 14 amino acids. Altogether, about 11% of the 64 Black subjects (7/64) with low LDL-C levels had these two nonsense mutations. Genotyping of these two mutations in >2600 US Blacks showed that about 1.8% or one in every 50 African Americans have a *PCSK9* nonsense mutation (Cohen et al. 2005). Genotyping of 549 Yoruba-speaking Nigerians found no example of the Tyr142Ter mutation, but the Cys679Ter mutation was present at 1.4% frequency and is similar to the 1.4% frequency found in US Blacks (Cohen et al. 2005). These data suggest that low LDL-C and *PCSK9* mutations are more likely to occur in those with African ancestry.

A compound heterozygous individual carrying a nonsense Tyr142Ter mutation and a 3bp deletion (290GCC) was identified who had no circulating PCSK9 and LDL-C level was only 14 mg/dL (Zhao et al. 2006). The predicted mechanism causes mRNA decay that results in the deletion in the prodomain disrupts the process of PCSK9 synthesis (Park et al. 2004). Another

PCSK9 mutation, Arg46Leu, with a frequency of 3.2% in US Whites and only 0.6% in US Blacks has been identified, which is associated with a 15% reduction in plasma LDL-C and 9% reduction total cholesterol (Cohen et al. 2006). Zhao et al observed significant low LDL-C levels with the Arg46Leu mutation as well. The substitution of arginine (polar) with leucine (hydrophobic) is predicted to interfere with the folding of the protein between the prodomain and C-terminal. This likely reduces the rate of PCSK9 processing (Park et al. 2004).

These data show that *PCSK9* is important in the determination of plasma LDL-C levels, which eventually led to the development of PCSK9 inhibitors as drug therapy for both FH hypercholesterolemia and polygenetic hypercholesterolemia in the general population (Rosenson et al. 2018). *PCSK9* studies can hold vital information for therapeutic studies. *PCSK9* variants are still not widely known in many understudied populations including African Blacks and NHWs. This study aims to identify significant *PCSK9* variants for further investigations in these populations.

1.2 Study Aims

The objective of this study is to confirm the relationship between the *PCSK9* variants and plasma lipid levels in two population-based samples that have not been previously studied with respect to the genetic effect of *PCSK9* on plasma lipid profile. One sample is composed of 788 African Blacks from Nigeria and the second is composed of 623 non-Hispanic Whites (NHWs) from the United States. *PCSK9* genotyping was completed using genome-wide chip followed by genotype-phenotype association analyses. This study will add to the field to further precision health research for those of African and European descents.

Aim 1: Utilize the genotype data on all *PCSK9 v*ariants from a new genome-wide data on African Blacks and NHWs and determine the allele frequencies on individual variants on chromosome 1.

Aim 2: Analyze the allele frequency differences of each *PCSK9* variants between the African Blacks and NHWs.

Aim 3: Perform association analyses on the effect of *PCSK9* variant on plasma lipid profile within each racial group and then contrast the results between racial groups.

2.0 Subjects and Methods

2.1 Background on the African Black Population and Non-Hispanic White Population

Populations differ in habits, genetics, and environment. Western societies play a large role in these factors. In general, the Nigerian population has lower risk for CAD due to low body mass index (BMI), low cholesterol intake from diet, and higher physical activity levels. Although, cholesterol levels are higher in African Americans, the West African population historically has lower cholesterol levels (Bunker et al. 1996). An early epidemiology study on healthy African Blacks found that they have favorable lipid profile and a lower prevalence of ischemic heart disease than in healthy Europeans (Onitiri et al. 1977). The same study also reported that the consumption of higher fat and protein diet was associated with higher lipid levels in African Blacks. Since there is a positive relationship between heart disease and age—it is possible that the difference in lipid levels and heart disease could be explained by age as a factor. Westernized populations have a higher life expectancy than African populations, allowing Westernized populations to have an increased risk of heart disease and higher lipid levels on average. However, age was controlled as a confounding factor in this study and does not impact the low heart disease prevalence finding in African Blacks. In **Table 2,** the average lipid profiles are "healthier" in our African Black sample compared to our NHW sample.

2.1.1 Study Samples

The Nigerian African Black sample (BENIN) is from a previous CHD-related risk factors study in Benin City, Nigeria. DNAs from 788 subjects, aged 19 to 70, was collected from blood clots (Bunker et al. 1996). The NHW samples, aged 24 to 75, are derived from the San Luis Valley Diabetes Study (SLVDS) in Colorado. All subjects were non-diabetic, and DNA was extracted from buffy coats (Rewers et al. 1996). African Black lipid levels were determined from fasting blood samples with a serum assay, and NHW lipid levels were determined from fasting plasma samples (Bunker et al. 1996; Rewers et al. 1996). **Table 2** provides a summary of phenotypes and demographics for both population samples and their detailed description is given elsewhere (Bryant et al. 2013; Pirim et al. 2019; Pirim et al. 2020). This study was approved by the University of Pittsburgh IRB.

Table 2. Demographics and Lipid Level Summary

Variable	NHWs (n=621) Mean±sd or count (percentage)	African Blacks (n=788) Mean±sd or count (percentage)	P-value
Males	294 (47.35%)	495 (62.81%)	<0.0001
Age (Yrs)	52.83 ± 11.41	40.95 ± 8.39	<0.0001
BMI (kg/m^2)	25.5 1± 4.06	22.87 ± 4.04	<0.0001
LDL-C (mg/dl)	136.9 ± 40.80	109.25 ± 34.40	<0.0001
HDL-C (mg/dl)	50.76 ± 14.35	47.88 ± 12.87	0.0001
TG (mg/dl)	142.72 ± 93.49	72.96 ± 39.32	<0.0001
TC (mg/dl)	217.0 ± 43.5	172.01 ± 38.47	<0.0001

TC: Total cholesterol; LDL-C: Low-density lipoprotein cholesterol; HDL-C; High-density lipoprotein cholesterol; TG: Triglycerides

- Adapted from (Bryant et al. 2013; Pirim et al 2019; Pirim et al 2020). Mean is unadjusted.

2.2 Genotyping Methods

Both sets of samples underwent genotyping using the Infinium™ Global Diversity Array-8 v.1.0 BeadChip. This chip uses Genome Reference Consortium Human Build 37 (GRCh37) to identify variants.

This array utilizes 1,825,277 markers. This chip provides broad coverage for diverse multiethnic populations. The Infinium™ Global Diversity Array-8 v1.0 tags both common and low-frequency variants selected from a wide variety of scientific databases, ancestries, and diseases. Quality control markers (QC) markers are used for sample tracking, ancestry determination, and sex determination (Ilumina 2021).

Since this study is focused on *PCSK9*, SNPs shown in **Table 3** are located near or in the *PCSK9* gene starting from 55504148 bp to 55531526 bp (including 1000bp upstream and downstream of *PCSK9* region). SNPs were annotated by using SNPNexus, BiomaRT, VEP, UCSC, Bioconductor, and Annovar. Of the 93 SNPs identified in the Infinium™ Global Diversity Array-8 v.1.0, in the *PCSK9* region, 90 had reference IDs (RS numbers) and 60 were polymorphic in African Blacks or NHWs. Two of the seven SNPs listed in **Table 1.** were present on the chip. The Phe216Leu mutation was found in African Blacks only, while the Ser127Arg mutation was not identified in either sample. Both of these mutations are highlighted in yellow in **Table 3**.

