THE GENETIC CONTRIBUTIONS TO HAART-ASSOCIATED DYSLIPIDEMIA

by

Matthew James Nicholaou

BS, University of Delaware, 2002

Submitted to the Graduate Faculty of

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Doctor of Public Health

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

Graduate School of Public Health

This dissertation was presented

by

Matthew J. Nicholaou

It was defended on

July 11, 2011

and approved by

Dissertation Advisor: Jeremy Martinson, D.Phil.
Assistant Professor
Department of Infectious Diseases and Microbiology and Dept. of Human Genetics
Graduate School of Public Health
University of Pittsburgh

Committee Member: Rhobert W. Evans, Ph.D.
Assistant Professor
Department of Epidemiology
Graduate School of Public Health
University of Pittsburgh

Committee Member: Eleanor Feingold, Ph.D.
Professor
Department of Human Genetics
Graduate School of Public Health
University of Pittsburgh

Committee Member: Lawrence Kingsley, Dr.PH.
Professor
Department of Infectious Diseases and Microbiology and Dept. of Epidemiology
Graduate School of Public Health
University of Pittsburgh

THE GENETIC CONTRIBUTIONS TO HAART-ASSOCIATED DYSLIPIDEMIA

Matthew J. Nicholaou, DrPH

University of Pittsburgh, 2011

Highly active anti-retroviral therapy (HAART) has been successful in delaying the progression to AIDS in HIV infected individuals. Exposure to HAART can result in metabolic side effects such as dyslipidemia and lipodystrophy in a subset of treated patients. We used a custom designed Illumina GoldenGate Genotyping assay to investigate the genetic susceptibility to dyslipidemia attributed to HIV infection and HAART treatment. 1,945 men were selected from the Multicenter AIDS Cohort Study (MACS) for genotyping and phenotypic analysis of serum lipid levels. This population was stratified by biogeographical ancestry and HIV/HAART status. Among men of European ancestry, those who were infected with HIV and receiving HAART had significantly lower serum low-density lipoprotein cholesterol (LDL-C, $P = 1.90 \times 10^{-4}$) and high-density lipoprotein cholesterol levels (HDL-C, $P < 1.00 \times 10^{-7}$), with significantly higher serum triglyceride (TRIG, $P < 1.00 \times 10^{-7}$) levels compared to HIV/HAART (-/-) controls. Among men of mixed African and European ancestry, those who were HIV/HAART (+/+) had significantly lower LDL-C ($P = 1.80 \times 10^{-4}$) levels compared to HIV/HAART (-/-) controls. Four SNPs; rs1532624 ($P = 1.66 \times 10^{-5}$), rs1532625 ($P = 2.36 \times 10^{-5}$), rs711752 ($P = 4.48 \times 10^{-5}$), and rs708272 ($P = 4.59 \times 10^{-5}$), located in the *CETP* gene region on chromosome 16 had statistically significant associations with serum HDL-C levels in HIV/HAART (+/+) European men. One SNP, rs261334 ($P = 6.53 \times 10^{-6}$), located in the *LIPC* gene region on chromosome 15 was associated with serum LDL-C levels and another SNP, rs4783961 ($P = 9.83 \times 10^{-6}$) located in the

CETP gene region, was associated with HDL-C levels in HIV/HAART (+/+) men of mixed African and European ancestry. These results show that dyslipidemia attributed to HAART varies depending on biogeographical ancestry and implicates two genes associated with serum lipid levels in these patients. Understanding the mechanism of HAART-associated dyslipidemia is important to global public health because nearly half of the estimated 30 million individuals infected with HIV are receiving or eligible to receive these drugs and are at risk of these HAART related side effects. Our results can also aid in identifying those individuals at greatest risk of developing HAART-associated dyslipidemia, which could improve monitoring and management of care given to these patients.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would first like to thank the faculty, which through their help and guidance, made this body of work possible: my advisor Dr. Jeremy Martinson for being my scientific mentor over the past five years; my doctoral committee, Dr. Rhobert Evans, Dr. Eleanor Feingold, and Dr. Lawrence Kingsley for providing invaluable input and guidance throughout my doctoral research; Dr. Daniel Weeks and Dr. Michael Barmada for helping me develop the computing skills needed to study the genetics of complex diseases; Dr. John Wilson and Dr. Ada Youk for teaching me the diligent art of statistical analysis. I would also like to thank the staff and faculty of the Department of Infectious Diseases and Microbiology for providing the framework and environment to allow me to conduct my research and earn my doctoral degree. I would like to thank my parents Frank and Terri Francino, my grandparents George and Lois Nicholaou, my friends Adam Soloff, Jessica Malzahn, Rebecca Bosko Marino, Tom Cordrey, and Thomas Housley for providing me inspiration, support and love. Without them, none of my achievements would be possible.

This dissertation is dedicated in loving memory of my father Robert J. Kingston.

Table 1 - Abbreviations

Table of Abbreviations

Acronym	Meaning
(-/-)	HIV negative / HAART naïve
(+/-)	HIV positive / HAART naïve
(+/+)	HIV positive / HAART positive
ADT	Assay design tool
AEA	African and European ancestry
AIDS	Acquired Immunodeficiency Syndrome
AIM	Ancestry Informative Markers
ANOVA	Analysis of Variance
APOA5	Apolipoprotein A-V
APOB	Apolipoprotein B
ASO	Allele-specific oligo
ASW	African ancestry in southwest USA
AZT	Zidovudine
BGA	Biogeographical Ancestry
BnH	Black non-Hispanic
BUD13	BUD13 homolog
CAC	Coronary Artery Calcification
CAD	Coronary Artery Disease
CAMACS	Center for the Analysis and Management of the Multicenter AIDS Cohort Study
CCR5	C-C Chemokine Receptor Type 5
CD4	Cluster of Differentiation 4
CETP	Cholesteryl ester transfer protein
CEU	Utah residents with ancestry from northern and western Europe
CHD	Chinese in Metropolitan Denver
CHD	Coronary Heart Disease
CNV	Copy number variation
CRF	Circulating Recombinant Forms
CRP	C-reactive protein
CT	Computed Tomography
CVD	Cardiovascular Disease
CXCR4	C-X-C Chemokine Receptor Type 4
D:A:D	Data Collection on Adverse Events of Anti-HIV Drugs
DC-SIGN	Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin
DNA	Deoxyribonucleic Acid
EA	European ancestry
Env	Envelope

Acronym	Meaning
FDA	Food and Drug Administration
FDR	False Discovery Rate
FP	Fluorescent Polarization
GALT	Gut Associated Lymphatic Tissue
GC	GenCall
GLM	Generalize linear model
gp	Glycoprotein
GPCL	Genomics and Proteomics Core Laboratories
GWAS	Genome-wide association study
HAART	Highly Active Antiretroviral Therapy
HDL-C	High-density lipoprotein cholesterol
HIV	Human Immunodeficiency Virus
HLA	Human Leukocyte Antigen
HMG-CoA	3-hydroxy-3-methyl-glutaryl-CoA reductase
IBS	Identity-by-State
IMT	Intima-media Thickness
LD	Linkage disequilibrium
LDL-C	Low-density lipoprotein cholesterol
LDLR	LDL receptor
LIPC	Lipase, hepatic
LSO	Locus-specific oligo
LTR	Long Terminal Repeat
MACS	Multicenter AIDS Cohort Study
MAF	Minor allele frequency
MDS	Multidimensional Scaling
MEX	Mexican ancestry in Los Angeles
mHDL-C	mean HDL-C
MI	Myocardial Infarction
mLDL-C	mean LDL-C
mRNA	Messenger RNA
MSM	Men who have sex with men
mTCHOL	mean TCHOL
mTRIG	mean TRIG
Nef	Negative Regulatory Factor
NNRTI	Non-nucleoside reverse transcriptase inhibitor
NRTI	Nucleoside analog reverse transcriptase inhibitor
PBMC	Peripheral Blood Mononuclear Cells
PCR	Polymerase Chain Reaction
PI	Protease inhibitor
Pol	Polymerase
QC	Quality control
QQ-plot	Quantile-Quantile plot