Table 3. 93 *PCSK9* SNPs present on the Infinium™ Global Diversity Array-8 v.1.0 BeadChip

Ref SNP ID	Position (GRCh37)	Alleles (Major:Minor)	Samples with Minor Alleles present
rs186669805	1:55505520	G: A	ABs
rs376819769	1:55505585	C: T	monomorphic
rs1553135400	1:55505586	GC: G	monomorphic
rs773660398	1:55505640	G: A	monomorphic
rs1372204035	1:55505650	C: G	monomorphic
rs28385701	1:55505651	C: T	ABs, NHWs
rs1278890129	1:55505652	G: A	monomorphic
rs867904088	1:55505698	C: A	ABs
rs1272703401	1:55505722	G: A	ABs
rs115219247	1:55509009	G: A	ABs
rs151095149	1:55509522	T: G	ABs
no known variants (nearest: 55509560)	1:55509556	A: C	monomorphic
rs373551845	1:55509575	G: A	ABs
rs147865087	1:55509584	G: A	monomorphic
rs376385276	1:55509598	G: A	ABs, NHWs
rs369067856	1:55509618	C: T	monomorphic

rs79805678	1:55509644	G: A	monomorphic
rs28942111 (Ser127Arg)	1:55509689	T: A	monomorphic
rs72646503	1:55509692	C: T	NHWs
rs1553135971	1:55509694	A: G	monomorphic
rs28362222	1:55511327	C: T	ABs
rs372600893	1:55512266	G: A	monomorphic
rs367620267	1:55512275	G: A	monomorphic
rs141593516	1:55512289	C: T	monomorphic
rs137878146	1:55512295	C: T	monomorphic
rs373018373	1:55512318	C: T	ABs
rs28385708	1:55513043	C: T	ABs, NHWs
rs41294821	1:55513183	C: T	ABs, NHWs
rs45576433	1:55517301	A: G	NHWs
rs200856421	1:55518036	C: T	monomorphic
rs28942112 (Phe216Leu)	1:55518073	T: C	ABs
rs7552350	1:55518198	C: A	ABs, NHWs
rs371336612	1:55518321	A: C	ABs
rs150169598	1:55518362	G: T	ABs
rs148195424	1:55518374	C: T	ABs
rs370145766	1:55518382	C: T	ABs, NHWs
rs41297883	1:55518385	C: T	ABs, NHWs

rs147478188	1:55518386	G: A	ABs, NHWs
rs376945520	1:55518417	G: A	ABs
rs201789841	1:55518456	C: T	monomorphic
rs865914100	1:55521806	G: A	ABs, NHWs
rs867916076	1:55521846	C: A	monomorphic
rs376753957	1:55521859	C: T	ABs, NHWs
rs142149050	1:55522869	C: T	ABs
rs509504	1:55523033	G: A	ABs
no known variants (nearest: 55523061)	1:55523068	G: A	ABs
rs754420902	1:55523181	G (del)	ABs
rs368257906	1:55523187	G: A	ABs
rs200091654	1:55523734	C: T	monomorphic
rs72646515	1:55523822	G: A	ABs
rs540796	1:55524197	G: A	ABs, NHWs
rs1553137693	1:55524219	A: G	monomorphic
rs1553137699	1:55524228	G: T	monomorphic
rs376388695	1:55524244	G: A	ABs, NHWs
rs28362268	1:55524248	C: T	ABs
rs375582388	1:55524249	G: A	ABs
rs142229832	1:55524281	T: C	monomorphic
rs139669564	1:55524304	G: A	ABs
rs147599496	1:55524308	C: T	ABs

rs201395805	1:55524312	C: T	ABs
rs143394031	1:55524313	G: C	ABs
rs28362269	1:55525143	C: T	ABs
rs866944660	1:55525186	A: C	monomorphic
rs72646525	1:55527093	C: T	ABs
rs367606156	1:55527158	G: A	ABs
rs755750316	1:55527213	C: A	ABs, NHWs
rs145264023	1:55527214	G: A	ABs
rs142118418	1:55527221	C: T	ABs
rs75266432	1:55528377	G: A	monomorphic
rs28362285	1:55529047	C: T	ABs, NHWs
rs199815786	1:55529056	C: T	monomorphic
rs145770391	1:55529107	C: T	ABs
rs138178437	1:55529122	C: T	monomorphic
rs201280059	1:55529132	A: G	ABs
rs139894975	1:55529137	G: A	monomorphic
rs745864657	**1:55529151-55529158**	GCCGGGAC (del)	monomorphic
rs147182054	1:55529153	C: G	monomorphic
rs780214893	1:55529154	G: T	monomorphic
rs371914056	1:55529158	C: T	ABs
rs369851423	1:55529169	C: T	ABs
rs142524469	1:55529206	C: T	ABs

rs533555352	1:55529216	C: T	ABs, NHWs
rs189293781	1:55529285	G: A	ABs, NHWs
rs28362287	1:55529332	C: T	ABs, NHWs
rs72646536	1:55530105	T: C	ABs
rs41294827	1:55531528	G: T	ABs, NHWs
rs138436072	1:55535688	G: A	ABs
rs142116310	1:55536386	G: A	NHWs
rs17410238	1:55537966	G: A	ABs, NHWs
rs756133834	1:55538497	G: A	monomorphic
rs866351596	1:55539543	C: A	monomorphic
rs770231986	1:55539564	C: A	NHWs
rs867610842	1:55539580	C: A	monomorphic
no known variants (nearest: 55539533)	1:55542933	G: A	monomorphic

- ABs: African Blacks; NHWs: non-Hispanic Whites. 60 SNPS were polymorphic in ABs or NHWs. The remaining 33 were monomorphic.

2.3 Statistical Methods

Call rates were calculated in Microsoft Excel. The calculation used the total number of identified genotyped samples divided by the total number of samples. Minor allele frequencies (MAF) were calculated using the minor allele count divided by the total number of alleles (total number genotyped for that SNP times two. Hardy-Weinberg Equilibrium (HWE) calculations were performed using a gene calculator online (gene-calc) to determine concordance. Each variant was tested using a χ^2 (chi-squared) test. If samples were smaller than five, then the p-value underwent correction testing with Yate's correction for continuity (Miks and Binkowski 2022). A p-value of 8.0E-4 is considered significant due to our sample size being 60 (0.05 divided by 60).

In order to normalize the distribution of lipid levels in the samples, the Box-Cox transformation was used for HDL-C and TG levels in NHWs and HDL-C, TG levels, and LDL-C levels in African Blacks. Lipid level and genotype associations were conducted using the R statistical software (version 3.6.1). Linear regression analysis, under the additive model, was completed for both population samples and adjusted for significant covariates. For African Blacks, the covariates used were sex, age, waist, jobmin (exercise minutes to work), and staff level (junior level or senior level). For NHWs, the covariates used were sex, age, BMI, and smoking. Covariates were determined from previous studies (Bunker et al. 1996; Rewers et al. 1996; Pirim et al 2019; Pirim et al 2020). P-values less than 0.05 are considered as evidence for association and confidence intervals were measured at 95 percent.

Linkage disequilibrium (LD) was also investigated between the 60 SNPs. The LDLink LDmatrix tool was used to investigate the R^2 values for African populations and European populations (Machiela and Chanock 2015). The LDProxy tool was also used to identify potential proxy SNPs for the known *PCSK9* variants mentioned in **Table 1.**

3.0 Results

3.1 Allele Frequencies, Genotyping Call Rates, and HWE

Of the 93 SNPs genotyped, 60 showed variations in one or the other group or in both. 19 SNPs were present in both, 37 SNPs were present in only African Blacks, and 4 SNPs were present in only NHWs (**Table 4**). MAFs ranged from 0.0001 to 0.2160. MAFs greater than five percent are bolded. MAFs between populations that are different are highlighted in yellow. For **rs28385708,** the African Black MAF was 0.0129, and the NHWs MAF was 0.540. For **rs45576433,** the African Black MAF was 0.0102, and the NHWs MAF was 0.1393. For **rs7552350,** the African Black MAF was 0.2160, and the NHWs MAF was 0.1625. For **rs28362287,** the African Black MAF was 0.0725, and the NHWs MAF was 0.011. For **rs17410238**, the African Black MAF was 0.0142, and the NHWs MAF was 0.1357. In **rs7552350** and **rs28362287,** the MAF was higher in African Blacks than in NHWs.

Call rates were high among both samples—the lowest call rate was 94.47% in the African Blacks. All other call rates were greater than 95%.

Two of the 60 SNPs had HWE p-values less than 8.0E-04 (rs373018373 and rs367606156 in ABs). After correction testing, rs373018373 is no longer significant, but rs367606156 is still significant and thus was excluded from the lipid association analyses.