Acronym	Meaning
RCT	Reverse Cholesterol Transport
RNA	Ribonucleic Acid
RT	Reverse Transcriptase
SIV	Simian Immunodeficiency Virus
SMAC	Supramolecular activation clusters
SNP	Single nucleotide polymorphism
SREBP2	Sterol regulatory element binding transcription factor 2
ss	Single Stranded
SRR	Self-Reported Race
T-Cell	T Lymphocyte
Tat	Trans-activator of Transcription
TCHOL	Total cholesterol
TRIG	Triglycerides
TukeyHSD	Tukey's Honestly Significant Difference test
UTR	Untranslated region
Vif	Viral Infectivity Factor
VLDL	Very-low density lipoprotein
Vpr	Viral Protein R
Vpu	Viral Protein U
WGA	Whole genome amplification
WnH	White non-Hispanic
YRI	Yoruba in Ibadan, Nigeria
ZNF259	Zinc finger protein 259

1.0 INTRODUCTION

1.1 HUMAN IMMUNODEFICIENCY VIRUS

1.1.1 History and Epidemiology

In June 1981 the Centers for Disease Control and Prevention (CDC) published a report of five homosexual men with Pneumocystis pneumonia in Los Angeles, California[1]. In July and August of that year the CDC published two more Morbidity and Mortality Weekly Reports (MMWR) papers that reported additional cases of Pneumocystis pneumonia and Kaposi's Sarcoma in homosexual men from New York City and California[2,3]. These three reports would be the first to identify patients suffering from an infectious disease that would become the most significant public health issue of the 20th century.

Pneumocystis pneumonia and Kaposi's Sarcoma are extremely rare diseases that are usually only seen in people with severely impaired immune systems. All of the patients identified in the MMWR reports lacked any known underlying condition that would explain this immunodeficiency. A CDC task force found a pattern of shared symptoms in these cases: low CD4+ T-cell counts, lymphadenopathy, and recurrent opportunistic infections. They would come

to name this condition Acquired Immune Deficiency Syndrome, AIDS. It was soon observed that a similar syndrome existed in hemophiliacs and blood transfused patients[4]. The hemophiliacs gave the first clue to the causative agent of AIDS. These patients received purified blood products, factor 8 or 9, from pooled blood donors for treatment of their hemophilia. These blood products are filtered to remove any bacterial or fungal contaminates and any soluble toxic compounds, which suggested the possibility of a viral agent. In 1983 two groups isolated this virus simultaneously. After some controversy, it was agreed to call this newly discovered virus Human Immunodeficiency Virus (HIV)[5,6].

In the 30 years since the first reports of AIDS in the United States, HIV has spread to every corner of the globe, currently infecting approximately 33 million people (Figure 1)[7]. The

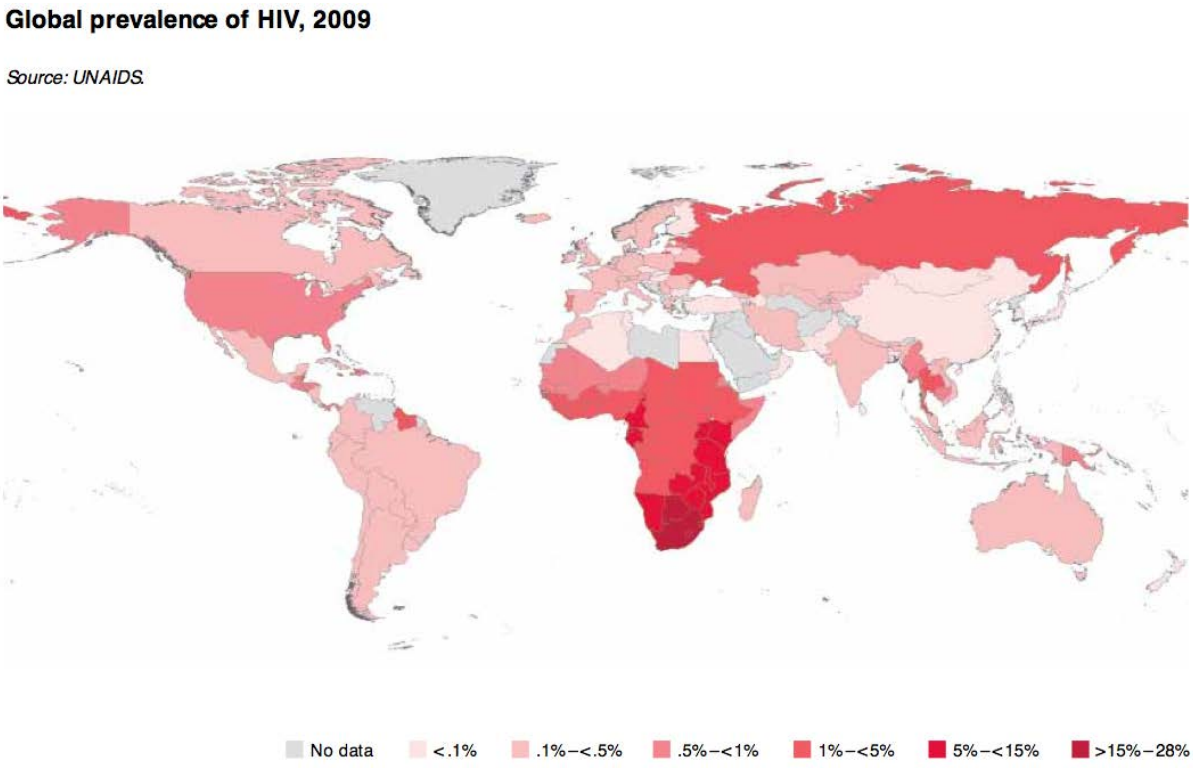


Figure 1 - Global Prevalence of HIV, 2009

global AIDS pandemic has essentially stabilized with rates of newly infected people declining since 1997 when new infections peaked at approximately 3.2 million. This decrease can be attributed to improved surveillance, prevention and education. However, not all countries are following this global trend of decreasing HIV incidence rates. AIDS related deaths are also on the decline, from a peak in 2004 at approximately 2.1 million to an estimated 1.8 million in 2009. This decrease can be attributed to increased availability of antiretroviral drugs and improved care of people living with HIV infection. The use of antiretroviral therapy, specifically highly active antiretroviral therapy (HAART), has altered the natural course of HIV infection and the progression to AIDS. HAART was established as the standard treatment for people infected with HIV in the late 90's. HAART consists of a combination of multiple antiretroviral drugs that inhibit different stages of the HIV life cycle. While there is currently no cure for HIV infection, HAART has extended the time by which HIV infection progresses to AIDS by over a decade. Globally, the number of HIV infected people receiving antiretroviral therapy has expanded 13-fold since 2004. This has contributed to the 19% decline in annual HIV related deaths since 2004.

With the prolonged lifespan of HIV infected individuals and the increasing number of patients receiving HAART, concerns have arisen regarding the side effects and potential health risks associated with long-term exposure to these drugs. HAART can result in metabolic disturbances such as insulin resistance, hypertriglyceridemia, hypercholesterolemia and morphologic changes pertaining to body fat distribution (lypodystrophy). These side effects are of particular concern because they resemble metabolic syndrome X, a combination of similar conditions seen in the HIV negative population that are significant cardiovascular disease risk

factors[8]. A subset of individuals receiving HAART experience drug related side effects, which suggests there may be a pre-existing genetic susceptibility present in these patients.