Table 4. Allele Frequencies, Call Rates, and HWE in African Blacks and NHWs

Ref SNP ID	Alleles (Major: Minor)	Position	ABs MAF	ABs Call Rate	ABs HWE P-value	NHWs MAF	NHWs Call rate	NHWs HWE P-value
rs186669805	G: A	1:55505520	0.0014	0.9811	0.9993	NA	1.000	NA
rs28385701	C: T	1:55505651	0.0014	0.9879	0.9993	0.0079	0.9896	0.9821
rs867904088	C: A	1:55505698	0.0021	0.9852	0.9985	NA	0.9965	NA
rs1272703401	G: A	1:55505722	0.0034	0.9933	0.9957	NA	0.9983	NA
rs115219247	G: A	1:55509009	0.0101	1.000	0.9620	NA	1.000	NA
rs151095149	T: G	1:55509522	0.0001	0.9730	0.9998	NA	0.9774	NA
rs373551845	G: A	1:55509575	0.0014	0.9825	0.9993	NA	1.000	NA
rs376385276	G: A	1:55509598	0.0041	0.9798	0.9938	0.0017	1.000	0.9991
rs72646503	C: T	1:55509692	NA	0.9879	NA	0.0009	0.9983	0.9998
rs28362222	C: T	1:55511327	0.0075	0.9865	0.9792	NA	1.000	NA
rs373018373	C: T	1:55512318	0.0069	0.9811	<0.0001	NA	0.9983	NA
rs28385708	C: T	1:55513043	0.0129	0.9960	0.9392	**0.0540**	0.9965	0.8420
rs41294821	C: T	1:55513183	0.0231	0.9933	0.8141	0.0297	0.9861	0.8060
rs45576433	A: G	1:55517301	0.0102	0.9933	0.9618	**0.1393**	0.9861	0.2399
rs28942112 (Phe216Leu)	T: C	1:55518073	0.0021	0.9784	0.9984	NA	0.9809	NA
rs7552350	C: A	1:55518198	**0.2160**	0.9717	0.9804	**0.1625**	0.9843	0.0471
rs371336612	A: C	1:55518321	0.0021	0.9771	0.9984	NA	0.9809	NA
rs150169598	G: T	1:55518362	0.0049	0.9717	0.9915	NA	0.9948	NA

rs148195424	C: T	1:55518374	0.0007	0.9757	0.9998	NA	0.9896	NA
rs370145766	C: T	1:55518382	0.0097	0.9703	0.9659	0.0009	0.9983	0.9998
rs41297883	C: T	1:55518385	0.0146	0.9676	0.9239	0.0052	0.9983	NA
rs147478188	G: A	1:55518386	0.0041	0.9771	0.9938	0.0009	1.000	0.9998
rs376945520	G: A	1:55518417	0.0027	0.9838	0.9973	NA	0.9948	NA
rs865914100	G: A	1:55521806	0.0007	0.9649	0.9998	0.0009	0.9948	0.9998
rs376753957	C: T	1:55521859	0.0100	0.9447	0.9649	0.0017	0.9965	0.9991
rs142149050	C: T	1:55522869	0.0191	0.9892	0.8703	NA	1.000	NA
rs509504	G: A	1:55523033	**0.0719**	0.9757	0.2994	NA	0.9983	NA
no known variants (nearest: 55523061)	G: A	1:55523068	0.0014	0.9663	0.9993	NA	0.9757	NA
rs754420902	G (del)	1:55523181	0.0028	0.9703	0.9972	NA	0.9983	NA
rs368257906	G: A	1:55523187	0.0069	0.9798	0.9827	NA	0.9965	NA
rs72646515	G: A	1:55523822	0.0056	0.9690	0.9888	NA	1.000	NA
rs540796	G: A	1:55524197	**0.1756**	0.9838	0.9968	**0.1522**	1.000	0.4478
rs376388695	G: A	1:55524244	0.0041	0.9771	0.9938	0.0009	0.9965	0.9998
rs28362268	C: T	1:55524248	0.0220	0.9852	0.8325	NA	1.000	NA
rs375582388	G: A	1:55524249	0.0055	0.9825	0.9890	NA	0.9965	NA
rs139669564	G: A	1:55524304	0.0020	0.9879	0.9985	NA	0.9983	NA
rs147599496	C: T	1:55524308	0.0035	0.9730	0.9967	NA	0.9913	NA
rs201395805	C: T	1:55524312	0.0014	0.9758	0.9993	NA	0.9878	NA

rs143394031	G: C	1:55524313	0.0055	0.9906	0.9890	0.0009	1.000	0.9998
rs28362269	C: T	1:55525143	**0.0940**	0.9906	0.4800	NA	0.9965	NA
rs72646525	C: T	1:55527093	0.0055	0.9730	0.9889	NA	0.9948	NA
rs367606156	G: A	1:55527158	0.0062	0.9730	<0.0001	NA	0.9965	NA
rs755750316	C: A	1:55527213	0.0041	0.9784	0.9938	NA	0.9948	NA
rs145264023	G: A	1:55527214	0.0034	0.9798	0.9957	NA	1.000	NA
rs142118418	C: T	1:55527221	0.0014	0.9798	0.9993	NA	0.9983	NA
rs28362285	C: T	1:55529047	0.0191	0.9879	0.8701	.0009	0.9983	0.9998
rs145770391	C: T	1:55529107	0.0014	0.9946	0.9993	NA	1.000	NA
rs201280059	A: G	1:55529132	0.0041	0.9798	0.9938	NA	0.9826	NA
rs371914056	C: T	1:55529158	0.0048	0.9865	0.9916	NA	0.9983	NA
rs369851423	C: T	1:55529169	0.0007	0.9622	0.9998	NA	1.000	NA
rs142524469	C: T	1:55529206	0.0041	0.9798	0.9938	NA	0.9983	NA
rs533555352	C: T	1:55529216	0.0090	0.9717	0.9706	0.0009	0.9948	0.9998
rs189293781	G: A	1:55529285	0.0021	0.9744	0.9984	0.0009	0.9983	0.9998
rs28362287	C: T	1:55529332	**0.0725**	0.9865	0.1072	0.0011	0.9965	0.9630
rs72646536	T: C	1:55530105	0.0069	0.9771	0.9827	NA	0.9774	NA
rs41294827	G: T	1:55531528	0.0413	0.9973	0.8390	0.0314	0.9965	0.7398
rs142116310	G: A	1:55536386	NA	0.9956	NA	.01304	1.000	0.9510
rs138436072	G: A	1:55535688	0.0150	0.9865	0.9192	NA	1.000	NA
rs17410238	G: A	1:55537966	0.0142	1.000	0.9263	**0.1357**	1.000	0.4289
rs770231986	C: A	1:55539564	NA	1.000	NA	0.0017	0.9983	0.9991

*MAF is minor allele frequency. **Bold** indicates MAFs ≥ 5% . Yellow highlight indicates vastly different MAFs between populations. Green highlight indicates <.0008 p-value for HWE. $\alpha = .05$. Red indicates SNP from table 1.

3.2 Lipid Association Analyses

Tables 5 and 6 show the lipid associations in African Blacks and NHWs, respectively. Only the significant genotype associations (p-values < 0.05) are shown. All association results can be found in Appendix A (African Blacks) and Appendix B (NHWs). Seven SNPs showed significant associations with at least one lipid trait in African Blacks (56 tested SNPs) and two additional SNPs showed significant associations in NHWs (23 tested SNPs). For NHWs, TG and LDL-C p-values were not significant in any of the genotyped SNPs.

3.2.1 Significant Lipid Associations for African Blacks

In African Blacks, two SNPs showed significance with total cholesterol (TC), two with TG, two with HDL-C, and four with LDL-C (**Table 5**). Of the seven significant SNPs, none showed significant associations with all four lipid traits. Rs28385701 showed significance with TC (p = 0.00269) and LDL-C (0.00172). Rs867904088 showed nominal significance with LDL-C (p = 0.0494). Rs1272703401 showed significance with TC (p = 0.0262) and LDL-C (p = 0.0245). Both rs151095149 (p = 0.0264) and rs41297883 (p = 0.000882) showed significance with TG. Rrs41294827 showed significance with HDL-C (p = 0.0238) and LDL-C (p = 0.0332). Lastly, rs138436072 showed nominal significance with HDL-C (p = 0.045).

3.2.2 Significant Lipid Associations for NHWs

One SNP showed significant association with HDL-C (rs72646503; p = 0.024), and one with TC (rs7552350; p = 0.0381).

Table 5. Significant *PCSK9* Variant Associations with Lipids in African Blacks

	TC*		TG*		HDL-C*		LDL-C*	
Ref SNP ID	Beta	P-Value	Beta	P-Value	Beta	P-Value	Beta	P-Value
rs28385701	-4.201	**2.69E-03**	-0.097	3.25E-01	-0.446	8.62E-01	-8.388	**1.72E-03**
rs867904088	-1.691	1.35E-01	-0.028	7.28E-01	1.786	3.89E-01	-4.249	**4.94E-02**
rs1272703401	-2.004	**2.26E-02**	-0.094	1.29E-01	-0.079	9.61E-01	-3.763	**2.45E-02**
rs151095149	1.034	5.96E-01	0.305	**2.64E-02**	-3.586	3.18E-01	1.55	6.77E-01
rs41297883	-0.106	8.07E-01	0.102	**8.82E-04**	-0.674	3.95E-01	-0.423	6.11E-01
rs41294827	-0.387	1.25E-01	-0.033	6.80E-02	1.045	**2.38E-02**	-1.027	**3.32E-02**
rs138436072	-0.392	3.53E-01	0.019	5.25E-01	-1.559	**4.50E-02**	-0.348	6.65E-01

*Indicates corrected value with Box-Cox transformation for trait. **Bolded** refers to significant p-values ($p < 0.05$).

Table 6. Significant *PCSK9* Variant Associations with Lipids in NHWs

	TC		TG*		HDL-C*		LDL-C	
Ref SNP ID	Beta	P-Value	Beta	P-Value	Beta	P-Value	Beta	P-Value
rs72646503	-5.453	8.97E-01	0.222	4.84E-01	-0.417	**2.40E-02**	24.617	5.43E-01
rs7552350	-7.319	**3.81E-02**	-0.037	1.61E-01	-0.003	8.57E-01	-4.586	1.78E-01

*Indicates corrected value with Box-Cox transformation for lipid trait. LDL-C and TC were not included as there were no significant associations. Full analysis data can be found in Appendix B for NHWs. **Bolded** refers to significant p-values ($p < 0.05$).

3.2.3 Comparison of Significant SNPs in African Blacks and NHWs

Significant polymorphic SNPs from the previous result tables in both populations are summarized in **Table 7**. The table shows a comparison for associations for each population sample side by side in the context of the MAF and lipid level associations. Red SNPs indicate significant associations in either African Blacks or NHWs. Blue SNPs indicate known SNPs mentioned in **Table 1**. Only Phe216Leu (rs28942112) was detected in African Blacks.

For **rs28385701,** the MAFs are similar in both populations, and this SNP has the most significant p-values for LDL-C and TC in African Blacks. While beta effects are of similar direction and magnitude in NHWs, they were not statistically significant. For **rs7552350,** the MAFs are different in both populations, with the MAF being five percent higher in African Blacks. However, the significant association of this SNP is only seen in NHWs for TC. The beta estimate for TC in African Blacks is much smaller and not significant ($p = 0.343$). For **rs41297883**, the MAF is higher in African Blacks, and the significant association is between African Blacks and TG. This is the largest p-value ($p = 0.000882$) of all the significant associations. The beta effects are of similar direction for NHWs, but the magnitude is larger in TC and LDL-C. However, they are not statistically significant. For **rs41294827,** MAFs are similar between both populations, but significant associations are with HDL-C and LDL-C in African Blacks. While beta effects are of similar direction and magnitude in NHWs, they were not statistically significant. The African Black MAF is higher than the NHWs for **rs41294827.**

Only one rare mutation referenced in **Table 1** was detected in African Blacks—**rs28942112** (Phe216Leu) with a MAF of 0.21%. Only three samples carry Phe216Leu mutation, and their lipid traits, age, BMI, and sex are shown in **Table 8**. Abnormal lipid levels compared to the average values given in Table 2 for African Blacks are highlighted in yellow. The only female

in the table appears to have the FH heterozygous phenotype by having very high levels of TC and LDL-C. On the other hand, one male aged 50 has the opposite trend by having lower than the average TC and LDL-C.

Table 7. Summary of Polymorphic SNPs with Lipid Traits and MAFs in African Blacks and NHWs

Ref SNP ID	**African Blacks**			**NHWs**		
Traits	**MAF**	**Association**		**MAF**	**Association**	
rs28385701	T-0.0014	**Beta**	**P-value**	T- 0.0079	**Beta**	**P-value**
TC		-4.201	0.00269		-11.68	0.410
TG		-0.097	0.325		-0.146	0.168
HDL-C		-0.446	0.862		0.032	0.611
LDL-C		-8.388	0.00172		-8.532	0.532
rs7552350	A- 0.216	**Beta**	**P-value**	A- 0.1625	**Beta**	**P-value**
TC		-0.115	0.343		-7.319	0.0381
TG		0.004	0.671		-0.037	0.161
HDL-C		-0.336	0.13		-0.003	0.857
LDL-C		-0.201	0.382		-4.586	0.178
rs41297883	T-0.0146	**Beta**	**P-value**	T-0.0052	**Beta**	**P-value**
TC		-0.106	0.807		-7.739	0.653
TG		0.102	0.000882		0.047	0.719
HDL-C		-0.674	0.395		-0.034	0.654

LDL-C		-0.423	0.611		-9.564	0.563
rs41294827	T-0.0413	**Beta**	**P-value**	T-0.0314	**Beta**	**P-value**
TC		-0.387	0.125		-0.877	0.904
TG		-0.033	0.068		-0.032	0.560
HDL-C		1.045	0.0238		-0.007	0.832
LDL-C		-1.027	0.0332		-1.907	0.785
rs28942112	C-0.0021	**Beta**	**P-value**	NA	**NA**	
TC		27.82	0.188			
TG		0.36	0.651			
HDL-C		0.513	0.806			
LDL-C		26.993	0.151			

-Red indicates significant SNPs in either ethnic population from Table 5 or Table 6. Blue indicates known SNPs mentioned in Chapter 1 detected in at least one ethnic population.
-TC indicates total cholesterol; TG indicates triglyceride levels; HDL-C indicates HDL cholesterol levels; LDL-C indicates LDL cholesterol levels.

Table 8. Lipid Profile in Three African Blacks with rs28942112 (Phe216Leu)

rs28942112 Allele (Major:Minor)	**SEX**	**AGE (Yrs)**	**BMI (kg/m^2)**	**CHOL (mg/dl)**	**TG (mg/dl)**	**HDL-C (mg/dl)**	**LDL-C (mg/dl)**
T:C	M	35	19.75	182	62	47	122.6
T:C	M	50	19.25	140	55	57.1	71.9
T:C	F	40	26.24	256	86	40.2	198.6

*Highlighted values indicate abnormal lipid levels. Identifiers have been removed from each sample.

4.0 Discussions and Conclusions

In this study we have examined the distribution of *PCSK9* variants and their associations with plasma lipid profile in African Blacks from Nigeria and a NHW population from Colorado. As expected, African Blacks showed more variation since there is more diversity in the genome of people with African descent than those of European descent (Campbell and Tishkoff 2008). We also observed more significant associations with lipid levels in African Blacks than NHWs.

Of the nine SNPs significantly associated with at least one lipid trait in at least one population, four were polymorphic only in African Blacks, one was polymorphic only in NHWs, and four were polymorphic in both African Black and NHWs. Upon further investigation of the four SNPs polymorphic only in the African Black population (rs867904088, rs1272703401, rs151095149, rs138436072), two were missense variants (rs867904088, rs151095149), and two were intronic variants (rs1272703401, rs138436072). Potential clinical significance was identified from ClinVar and dbSNP (NCBI 2022). For rs867904088 (Thr61Asn), the clinical significance is uncertain for FH. For rs151095149 (Trp72Gly), there is no reported clinical significance. For rs1272703401, there is no known amino acid consequence, but the reported clinical significance is pathogenic for FH. For rs138436072, there is no known amino acid consequence, and there is no reported clinical significance. The only polymorphic SNP in NHWs (rs72646503) is synonymous (Gly128Gly).