1.1.2 Biology & Infection

HIV is a *Lentivirus* in the family *Retroviridae*. Viruses in the family *Retroviridae* posses a RNA genome and utilize reverse transcriptase (RT), an RNA-dependent DNA polymerase, to synthesize a double-stranded DNA copy of the RNA viral genome during its infection cycle. The viruses in the genus *Lentivirus* are associated with long incubation times and are often called slow viruses. HIV is an enveloped virus with a single-stranded positive-sense RNA genome approximately 10kb long that codes for three structural polyproteins common to all retroviruses (Gag, Pol, and Env) and auxiliary proteins it uses for regulatory and accessory functions (Tat and Rev & Nef, Vif, Vpr, and Vpu, respectively). There are two types of HIV; HIV-1, which is the most prevalent globally, and HIV-2, which predominates mostly in West Africa. HIV-1 contains three major groups based on nucleotide sequence analysis; group M that includes most HIV-1 isolates and groups O and N, which represent rare outlier viruses. HIV-1 group M viruses can be broken up into ten distinct subtypes called clades, A to K. These clades are often distributed by geographical location. Coinfection of an individual with different subtypes can lead to circulating recombinant forms (CRFs). It is believed that HIV-1 and -2 are zoonotic infections that originated from nonhuman primates in Africa. Specifically, sequence analysis of HIV-1 shows that it is highly homologous to Simian Immunodeficiency Virus (SIV_{cpz}) isolated from chimpanzees and HIV-2 has a high homology with SIV_{smm} strains isolated from sooty mangabeys[9,10]. SIV strains are often used to infect Asian macaques to create an AIDS-like

disease model used for research, but wild-type SIV strains that infect the chimpanzee and sooty mangabey are well tolerated and do not result in an AIDS-like syndrome.

HIV can infect a variety of immune cells, principally through the interaction of the viral envelope protein gp160 (a dimmer of the gp120 + gp41 peptides) and cellular surface receptors[11]. (Figure 2) HIV requires two cellular receptors for viral fusion and entry; CD4 as its primary host receptor, and either CCR5 or CXCR4 as coreceptors. The use of specific

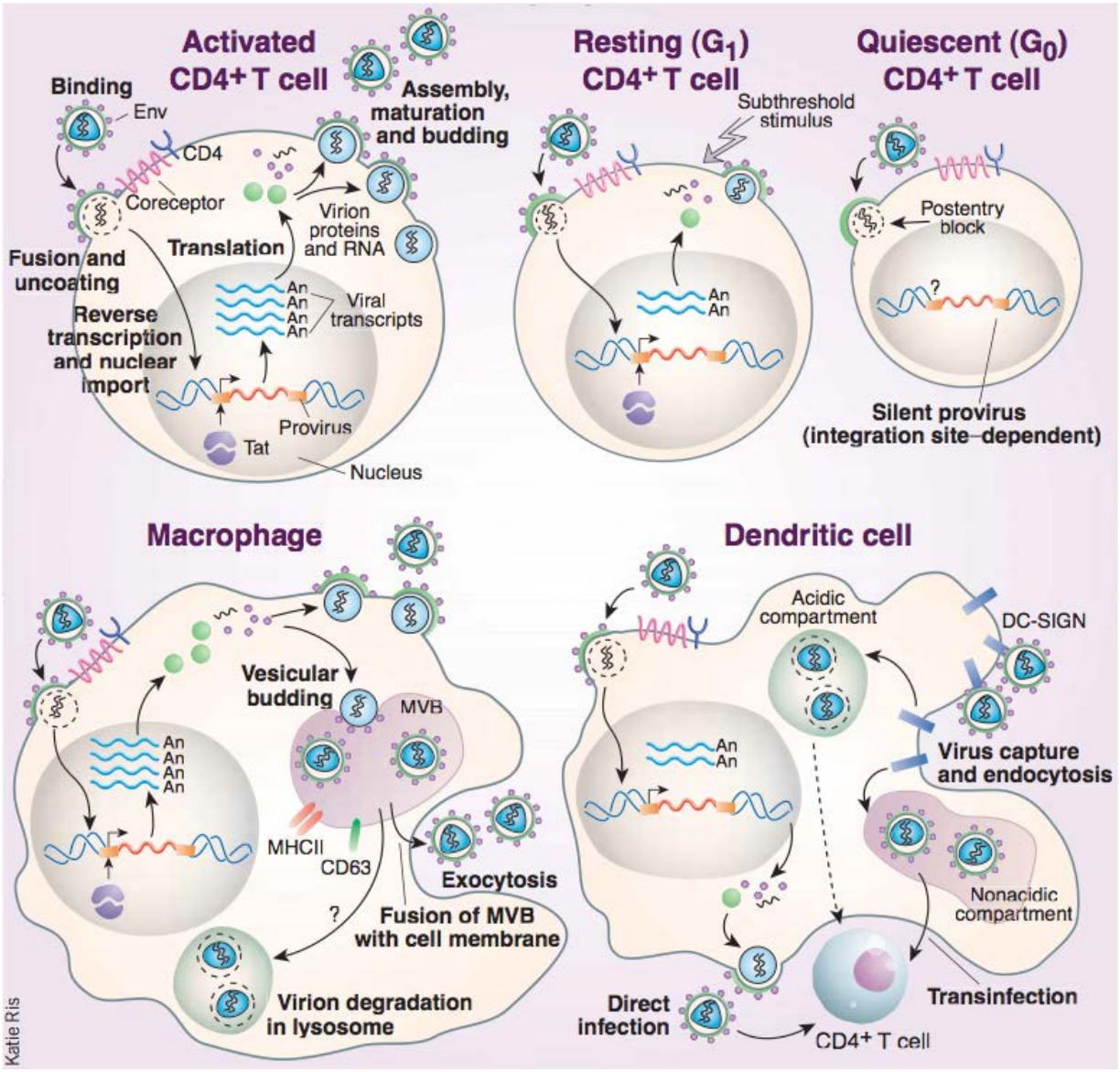


Figure 2 - HIV Lifecycle

coreceptors gives HIV its tropism, or selectivity for infecting certain cells; CCR5-tropic viruses predominately infect macrophages, and are often called Macrophage-tropic strains, whereas CXCR4-tropic viruses are called T-cell-line tropic strains. Mucosal infection of HIV-1 primarily occurs from CCR5-tropic viruses and progression to AIDS is often associated with a switch from CCR5 to CXCR4 tropic strains within an individual[12]. After viral fusion and uncoating the HIV-1 genome is released into the host cytoplasm and the ss(+)RNA is reverse transcribed by RT into a double-stranded DNA provirus. The proviral genome is then translocated to the nucleus by the viral integrase protein (a cleavage product of the Pol polyprotein) and is integrated into the host chromosome. The long-terminal repeats (LTR) in the integrated proviral genome contain enhancer and regulatory elements that allow for full-length viral RNA to be transcribed by host transcription factors with help from the viral accessory protein Tat. These full-length RNA transcripts are either alternatively spliced into viral messenger RNA (mRNA) or remain as full-length RNA that becomes packaged into new viral particles, serving as the viral genome for the next generation of progeny HIV. Some of the viral mRNA enters the cytoplasm to be translated by host ribosomes into the Gag-Pol precursor polyprotein, which is then proteolytically cleaved into the capsid, matrix, RT and integrase proteins. These viral proteins then form a complex and translocate to the inner surface of the host plasma membrane for viral assembly. Other viral mRNA molecules enter the endoplasmic reticulum where they are translated into the Env-polyprotein precursor. The Env proteins are transported into the Golgi apparatus to be glycosylated and cleaved to form mature viral Env proteins, which are then delivered to the host plasma membrane for viral particle assembly. The viral components assemble at the host plasma membrane and then bud from the infected cell. After budding, viral proteases cleave specific sites on the Gag and Gag-Pol precursors to produce a mature and infectious virion.