Four polymorphic SNPs in both populations were significant in either African Blacks or NHWs— rs28385701, rs41297883, rs7552350, and rs41297883 (**Table 7**). None of the SNPs were missense variants. According to ClinVar, two SNPs are synonymous (rs28385701, rs41297883), one SNP is an intronic variant (rs7552350), and one SNP has an unknown consequence

(rs41294827). For rs28385701 (Ser47Ser), this variant is likely benign for FH. For rs7552350, there is no known amino acid consequence, but this variant is reported as benign for FH. For rs41297883 (Gly240Gly), the minor allele reported in our samples was T, but A has been reported as another possible minor allele. Both result in a synonymous mutation, and they are reported as likely benign for FH. For rs41294827, there is no reported consequence and no reported clinical significance as of April 2021 (NCBI 2022). Of the SNPs reported in ClinVar, it is not surprising that none have confirmed significant interpretations because these are rare variants and we aimed to study population variation in lipid traits, not FH patients. The contribution of our data on these associations may help further research on rare *PCSK9* variants, especially in understudied ethnic groups.

Five of the seven *PCSK9* functional variants listed in **Table 1** were not present on the genotyping chip. Of the two FH variants that were present on the genotyping chip, only one (Phe216Leu) showed variation and was present in three African Blacks, including two males and one female (**Table 8**). Compared to average LDL-C of 109 mg/dl and TC of 217 mg/dl in ABs, the only female with this mutation had very high LDL-C (199 mg/dl) and TC (256 mg/dl, which appears to be the case of FH heterozygous. On the other hand, one male at age 50, had very low LDL-C (72 mg/dl) and slightly below average TC (140 mg/dl), which is unexpected finding. Perhaps the latter subject has some other LDL-C lowering variants that might have contributed to low LDL-C despite carrying the functional mutation. High cholesterol levels or “bad” cholesterol levels are not just caused by genetics, but a combination of environmental and genetic factors including BMI and diet.

A limitation of this study is that only two of seven previously implicated functional mutations with FH were present on the genotype chip, and we did not genotype the remaining five mutations. This makes it difficult to confirm the previous literature with our genotype data. Rare variants in complex diseases like heart disease are difficult to investigate (Auer and Lettre 2015). We attempted to use proxy variants for these five functional variants (Asp374Tyr, Tyr142Ter, Cys679Ter, Arg46Leu, 3bp Del 290GCC). We used the LDProxy tool to determine nearby SNPs that could be proxy for the five *PCSK9* variants not genotyped (Machiela and Chanock 2015). The LDProxy tool did not have many of our genotyped SNPs in the database, therefore, we were unable to determine if any SNPs were acceptable proxies. The same was also true for the *Ensembl!(c)* tool (Howe et al. 2021). A future study should impute the SNPs in the *PCSK9* region as this is currently the suggested technology for rare variant association studies (Auer and Lettre 2015). Imputation increases the genetic information; this is more reflective of the total genetic variation in the *PCSK9* region. Similarly, linkage disequilibrium (LD) analysis may help to further investigate the 60 SNPs genotyped. However, an LD heat plot made from this database (LDLink) did not show any significant LD between SNPs when looking at the R^2 values in red shading **(Figure 2.)**

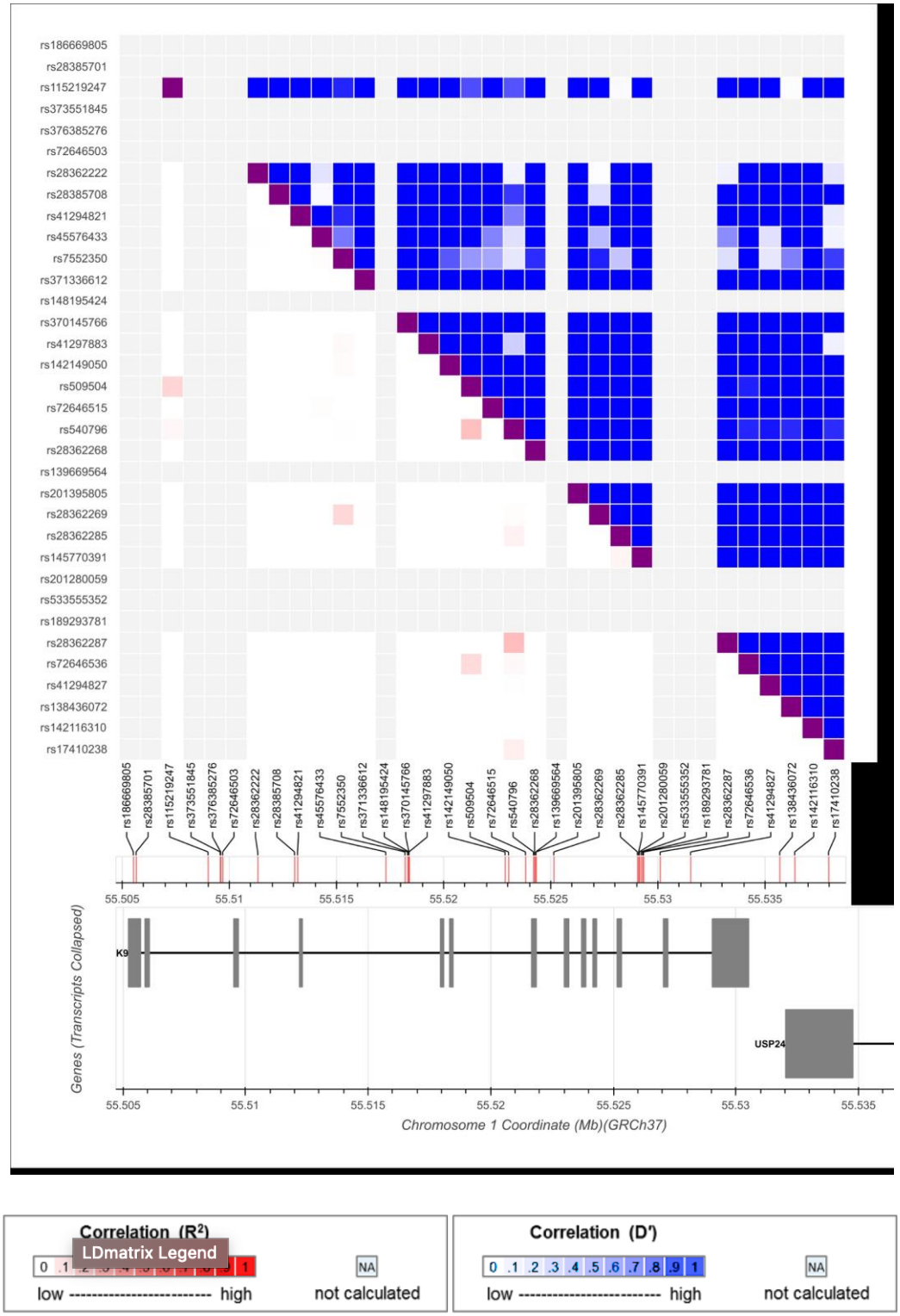

Figure 2. Linkage Disequilibrium Heat Plot Example

-Linkage Disequilibrium Heat Plot made from LDLink. Red indicates high LD between SNPs.

Overall, this study successfully investigated the distribution pattern of *PCSK9* allele frequencies and their associations with plasma lipid profile in two previously uncharacterized populations, including African Blacks and NHWs. This study identified multiple common and rare *PCSK9* variants in both populations. However, only nine SNPs, seven in African Blacks and two in NHWs, demonstrated significant associations with plasma lipid profile. This is most like due the relatively smaller sample sizes in both groups. The relationship between rs28942112 (Phe216Leu) and abnormal lipid levels in one African Black subject appears to confirm its association with the FH pattern from literature. Future studies on larger African Blacks and NHWs samples will probably identify more significant associations with both common and rare variants.