HIV-1 is transmitted primarily via unprotected sexual contact, via blood, and from mother to child. The predominant routes of transmission differ by geographical region. In the United States HIV transmission is predominately by men who have sex with men (MSM) and by intravenous drug use. In contrast, transmission is mostly by heterosexual contact in the rest of the world. CCR5 tropic HIV enters the host through genital or rectal mucosal surfaces. The exact mechanism of viral translocates across the mucosal surface is not known. It is thought that, amongst a pool of HIV viruses that are exposed to the mucosal surface during sexual intercourse, most infections start with the crossing of very few 'seed' viruses[13,14]. Once HIV has crossed the mucosa it must find permissive cells to infect. The primary targets for HIV are activated cells expressing the CD4 surface receptor, e.g. activated CD4+ T-Cells located at the genital or rectal mucosa. HIV can also be bound and phagocytized by Dendritic cells through their surface receptor DC-SIGN. These dendritic cells are not actively infected by HIV, but act as a transporter to lymphatic tissues present at the site of infection, such as Peyer's patches or peripheral lymph nodes where HIV is brought into contact with an abundance of activated CD4+ cells. HIV establishes local sites of infection at these lymphoid tissues because there is a large pool of cells for it to infect, resulting in lymphadenopathy and loss of gut associated lymphatic tissues (GALT) in acutely infected individuals[15].

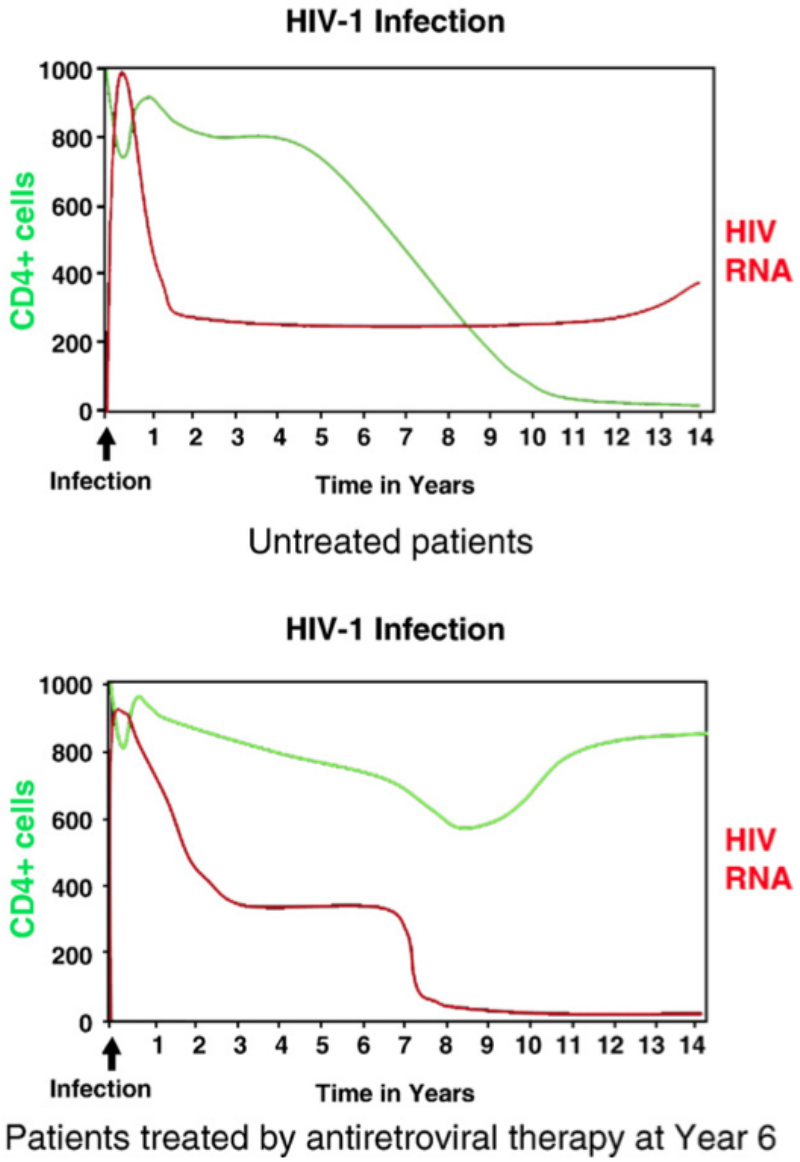


Figure 3 - HIV Infection

The course of HIV-1 infection is shown in Figure 3, and is characterized by three distinct stages; the acute phase, asymptomatic phase (latent phase) and the symptomatic phase (AIDS)[16]. During the acute phase of infection viral RNA copies in the plasma peak with an associated drop in CD4+ T-cell count. During this stage of infection viral replication is extremely high with up to $1x10^{10}$ virions being produced per day[17]. As the adaptive immune response increases the host begins to control infection resulting in a decrease in plasma viral load and an increase in CD4+

T-cells. This leads to the second stage of infection, the asymptomatic phase, lasting years where plasma viral loads stay low along with a progressive loss of CD4+ T-cells. The steady-state viremia during this stage of infection predicts how quickly the host will progress to AIDS[18]. HIV infected individuals begin to experience symptoms of AIDS when serum CD4+ T-cell levels drop below 200 cells per µl. This drop is associated with increased plasma viral load and loss of HIV specific CD8+ T-cells. The loss of these immune cells results in the immunodeficiency, and inability to fight opportunistic infections that is the hallmark of individuals with AIDS.

1.1.3 Antiretroviral Therapy

While there is still no cure for HIV infection, the use of antiretroviral drugs has greatly altered the natural course of disease progression. Zidovudine (AZT) was the first drug to be used in the Unites States for the treatment of AIDS. AZT was first used in 1987 as an antiretroviral drug but was initially developed as a potential anticancer drug. AZT was the first of a class of drugs known as, nucleoside analog reverse transcriptase inhibitors (NRTI). NRTIs inhibit the function of RT by competing with the cellular deoxynucleosides used during reverse transcription. NRTIs are incorporated into the viral replication products resulting in chain-termination due to the presence of the azido group in AZT. As soon as 6 months after AZT was implemented as a treatment for AIDS, viruses could be isolated from patients whose RT had become resistant to AZT inhibition[19]. HIV is able to develop these resistance mutations so quickly due to the high rate of replication and error-prone nature of RT itself. RT, unlike other DNA polymerases, lacks proofreading capability, resulting in a large amount of misincorporation – approximately 1 error for every 2000-5000 polymerized nucleotides[20].

With the development of AZT as a RT inhibitor work soon began on other drugs able to inhibit RT function. A second class of drugs was developed in the late 1980s and early 1990s called non-nucleoside RT inhibitors (NNRTI). NNRTIs inhibit RT differently than NRTIs; this class of drugs does not bind to the RT active site, but to a second allosteric site on RT that alters the enzyme's conformation and inhibits its function. The first NNRTI approved for use in the United States was nevirapine in 1996. Even before the FDA approved the drug, *in vitro* studies began to show HIV's ability to develop resistance mutations to nevirapine[21]. Again, HIV's error prone RT results in a high number of mutant viruses, which allows for the development of a resistant RT molecule very quickly.