5.0 Public Health Significance

Heart disease is the most common cause of death in the United States. Previous genetic studies have identified and functionally characterized multiple rare and causal *PCSK9* variants associated with either FH or extremely low plasma LDL-C. The identification of latter mutations enabled the development of pharmacological treatment for FH by inhibiting PCSK9 that results in up-regulating LDL-receptors and consequently low plasma LDL-C. Additional genetic studies like ours, are useful to identify potential novel rare and functional variants associated with plasma lipid profile and to enrich the public databases. It is equally important to investigate variants in diverse ethnic populations, as done here, to further the goal of precision medicine the general population (Marc et al. 2013). As the population size grows in the United States and world-wide, chronic diseases will only increase, and it is important to investigate the ethnic-specific differences in the distribution of genetic variants associated with disease. We found that Nigerian African Blacks appear to be genetically more diverse in the *PCSK9* region and have a healthier lipid profile on average compared to a NHW sample from Colorado. A deeper investigation of SNPs identified in this study can contribute to precision medicine as well as to *PCSK9* inhibitor research related to heart disease research.

Appendix A

Appendix Table 1.

African Black	TC*		TG*		HDL-C*		LDL-C*	
Ref SNP ID	**Beta**	**P-Value**	**Beta**	**P-Value**	**Beta**	**P-Value**	**Beta**	**P-Value**
rs186669805	-0.755	5.85E-01	0.028	7.71E-01	1.784	4.82E-01	-2.617	3.22E-01
rs28385701	-4.201	2.69E-03	-0.097	3.25E-01	-0.446	8.62E-01	-8.388	1.72E-03
rs867904088	-1.691	1.35E-01	-0.028	7.28E-01	1.786	3.89E-01	-4.249	4.94E-02
rs1272703401	-2.004	2.26E-02	-0.094	1.29E-01	-0.079	9.61E-01	-3.763	2.45E-02
rs115219247	-0.859	9.18E-02	-0.003	9.32E-01	-0.455	6.39E-01	-1.206	2.16E-01
rs151095149	1.034	5.96E-01	0.305	2.64E-02	-3.586	3.18E-01	1.55	6.77E-01
rs373551845	1.454	2.93E-01	0.054	5.78E-01	1.869	4.60E-01	1.93	4.65E-01
rs376385276	0.302	7.07E-01	-0.043	4.42E-01	2.932	4.63E-02	-0.385	8.01E-01
rs28362222	-0.794	1.84E-01	0.034	4.18E-01	-1.973	7.04E-02	-1.163	3.09E-01
rs373018373	-0.004	9.95E-01	0.073	7.16E-02	-0.679	5.02E-01	0.416	6.92E-01
rs28385708	-0.178	6.93E-01	-0.019	5.46E-01	0.331	6.90E-01	-0.593	4.92E-01
rs41294821	-0.115	7.38E-01	-0.015	5.37E-01	0.897	1.63E-01	-0.041	9.51E-01
rs45576433	0.479	3.47E-01	0.003	9.31E-01	-0.231	8.05E-01	1.232	2.05E-01
rs28942112	1.42	2.11E-01	0.036	6.51E-01	0.513	8.06E-01	2.716	2.08E-01
rs7552350	-0.115	3.43E-01	0.004	6.71E-01	-0.336	1.30E-01	-0.201	3.82E-01
rs371336612	0.054	9.62E-01	-0.027	7.37E-01	1.152	5.78E-01	-0.231	9.15E-01

rs150169598	-0.043	9.54E-01	-0.001	9.88E-01	1.366	3.19E-01	-0.629	6.56E-01
rs148195424	-0.417	8.33E-01	0.083	5.52E-01	-3.028	4.02E-01	-0.14	9.70E-01
rs370145766	-0.684	1.99E-01	0.05	1.85E-01	-0.257	7.92E-01	-1.56	1.24E-01
rs41297883	-0.106	8.07E-01	0.102	8.82E-04	-0.674	3.95E-01	-0.423	6.11E-01
rs147478188	-1.145	1.56E-01	-0.003	9.63E-01	1.096	4.58E-01	-3.004	5.16E-02
rs376945520	-1.727	7.83E-02	-0.115	9.57E-02	-0.42	8.16E-01	-2.944	1.16E-01
rs865914100	-0.638	7.44E-01	0.022	8.74E-01	-2.811	4.33E-01	-0.617	8.69E-01
rs376753957	-0.425	4.23E-01	0.022	5.57E-01	-1.583	1.08E-01	-0.539	5.93E-01
rs142149050	-0.367	3.29E-01	0	9.92E-01	-0.516	4.56E-01	-0.752	2.95E-01
rs509504	-0.34	1.03E-01	0.016	2.93E-01	-0.566	1.42E-01	-0.505	2.05E-01
no known variants (nearest: 55523061)	-1.45	2.95E-01	-0.047	6.29E-01	-0.771	7.60E-01	-2.373	3.70E-01
rs754420902	0.354	7.18E-01	0.021	7.58E-01	-0.126	9.44E-01	1.04	5.80E-01
rs368257906	0.169	7.87E-01	0.088	6.04E-02	0.026	9.82E-01	0.568	6.35E-01
rs72646515	-0.249	7.38E-01	-0.063	1.97E-01	0.238	8.63E-01	-0.506	7.05E-01
rs540796	0.061	6.46E-01	0.008	4.00E-01	-0.303	2.15E-01	0.233	3.58E-01
rs376388695	-0.181	8.23E-01	0.04	5.15E-01	1.199	4.18E-01	0.003	9.99E-01
rs28362268	-0.271	4.59E-01	-0.009	7.15E-01	0.004	9.95E-01	-0.656	3.38E-01
rs375582388	-0.553	4.26E-01	-0.036	4.62E-01	0.565	6.59E-01	-1.31	3.22E-01
rs139669564	0.499	6.59E-01	0.014	8.84E-01	-1.076	6.03E-01	3.431	1.12E-01
rs147599496	0.674	4.41E-01	0.066	2.87E-01	0.786	6.27E-01	0.878	5.98E-01
rs201395805	-0.972	4.80E-01	-0.013	8.91E-01	1.237	6.25E-01	-2.796	2.89E-01

rs143394031	-0.468	5.00E-01	0.017	7.22E-01	1.159	3.64E-01	-1.498	2.56E-01
rs28362269	-0.239	1.77E-01	-0.003	8.11E-01	-0.051	8.76E-01	-0.563	9.65E-02
rs72646525	-0.95	1.73E-01	-0.025	6.13E-01	1.332	2.96E-01	-2.602	5.14E-02
rs367606156	0.065	9.10E-01	0.039	3.52E-01	0.793	4.48E-01	0.282	7.96E-01
rs755750316	-0.595	4.60E-01	-0.004	9.44E-01	-0.856	5.63E-01	-1.038	5.01E-01
rs145264023	-0.119	8.91E-01	0.102	9.74E-02	0.872	5.86E-01	-1.128	4.96E-01
rs142118418	1.908	1.70E-01	0.028	7.77E-01	-0.954	7.09E-01	4.212	1.12E-01
rs28362285	0.495	1.96E-01	-0.007	7.94E-01	0.556	4.30E-01	1.057	1.49E-01
rs199815786	-0.505	2.58E-03	0.073	1.64E-09	-1.59	3.30E-07	-0.848	8.07E-03
rs145770391	0.878	5.26E-01	0.033	7.36E-01	0.5	8.44E-01	1.345	6.11E-01
rs201280059	1.537	6.07E-02	0.016	7.75E-01	1.54	3.06E-01	2.67	8.70E-02
rs371914056	0.236	7.50E-01	0.032	5.38E-01	0.816	5.51E-01	-0.114	9.35E-01
rs369851423	0.332	8.66E-01	0.115	4.03E-01	1.128	7.53E-01	-0.294	9.37E-01
rs142524469	-0.643	4.29E-01	-0.005	9.24E-01	0.482	7.47E-01	-1.512	3.28E-01
rs533555352	0.132	8.11E-01	0.024	5.39E-01	0.601	5.52E-01	0.114	9.13E-01
rs189293781	1.534	1.73E-01	0.09	2.62E-01	1.832	3.78E-01	2.039	3.44E-01
rs28362287	0.177	4.02E-01	-0.017	2.51E-01	-0.077	8.41E-01	0.407	3.17E-01
rs72646536	-0.077	9.01E-01	0.017	7.02E-01	0.22	8.48E-01	-0.296	8.04E-01
rs41294827	-0.387	1.25E-01	-0.033	6.80E-02	1.045	2.38E-02	-1.027	3.32E-02
rs138436072	-0.392	3.53E-01	0.019	5.25E-01	-1.559	4.50E-02	-0.348	6.65E-01
rs17410238	-0.021	9.62E-01	0.031	3.17E-01	-0.684	4.01E-01	0.347	6.82E-01

*Indicates corrected value with Box-Cox transformation for lipid trait.