The third major class of HIV antiretroviral drugs is protease inhibitors (PI). PIs prevent HIV infection by inhibiting the HIV-1 protease from cleaving viral proteins during virion maturation. This prevents the virion from maturing into an infective HIV particle. The first PI approved by the FDA was saquinavir in 1995. Similarly to the other classes of antiretrovirals, resistance mutations were observed *in vitro* before the FDA approved the drug. In 1996, after observing HIV's ability to rapidly evolve resistance mutations to singular drug therapies, physicians began to treat HIV patients with a combination of drugs. This treatment is called highly active antiretroviral therapy (HAART) and consists of at least one drug from each of the major classes of antiretrovirals; NRTIs, NNRTIs, and PIs. The premise of this therapy is that using combinations of drugs makes it very difficult for HIV to simultaneously develop resistance mutations. The use of HAART has been very effective in keeping viral replication low and CD4+ T-cell counts high in chronically HIV infected individuals. While HAART has greatly altered the natural history of HIV infection, there are a few caveats that come with this treatment; HAART is very expensive, it requires very rigid adherence to maintain effective levels in the

blood, if therapy is stopped HIV rebounds very quickly so HAART must be continually given[22], and there is a substantial risk of drug related side effects.

1.1.4 Infection and Cholesterol

A number of studies have shown that HIV infection and replication are intimately connected to cellular cholesterol. HIV entry occurs at sites in the cellular membrane called lipid rafts[23]. Lipid rafts are dynamic microdomains of proteins, cholesterol and sphingolipids within the cellular membrane. Lipid rafts are important to cell function because they bring many surface proteins and adaptor proteins in close proximity to each other, which enhances cellular signal transduction. Lipid rafts are important to a number of immune related functions, such as immune synapse formation and enhanced T-cell signaling by the formation of supramolecular activation clusters (SMACs)[24]. Lipid rafts are ideal sites for HIV-1 entry because they bring CD4 and CCR5/CXCR4 receptors in close proximity, which enhances viral attachment and fusion. HIV-1 entry is also dependent on the presence of cholesterol that is associated with lipid rafts[25]. In fact, the treatment of cells with lovastatin, a strong inhibitor of the rate-limiting enzyme during cholesterol biosynthesis, inhibited HIV-1 infection at the level of virus entry[26]. Additionally, treatment of cells with oxysterol, a known suppressor of cholesterol biosynthesis by inhibition of SREBP2 (an important transcriptional controller of cholesterol synthesis), inhibited viral replication[27].

The HIV accessory protein Nef is important to viral infection and has a number of roles in modulated cellular functions. During assembly of new viral particles, Nef binds and transports cholesterol to lipid rafts present at the cellular membrane by its myristoylation. This interaction of Nef with cholesterol appears to be crucial in HIV infectivity[28]. Nef also has the ability to

interact with a number of cellular signaling proteins. Specifically, Nef has been shown to activate a number of genes involved in cholesterol biosynthesis as well as being able to decrease cellular cholesterol efflux[28-31]. There is also evidence that HIV infection has effects on cholesterol metabolism at a systemic level. In the absence of HAART, HIV infection itself has been associated with hypocholesterolemia and hypertriglyceridemia[32-34]. There is also evidence that HIV infection causes a 3- to 4-fold increase in *de novo* hepatic lipogenesis when compared to HIV negative controls[35].

1.2 DYSLIPIDEMIA

1.2.1 Dyslipidemia, Atherosclerosis and Cardiovascular Disease Risk

Dyslipidemia is defined as abnormal levels of lipids in the blood. This includes both increases (hyper-) and/or decreases (hypo-) of serum cholesterol, fats, and lipoproteins. An increased level of certain lipids, such as low-density lipoprotein cholesterol (LDL-C), is associated with the development of atherosclerosis, which is the thickening of the arterial walls throughout the body. Cardiovascular disease (CVD), also called coronary heart disease (CHD) or coronary artery disease (CAD) when refereeing to the vasculature of the heart, is the clinical manifestation of atherosclerosis resulting in myocardial infarction and stroke. CVD is the number one killer in the United States, accounting for one in three deaths among men and women in 2007[36]. The major risk factors that have been associated with CHD in the general population include; blood pressure, cigarette smoking, total cholesterol (TCHOL), LDL-C, high-density lipoprotein (HDL-C), lipoprotein (a), triglycerides (TRIG), diabetes, obesity and family history of CVD[37-39].

Atherosclerosis is a complex disease process mediated largely by an inflammatory response to the deposition of lipids in arterial walls; because of this, certain indicators of immune function, e.g. C-reactive protein (CRP), have been implicated as CVD risk factors[40]. CVD and the risk factors associated with CVD vary by race and gender[41-43]. Dyslipidemia profiles tend to also vary by race and gender. Whites often have higher TRIG and lower HDL-C levels when compared to blacks in both men and women. Women tend to have higher TCHOL, LDL-C, and HDL-C levels compared to men.

Atherosclerosis starts at sites in arterioles where laminar flow is disrupted. LDL-C particles enter the arterial intima and begin to elicit an inflammatory response that results in the recruitment of monocytes from the blood. Upon entering the intima space, monocytes differentiate into macrophages and begin to ingest oxidized lipid particles, taking on the appearance of lipid-laden foam cells. The accumulation of foam cells results in fatty streaks or early atherosclerotic lesions. In addition to monocytes other immune cells, e.g. T- cells, are also recruited to the lesion further perpetuating the local inflammation. Over time, foam cells begin to undergo apoptosis releasing cellular debris and crystalline cholesterol resulting in the formation of a necrotic core, which further enhances the inflammatory response. At this stage the plaque can either stabilize preserving the luminal space or continue to grow, leading to a rupture of the fibrous cap and release of thrombotic material into the lumen. The thrombosis associated with this plaque rupture is what causes acute myocardial infarctions seen in CHD[44,45].

Even though the risk factors associated with CVD are well established, most individuals diagnosed with acute coronary events were not identified as being 'high risk'[45]. This observation has lead to the investigation of a number of clinical indicators of subclinical atherosclerosis. Coronary artery calcification (CAC) using computed tomography (CT) and B-mode ultrasound

measurement of carotid intima-media thickness (IMT) are noninvasive methods for estimating the plaque burden present in individuals that have not suffered acute cardiac events. These methods have found to be reliable independent predictors of cardiovascular disease events.

1.2.2 HAART-Associated Dyslipidemia and Cardiovascular Disease Risk

Infection with HIV-1 is often associated with metabolic changes that affect lipid levels, and the use of Highly Active Antiretroviral Therapy (HAART) can compound these effects. Despite their benefits in lowering plasma HIV levels, exposure to certain antiretroviral drugs can affect both lipid metabolism (dyslipidemia) and body fat distribution (lipodystrophy) in HIV+ patients[46-52]. These metabolic side effects, along with reports of higher cardiovascular disease (CVD) risk and myocardial infarction in HIV positive patients, have raised concerns of the potential dangers associated with HAART[53-58].

Very early on in the age of HAART use, studies began to report evidence of premature CAD associated with PIs[57]. Use of PIs is also associated with increased levels of TCHOL, LDL-C, and TRIG along with decreased HDL-C levels[46-50,59,60]. A recent review[61] outlined data from four studies comparing lipid levels in HIV-seronegative individuals exposed to PIs for short time periods to the levels seen in HIV-seropositive individuals receiving PI therapy[62-65]. These data showed that dyslipidemia associated with PIs is restricted to elevated TRIG levels and was seen only in response to one particular PI, ritonavir, in HIV-seronegative individuals. Also, when comparing HIV (-) to HIV (+) individuals an increased level of LDL-C was observed in the HIV (+) but not HIV (-) individuals after exposure to various PIs. This suggests that the increase in LDL-C levels associated with PI treatment in HIV (+) individuals is not due to the drug treatment but rather to a restoration to health. The other components of HAART – NRTIs and

NNRTIs – have also been associated with the development of dyslipidemia[52,61]. These lipid profiles would be considered atherogenic in the general population and may suggest an increased risk of atherosclerosis and CVD in patients receiving HAART. Often patients receive combinations of these drugs over varying periods of exposure making definitive conclusions about HAART-associated dyslipidemia and CVD risk very difficult. In addition, not every individual responds to these drugs in similar ways. For example, patients with the HLA-B*5701 allele have a hypersensitivity to treatment with the NRTI abacavir[66]. We hypothesize there is a similar genetic susceptibility to dyslipidemia in response to HAART.

There is evidence that HIV infection and HAART increases the risk of CVD but these studies often fail to correct for traditional risk factors of CVD (obesity, smoking, genetics, lipid levels)[58,67,68]. For instance, the D:A:D (Data Collection on Adverse Events of Anti-HIV Drugs) study initially reported a higher risk of CVD in HIV infected individuals and an association of antiretroviral therapy with risk of MI, but follow-up analyses revealed that this association was abrogated when adjusting for traditional risk factors[69,70]. Studies on indicators of subclinical coronary atherosclerosis (IMT and CAC) in HIV infected individuals have reported mixed results on the association of HIV infection and HAART with increased CVD risk[58,71-73].

1.2.3 Dyslipidemia and Cardiovascular Disease Genetics

The first indication that dyslipidemia was linked to CVD, and that genetics played a role, came from an investigation of people who suffer from homozygous familial hypercholesterolemia. This inherited condition is caused by a mutation in the *LDLR* gene, which encodes the LDL-C receptor and is characterized by increased levels of serum LDL-C along with substantial premature atherosclerosis. The LDLR protein is located in the liver and is responsible for

mediating the uptake of LDL-C and very-low density lipoprotein (VLDL) from the serum into the liver. The mutation that causes familial hypercholesterolemia results in a defective form of LDLR, which can no longer effectively transport LDL-C particles from the serum into the liver. This association between increased serum LDL-C levels and development of atherosclerosis was the first of its kind and led to the understanding that modulating LDL-C levels could impact the development of atherosclerosis, which ultimately resulted in the use of HMG-CoA reductase inhibitors (statins) for use in treatment of atherosclerosis and CVD. Statins inhibit HMG-CoA reductase, the rate-limiting enzyme during cholesterol biosynthesis, which lowers serum LDL-C levels.

CVD in the general population does not arise from just one mutation in one gene. Rather, CVD is a complex human disease that involves a number of metabolic processes involved in cholesterol metabolism and inflammation, as well as external environmental influences. Genetic linkage studies of inherited traits like familial hypercholesterolemia have shed light on a few of the processes, e.g. reverse cholesterol transport (RCT), involved in CVD but are limited in their scope to identifying only those genes/mutations that have a profoundly detrimental effect. Another way to investigate genes involved in complex human disease is to perform a genome-wide association study (GWAS). In general, a GWAS tests for thousands to millions of single nucleotide polymorphisms (SNPs) that are evenly spaced across the human genome under the hypothesis that a common trait is associated with a common [genetic] variant. When certain mutations are found to be associated with a disease trait, like increased levels of LDL-C, a genomic region or a particular gene is identified as a potential ‘candidate’ for further investigation. While GWAS provide a powerful tool in investigating the genetic components to human disease, they do have some fundamental flaws that need to be addressed. There is a

potential for increased Type I errors (false positive) because of the larger number of independent tests being performed. To address this, a stringent statistical correction (such as Bonferroni's correction) is applied or a False Discovery Rate (FDR) analysis is performed, but replication of positive associations across different studies remains an issue. Additionally, the SNPs identified by these studies often only explain a small portion of the trait variance (e.g. LDL-C levels) despite the large number of SNPs identified. This is either because the statistical corrections themselves are excessively conservative, over correcting the significance levels in an attempt to eliminate false positives, and/or the studies are underpowered to detect the effects of numerous SNPs that are individually contributing small effects of the trait of interest but collectively can explain a larger portion of the trait variance.

Another concern when performing genetic association studies is population stratification. As populations continued to migrate out of Africa and across the globe, their genetic pools became increasingly more homogeneous[74]. As time passed, these populations of limited genetic information diverged from each other forming divergent groups of what we today call race or ethnicity. When selecting a study population consisting of many races, or ancestral populations, there is a potential to increase the Type I error, especially in a case-control study design[75]. This is because of the fundamental design of these types of associations and statistical procedures. Often allele frequencies are compared between the case and control groups under the assumption that people in the case group will share a common genetic variant (e.g. the allele associated with the disease phenotype) and have a higher frequency of this variant compared to controls. If there is an unequal distribution of individuals who share genetic variants due to ancestry, and not disease status, then this will result in a false positive association result. So, it is important to accurately identify what ancestral populations are present in the study population (population substructure)

and if possible, to keep them separate when performing association studies. This can be accomplished by either using self-reported race (ethnicity) or by using genetic markers that have very different frequencies in different ancestral populations (ancestry informative markers, AIMs). Self-reported race is often used to stratify populations but this may not always be adequate, especially when dealing with populations of mixed genetic ancestry (e.g. African Americans)[76].

Over a decade of genetic association studies have investigated the genetic contributions to CVD and dyslipidemia with an ever-growing number of potential candidate genes and SNPs[77-97]. While the numbers of candidate genes continues to expand, there are recurrent biological processes that have been identified as being important to dyslipidemia, atherosclerosis, and CVD risk. The most common of these genes are those involved in cholesterol metabolism and reverse cholesterol transport. This is a complex process by which cholesterol is synthesized or absorbed then transported to the periphery and back to the liver where it is ultimately excreted. Briefly, the main players in this process are the transporters of cholesterol in the blood; LDL-C which predominately carries cholesterol to the peripheral tissues, and HDL-C, which predominately transports cholesterol back from these tissues to the liver for digestion or recycling. Many of the candidate genes identified to date encode proteins or enzymes directly related to this transport mechanism, e.g. *LDLR*, hepatic lipase (*LIPC*), apolipoprotein B (*APOB*), cholesteryl ester transfer protein (*CETP*), and lipoprotein lipase (*LPL*) to name a few.

1.2.4 Investigating the Genetic Contributions to HAART-Associated Dyslipidemia

Dyslipidemia is common in patients receiving HAART, yet not all of these patients suffer these side effects. This suggests there may be an underlying genetic contribution to HAART-

associated dyslipidemia. We hypothesized that this variation is due to the joint effect of HIV infection and HAART, together with the underlying genetic predisposition present in these individuals. To test this hypothesis we designed an Illumina GoldenGate genotyping array consisting of single nucleotide polymorphisms (SNPs) that had been previously associated with dyslipidemia and/or CVD in the general population[76-86,88-90,93-95,97-108]. We also included SNPs to provide effective coverage of the variation present in 75 candidate gene regions with similar associations in the general population[77-87,89-91,94-96,98-109]. Currently, there have been no studies investigating the possible differences in HAART-associated dyslipidemia by ethnicity. Since our study cohort is multiethnic, we included ancestry informative markers (AIMs) to accurately stratify our population based on genetically defined biogeographical ancestry (BGA). We used BGA to stratify our population because it has been shown to be superior to self-reported racial classifiers when investigating genetic variations associated with a disease phenotype[74].