Appendix B

Appendix Table 2.

NHWs	TC		TG*		HDL-C*		LDL-C	
Ref SNP ID	**Beta**	**P-value**	**Beta**	**P-value**	**Beta**	**P-value**	**Beta**	**P-value**
rs28385701	-11.675	4.10E-01	-0.146	1.68E-01	0.032	6.11E-01	-8.532	5.32E-01
rs376385276	-22.207	4.55E-01	-0.217	3.34E-01	0.079	5.48E-01	-23.009	4.21E-01
rs72646503	-5.453	8.97E-01	0.222	4.84E-01	-0.417	2.40E-02	24.617	5.43E-01
rs28385708	-0.872	8.74E-01	0.027	5.17E-01	0.013	5.86E-01	-4.065	4.44E-01
rs41294821	3.18	6.52E-01	-0.034	5.30E-01	-0.012	7.05E-01	3.162	6.41E-01
rs45576433	-3.027	4.10E-01	-0.01	7.15E-01	0.016	3.12E-01	-3.132	3.76E-01
rs200856421	-5.055	1.63E-01	0.072	8.82E-03	-0.19	3.22E-29	1.822	6.01E-01
rs7552350	-7.319	3.81E-02	-0.037	1.61E-01	-0.003	8.57E-01	-4.586	1.78E-01
rs370145766	-37.323	3.74E-01	-0.261	4.10E-01	0.069	7.07E-01	-34.004	4.00E-01
rs41297883	-7.739	6.53E-01	0.047	7.19E-01	-0.034	6.54E-01	-9.564	5.63E-01
rs147478188	32.502	4.40E-01	-0.131	6.80E-01	-0.013	9.43E-01	36.304	3.70E-01
rs865914100	-6.809	8.72E-01	-0.17	5.92E-01	0.088	6.37E-01	-11.76	7.72E-01
rs376753957	12.913	6.64E-01	-0.15	5.04E-01	0.038	7.72E-01	12.273	6.67E-01
rs540796	4.495	1.60E-01	0.029	2.30E-01	-0.001	9.37E-01	2.613	3.96E-01

rs376388695	-6.985	8.68E-01	-0.172	5.89E-01	0.088	6.36E-01	-11.906	7.69E-01
rs143394031	-5.055	1.63E-01	0.072	8.82E-03	-0.19	3.22E-29	1.822	6.01E-01
rs28362285	-6.234	8.82E-01	0.236	4.58E-01	-0.201	2.77E-01	-7.051	8.62E-01
rs533555352	-6.733	8.73E-01	-0.17	5.94E-01	0.089	6.29E-01	-11.833	7.70E-01
rs189293781	-29.305	4.86E-01	-0.297	3.49E-01	0.123	5.06E-01	-31.841	4.31E-01
rs28362287	0.373	9.75E-01	-0.007	9.39E-01	0.027	6.00E-01	-2.399	8.32E-01
rs41294827	-0.877	9.04E-01	-0.032	5.60E-01	-0.007	8.32E-01	-1.907	7.85E-01
rs142116310	-16.736	1.28E-01	-0.149	7.16E-02	0.022	6.45E-01	-15.973	1.31E-01
rs17410238	3.749	2.65E-01	0.022	3.93E-01	-0.002	8.96E-01	2.448	4.49E-01
rs770231986	-15.589	5.99E-01	-0.175	4.34E-01	0.037	7.78E-01	-15.652	5.83E-01

*Indicates corrected value with Box-Cox transformation for lipid trait.

Bibliography

Abifadel M, Varret M, Rabès J-P, Allard D, Ouguerram K, Devillers M, Cruaud C, Benjannet S, Wickham L, Erlich D, Derré A, Villéger L, Farnier M, Beucler I, Bruckert E, Chambaz J, Chanu B, Lecerf J-M, Luc G, Moulin P, Weissenbach J, Prat A, Krempf M, Junien C, Seidah NG and Boileau C. "Mutations in PCSK9 cause autosomal dominant hypercholesterolemia." *Nature Genetics* 2003; 34 (2): 154-156. *https://doi.org/10.1038/ng1161*.

Auer PL and Lettre G. "Rare variant association studies: Considerations, challenges and opportunities." *Genome Medicine* 2015; 7 (16): 1-11. https://doi.org/10.1186/s13073-015-0138-2.

Bunker CH, Ukoli FA, Okoro FI, Olomu AB, Kriska AM, Huston SL, Markovic N and Kuller LH. "Correlates of serum lipids in a lean black population." *Atherosclerosis* 1996; 123 (1-2): 215-225. *https://doi.org/10.1016/0021-9150(96)05810-8*.

Bryant EK, Dressen AS, Bunker CH, Hokanson JE, Hamman RF, Kamboh MI and Demirci FY. "A multiethnic replication study of plasma lipoprotein levels-associated SNPs identified in recent gwas." *PLoS ONE* 2013; 8 (5): e63469. https://doi.org/10.1371/journal.pone.0063469.

Campbell MC and Tishkoff SA. "African genetic diversity: Implications for human demographic history, modern human origins, and complex disease mapping." *Annual Review of Genomics and Human Genetics* 2008; 9 (1): 403-433. https://doi.org/10.1146/annurev.genom.9.081307.164258.

Chemello K, García-Nafría J, Gallo A, Martín C, Lambert G and Blom D. "Lipoprotein metabolism in familial hypercholesterolemia." *Journal of Lipid Research* 2021; 62 (100062): 1-16. *https://doi.org/10.1016/j.jlr.2021.100062*.

Cohen J, Pertsemlidis A, Kotowski IK, Graham R, Garcia CK and Hobbs HH. "Low ldl cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9." *Nature Genetics* 2005; 37 (2): 161-165. *https://doi.org/10.1038/ng1509*.

Cohen JC, Boerwinkle E, Mosley TH and Hobbs HH. "Sequence variations in PCSK9,low ldl, and protection against coronary heart disease." *New England Journal of Medicine* 2006; 354 (12): 1264-1272. *https://doi.org/10.1056/nejmoa054013*.

Coram MA, Duan Q, Hoffman TJ, Thornton T, Knowles JW, Johnson NA, Ochs-Balcom HM, Donlon TA, Martin LW, Eaton CB, Robinson JG, Risch NJ, Zhu X, Li Y, Reiner AP, and Tang, H. "Genome-wide characterization of shared and distinct genetic components that influence blood lipid levels in ethnically diverse human populations." *The American Journal of Human Genetics* 2013; 92 (6): 904-916. *https://doi.org/10.1016/j.ajhg.2013.04.025.*

El Khoury P, Elbitar S, Ghaleb Y, Khalil YA, Varret M, Boileau C and Abifadel M. "PCSK9 mutations in familial hypercholesterolemia: From a groundbreaking discovery to anti-PCSK9 therapies." *Current Atherosclerosis Reports* 2017; 19 (12): *https://doi.org/10.1007/s11883-017-0684-8*.

Expert Panel on Detection E, and Treatment of High Blood Cholesterol in Adults. "Executive summary of the third report of the national cholesterol education program (ncep) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (adult treatment panel iii)." *JAMA* 2001; 285 (19): 2486–2497. *https://doi.org/doi:10.1001/jama.285.19.2486*.

Guo Q, Feng X and Zhou Y. "PCSK9 variants in familial hypercholesterolemia: A comprehensive synopsis." *Front Genet* 2020; 11 (1020): 1-13. *https://doi.org/10.3389/fgene.2020.01020*.