2.0 STUDY PREMISE

2.1 HYPOTHESIS

Dyslipidemia side effects are often seen in HIV-positive individuals undergoing HAART, but the fact that these side effects are not seen in every such individual points to a host genetic contribution. Many human genes are already known to affect lipid metabolism and the susceptibility to cardiovascular disease, and so we hypothesize that variation in human genes involved in lipid metabolism will contribute to the development of HAART-associated dyslipidemia in HIV-positive individuals. Further, as both the incidence and type of cardiovascular disease varies between ethnic groups in the general population, we further hypothesize that the ethnic groups present in the MACS may vary in their response to HAART in ways that influence the development of dyslipidemia. We have developed a SNP genotyping array to test these hypotheses in the MACS.

2.1.1 Specific AIM #1: Population Stratification

Define the genetic population substructure of the MACS. The MACS is a multiethnic cohort, and studies of similar cohorts have shown that self-reported ethnicity does not always adequately describe a cohort's genetic composition. We have used a panel of Ancestry Informative Marker

SNPs to characterize the biogeographical ancestry of each MACS participant included in our study.

2.1.2 Specific AIM #2: Dyslipidemia Present in the MACS

Determine the extent of HAART-associated dyslipidemia in the MACS. We have determined the mean, pre-statin, values of the key biomarkers of dyslipidemia in the MACS, and stratified these by biogeographical ancestry to reveal variation between ethnic groups in their dyslipidemic response to HIV infection and HAART.

2.1.3 Specific AIM #3: Genetic Associations with Dyslipidemia

Identify key human genes that mediate the development of the dyslipidemic response to HAART in the MACS. We have performed both single-locus SNP and multi-locus haplotype analyses of MACS participants, stratified by lipid biomarkers and by biogeographical ancestry, and have identified specific lipid metabolism genes that significantly influence the development of HAART-associated dyslipidemia in the MACS. We have found that this genetic predisposition to the development of HAART-associated dyslipidemia varies between biogeographical ancestry groups in the MACS.

3.0 MATERIAL AND METHODS

3.1 STUDY DESIGN

Study participants were chosen from the Multicenter Aids Cohort Study (MACS) that had metabolic and lipid data recorded in the year 2005. Longitudinal data was obtained from time of enrollment to the last visit in 2009. The MACS is a prospective multiethnic cohort study of the natural history of HIV in homosexual and bisexual men in four major metropolitan areas; Los Angeles, CA; Chicago, IL; Baltimore, MD; and Pittsburgh, PA. Participants in the MACS are seen semi-annually providing us with detailed longitudinal lipid profiles from 1,945 HIV seropositive (n=955) and seronegative (n=950) individuals.

3.2 PHENOTYPES AND EXPOSURES

The mean LDL-C, HDL-C, TRIG, and TCHOL for patients were calculated according to their exposure group. Lipid levels used in this study were obtained solely from visits at which patients reported not receiving lipid-lowering medication. Three exposure groups can be defined based on HIV-1 infection status and highly active antiretroviral therapy (HAART) use; HIV and HAART negative controls (-/-), HIV positive but HAART naïve (+/-), and HIV positive receiving HAART (+/+). HAART use was defined in accordance with the DHHS/Kaiser Panel

guidelines[110]. For the HIV/HAART (+/-) patients (reported as not receiving HAART as of the last visit in 2009), the lipid phenotype values were used from visits after they had become HIV positive. For HIV/HAART (+/+) patients the only lipid phenotype values used were those obtained after the patient had been put on HAART. Only blood samples from fasting patients were used for the evaluation of serum lipids and the LDL-C was calculated using the Friedewald equation [LDL-C = TC - (HDL-C + (TRIG/5)] or by direct measurement if the TRIG level was above 400mg/dl. Additional covariates were obtained from CAMACS and are listed in Table 2.

Table 2 - MACS Covariates

MACS Covariates	
Self Reported Race	BMI
Date of Birth	Glucose
HIV sero-status	Hours fasting before blood draw
Last seen seronegative	Total cholesterol
First seen seropositive	High-density lipoprotein
AIDS diagnosis (i.e. Karposi's Sarcoma)	Low-density lipoprotein
Last reported ART - free	Triglycerides
First reported ART	Insulin
First seen with AIDS	Apolipoprotein A
Visit number	Apolipoprotein B
Visit Date	Date-Type of cholesterol/lipid lowering drugs
Diagnosed with angina/chest pain	
Relatives diagnosed with heart attack	
Relatives diagnosed with congestive heart failure or CHF	
Relatives diagnosed with stroke	

3.3 SAMPLE PROCESSING

Peripheral blood mononuclear cell (PBMC) pellets were obtained from the MACS and genomic DNA was extracted using the Qiagen QIAmp DNA Blood Mini Kit using the Blood or Body

Fluid Spin Protocol. [Qiagen Inc., Valencia, CA] The QIAmp DNA Blood Mini Kit is a silica-gel membrane based extraction where DNA binds to the membrane while contaminants are eluted with successive wash steps. The extracted genomic DNA was then stored at 4°C at a total volume of 400 µl. To obtain the DNA quantities needed for the Illumina GoldenGate assay a whole genome amplification (WGA) was performed on all of the samples using the Illustra Genomphi V2 DNA Amplification Kit. [GE Healthcare Biosciences, Pittsburgh, PA] The Genomphi V2 DNA Amplification Kit yields 4-7 µg per 20 µl reaction. Extracted genomic DNA was first denatured at 95°C for three minutes then cooled to 4°C. The kit's master mix, which contains random hexamer primers, dNTPs, and Phi29 DNA polymerase, was added and then incubated at 30°C for 2 hours. The WGA DNA was quantified using Quant-iT PicoGreen dsDNA reagent kit [Invitrogen Ltd., Paisley, Carlsbad, CA] and diluted to a concentration of 50 ng/ml for use in the genotyping assay. PicoGreen is a sensitive fluorescent nucleic acid stain that binds to double-stranded DNA allowing for quantification of a wide range of DNA concentrations[111].

3.4 GENOTYPING AND ASSAY DESIGN

Initially, PCR based Fluorescent Polarization (FP) was used to genotype our samples. This procedure only allows for the detection of one particular single nucleotide polymorphism (SNP) at a time. Instead we chose to use the Illumina GoldenGate technology for the genotyping of our samples. [Illumina Inc., San Diego, CA] This allowed us to construct a custom array of 1,536 SNPs covering selected genes and regions of interest previously found to be associated with lipid metabolism and cardiovascular disease in the general population.