Heron M. "Deaths: Leading causes for 2017." *National Vital Statistics Reports* 2019; 68 (6): 1-77. *https://www.cdc.gov/nchs/data/nvsr/nvsr68/nvsr68_06-508.pdf*

Howe KL, Achuthan P, Allen J, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, Azov AG, Bennett R, Bhai J, Billis K, Boddu S, Charkhchi M, Cummins C, Luca, Davidson C, Dodiya K, Bilal, Fatima R, Gall A, Carlos, Grego T, Guijarro-Clarke C, Haggerty L, Hemrom A, Hourlier T, Izuogu OG, Juettemann T, Kaikala V, Kay M, Lavidas I, Le T, Lemos D, Jose, Marugán JC, Maurel T, McMahon AC, Mohanan S, Moore B, Muffato M, Oheh DN, Paraschas D, Parker A, Parton A, Prosovetskaia I, Sakthivel MP, Ahamed, Schmitt BM, Schuilenburg H, Sheppard D, Steed E, Szpak M, Szuba M, Taylor K, Thormann A, Threadgold G, Walts B, Winterbottom A, Chakiachvili M, Chaubal A, Nishadi, Flint B, Frankish A, Hunt SE, Iisley GR, Langridge N, Loveland JE, Martin FJ, Mudge JM, Morales J, Perry E, Ruffier M, Tate J, Thybert D, Trevanion SJ, Cunningham F, Yates AD, Zerbino DR and Flicek P. "Ensembl 2021." *Nucleic Acids Research* 2021; 49 (D1): 884-891. *https://doi.org/10.1093/nar/gkaa942.*

Iluminia. "Infinium™ global diversity array-8 v.1.0 datasheet." Published 2021. *https://www.illumina.com/content/dam/illumina/gcs/assembled-assets/marketing-literature/infinium-global-diversity-array-data-sheet-m-gl-00153/infinium-global-diversity-array-data-sheet-m-gl-00153.pdf.*

Kwon HJ, Lagace TA, McNutt MC, Horton JD, Deisenhofer J. "Molecular basis for ldl receptor recognition by PCSK9." *Proc Natl Acad Sci U S A* 2008; 105 (6): 1820-1825. *https://doi.org/10.1073/pnas.0712064105*.

Machiela MJ and Chanock SJ. “Ldlink: A web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants.” *Bioinformatics* 2015; 31 (21): 3555-3557. *https://doi.org/10.1093/bioinformatics/btv402.*

Maxwell KN, Breslow J. L. “Proprotein convertase subtilisin kexin 9: The third locus implicated in autosomal dominant hypercholesterolemia.” *Current Opinion in Lipidology* 2005; 16 (2): 167-172. *https://doi.org/10.1097/01.mol.0000162321.31925.a3*.

National Center for Biotechnology Information. ClinVar; [VCV000922888.2], [VCV000440709.1], [VCV000262902.10], (accessed April 7, 2022).

Onitiri AC, Sander, M., Boyo, A. E.. “Serum lipids and lipoproteins in healthy africans.” *Clinica Chimica Acta* 1977; 81 57-61. *https://doi.org/10.1016/0009-8981(77)90413-2*

Park SW, Moon Y-A and Horton JD. “Post-transcriptional regulation of low-density lipoprotein receptor protein by proprotein convertase subtilisin/kexin type 9a in mouse liver.” *Journal of Biological Chemistry* 2004; 279 (48): 50630-50638. https://doi.org/10.1074/jbc.m410077200.

Pham NH, Truong, P. K., Lao, T. D., & Le, T. A. . “Proprotein convertase subtilisin/kexin type 9 gene variants in familial hypercholesterolemia: A systematic review and meta-analysis.” *Processes* 2021; 9 (2): 283. *https://doi.org/10.3390/pr9020283*.

Pirim D, Bunker CH, Hokanson JE, Hamman RF, Demirci FY and Kamboh MI. “Hepatic lipase (lipc) sequencing in individuals with extremely high and low high-density lipoprotein cholesterol levels.” *PLOS ONE* 2020; 15 (12): e0243919. https://doi.org/10.1371/journal.pone.0243919.

Pirim D, Radwan ZH, Wang X, Niemsiri V, Hokanson JE, Hamman RF, Feingold E, Bunker CH, Demirci FY and Kamboh MI. “Apolipoprotein e-c1-c4-c2 gene cluster region and inter-individual variation in plasma lipoprotein levels: A comprehensive genetic association study in two ethnic groups.” *PLOS ONE* 2019; 14 (3): e0214060. *https://doi.org/10.1371/journal.pone.0214060*.

Rafieian-Kopaei M, Setorki, M., Doudi, M., Baradaran, A., & Nasri, H. “Atherosclerosis: Process, indicators, risk factors and new hopes.” *International Journal of Preventative Medicine* 2014; 5 (8): 927-946. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4258672/

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rewers M, Shetterly, S. M., & Hamman, R. F. “Hypertension among rural hispanics and non-hispanic whites: The san luis valley diabetes study.” *Public Health Reports (Washington, D.C.: 1974)* 1996; 111 27-29. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1381658/

Rosenson RS, Hegele RA, Fazio S and Cannon CP. "The evolving future of pcsk9 inhibitors." *Journal of the American College of Cardiology* 2018; 72 (3): 314-329. *https://doi.org/10.1016/j.jacc.2018.04.054*.

Sabatine MS, Giugliano RP, Keech AC, Honarpour N, Wiviott SD, Murphy SA, Kuder JF, Wang H, Liu T, Wasserman SM, Sever PS and Pedersen TR. "Evolocumab and clinical outcomes in patients with cardiovascular disease." *New England Journal of Medicine* 2017; 376 (18): 1713-1722. https://doi.org/10.1056/nejmoa1615664.

Sanchis-Gomar F, Perez-Quilis C, Leischik R and Lucia A. "Epidemiology of coronary heart disease and acute coronary syndrome." *Ann Transl Med* 2016; 4 (13): 256. *https://doi.org/10.21037/atm.2016.06.33*.

Schwartz GG, Steg PG, Szarek M, Bhatt DL, Bittner VA, Diaz R, Edelberg JM, Goodman SG, Hanotin C, Harrington RA, Jukema JW, Lecorps G, Mahaffey KW, Moryusef A, Pordy R, Quintero K, Roe MT, Sasiela WJ, Tamby J-F, Tricoci P, White HD and Zeiher AM. "Alirocumab and cardiovascular outcomes after acute coronary syndrome." *New England Journal of Medicine* 2018; 379 (22): 2097-2107. https://doi.org/10.1056/nejmoa1801174.

Soutar AK and Naoumova RP. "Mechanisms of disease: Genetic causes of familial hypercholesterolemia." *Nat Clin Pract Cardiovasc Med* 2007; 4 (4): 214-25. *https://doi.org/10.1038/ncpcardio0836*.

Sun XM, Eden ER, Tosi I, Neuwirth CK, Wile D, Naoumova RP and Soutar AK. "Evidence for effect of mutant pcsk9 on apolipoprotein b secretion as the cause of unusually severe dominant hypercholesterolaemia." *Hum Mol Genet* 2005; 14 (9): 1161-9. *https://doi.org/10.1093/hmg/ddi128*.

Virani SS, Alonso A, Aparicio HJ, Benjamin EJ, Bittencourt MS, Callaway CW, Carson AP, Chamberlain AM, Cheng S, Delling FN, Elkind MSV, Evenson KR, Ferguson JF, Gupta DK, Khan SS, Kissela BM, Knutson KL, Lee CD, Lewis TT, Liu J, Loop MS, Lutsey PL, Ma J, Mackey J, Martin SS, Matchar DB, Mussolino ME, Navaneethan SD, Perak AM, Roth GA, Samad Z, Satou GM, Schroeder EB, Shah SH, Shay CM, Stokes A, Vanwagner LB, Wang N-Y and Tsao CW. "Heart disease and stroke statistics—2021 update." *Circulation* 2021; 143 (8): 254-743. *https://doi.org/10.1161/cir.0000000000000950*.

Youngblom E. PM, Knowles JW. "Familial hypercholesterolemia." *GeneReviews*® 2016, *https://www.ncbi.nlm.nih.gov/books/NBK174884/*

Zhao Z, Tuakli-Wosornu Y, Lagace TA, Kinch L, Grishin NV, Horton JD, Cohen JC and Hobbs HH. "Molecular characterization of loss-of-function mutations in pcsk9 and identification of a compound heterozygote." *The American Journal of Human Genetics* 2006; 79 (3): 514-523. *https://doi.org/10.1086/507488*.