3.4.1 Fluorescent Polarization Genotyping

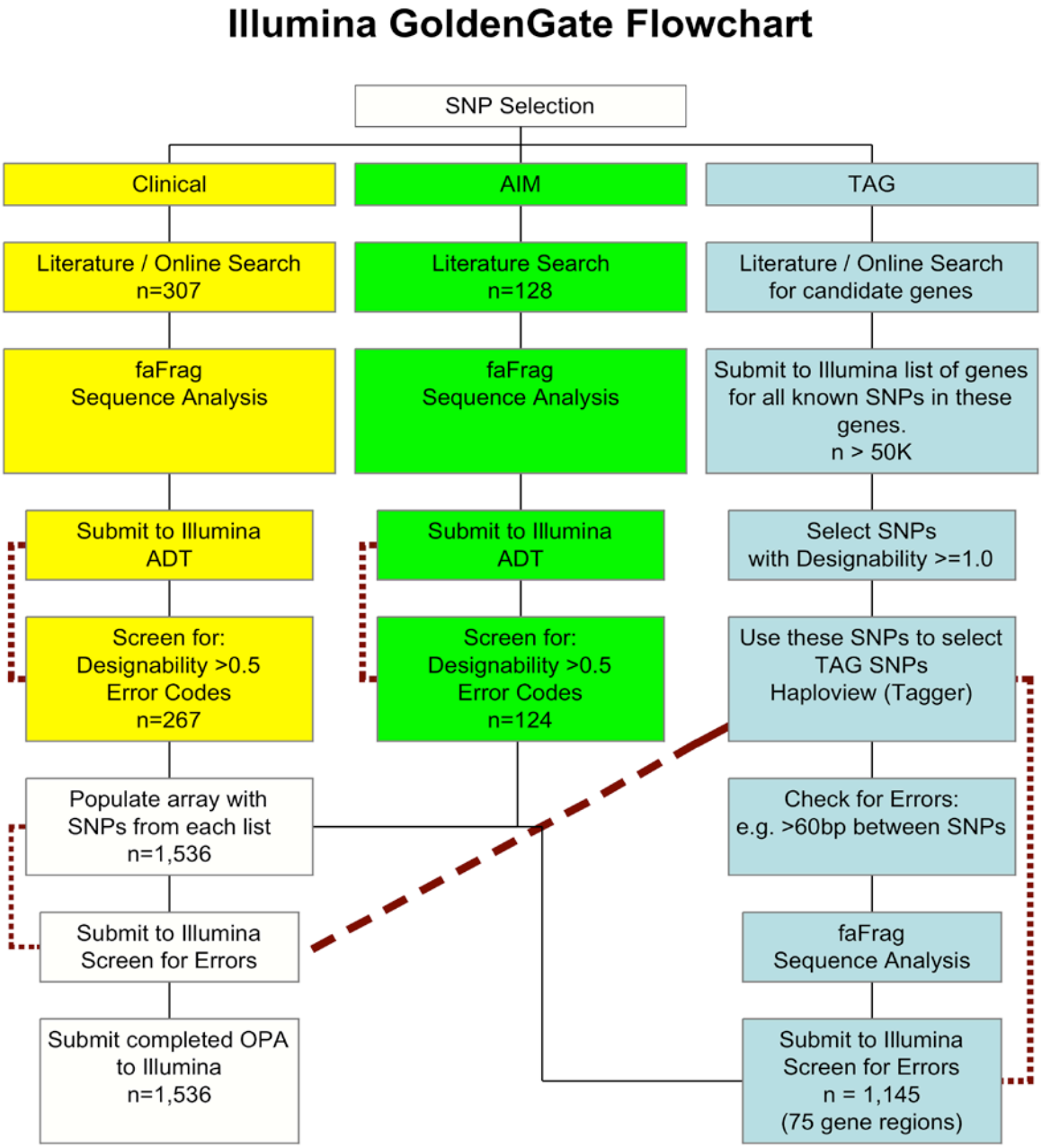
Preliminary genotyping of four SNPs associated with dyslipidemia in the general population was performed on genomic DNA extracted from a subset of study participants. The PCR based Fluorescent Polarization (FP) assay was used to genotype these samples[112,113]. First a PCR reaction was performed on the genomic DNA to amplify the region around the SNP of interest. Then a PCR clean up (Exonuclease I/Shrimp Alkaline Phosphatase) was performed to remove any single stranded DNA molecules and excess primers. Following the PCR clean up a single base extension was performed using allele-specific dye-terminators. Two fluorophore dyes were used; Tamra and R110. The FP assay is based on the principle that when a fluorescent molecule is excited with plane-polarized light, it emits light in a fixed plane. The amount of light emitted and the extent at which it is polarized or depolarized is directly related to the molecular volume, holding temperature and viscosity of the media constant. This is due to the difference in molecular rotation of different sized molecules. A larger molecule will rotate and tumble less than a smaller molecule and thus maintain the polarization of the light emitted. The samples are then plotted with one dye being on the x-axis and the other on the y-axis. Samples that are homozygous for the SNP alleles have high FP values for one dye but not the other and heterozygous samples have high FP values for both dyes. Using these plots, genotype clusters can be defined and each samples genotype inferred.

3.4.2 SNP Selection

An extensive literature search was performed to identify SNPs that had been previously associated with dyslipidemia and/or cardiovascular disease (CVD) in the general population[75-

[95,99-107]. We identified 307 clinically relevant SNPs and approximately 90 genes and gene regions that had been associated with dyslipidemia and CVD. In the design phase, SNPs were selected based on the highest likelihood of success using Illumina's design scores and validation status. The overall process of designing our custom array is outlined in Table 3. The list of 1,536 SNPs

Table 3 - GoldenGate Design Flowchart



on our array can be divided into three groups; clinically relevant SNPs identified from our preliminary literature search, ancestry informative markers (AIMs) used for identifying biogeographical ancestry (BGA) and Tag SNPs that capture the allelic variation across the candidate genes previously identified using the linkage disequilibrium (LD) present in these regions.

The process of SNP selection was similar for the clinically relevant SNPs and AIMs groups. Each SNP's probe sequence (+/- 60 bps from the SNP) was first analyzed using faFrag, one of the Kent Source Utilities available at the University of California Santa Cruz Genome Browser [http://genome.ucsc.edu], which allows for the identification of polymorphic sites adjacent to the SNP of interest. This was important because if the patient's sample being tested also had mutations at these sites the binding efficacy of the probe may be inhibited. We were able to annotate these SNPs in the probe sequence using degenerate nucleotide codes, which allowed Illumina to construct probes that accounted for the additional SNPs. After sequence analysis was performed the probe sequences for each SNP were submitted to Illumina's online assay design tool (ADT). The ADT software uses a proprietary design algorithm to assign minor-allele frequencies (MAF), validation status, error codes, SNP scores and designability ranks. The SNP score and designability rank were used as the primary quality metrics. To make sure as many clinically relevant and AIM SNPs were included on the array a designability rank cutoff of ≥ 0.5 was used for these probes. The Illumina designability rank is a number ranging from 0 to 1 that summarizes the estimated viability of an assay. Illumina also uses a rank score of 1.1 to denote assays that have previously been validated as successful. Due to the length of the primers used in the Illumina GoldenGate assay, no SNP on the array could be within 60 bps of another

SNP included on the array. We used an iterative process of adding and removing SNPs based on this 60 bp criteria to maximize the number of SNPs included on the array.

A different selection process was used for the Tag SNP group. After identifying the candidate gene regions from the literature search, a list of 90 genes was submitted to the Illumina ADT software. Each candidate gene region was defined to include 10kb upstream and downstream of the 5' and 3' UTRs, respectively. The 10kb extension allows us to test the variation present in the promoter and other regulatory regions adjacent to the candidate gene. (Table 4) The ADT software, using human Genome Build 36 and dbSNP Version 126, returned a list of over 50k Illumina SNPs known and validated in these gene regions. The general selection scheme is outlined in Table 3. These SNPs were then used as a reference panel to create a set of Tagging SNPs for each candidate gene region using the Tagger algorithm in Haploview[114]. Only SNPs from the Illumina list that had a designability rank = 1.0 were used in the reference panel. The selection of Tag SNPs based on the LD present in each gene allows us to genotype a minimal number of SNPs and still provides information about a larger number of nearby variants that were not genotyped. The tagging strategy was as follows: Haploview was set to version [2], release [22], and analytical panel CEU (Utah residents with ancestry from northern and western Europe); a pair-wise $r^2 \geq 0.8$ was used to identify the Tag SNPs and a minimum distance of 60 bps was set between each SNP. Even though the MACS is a multi-ethnic cohort we chose to use the HapMap CEU reference population to design the Tag SNP panel because the LD structure is better defined in the CEU population compared to the other admixed populations, e.g. the ASW (African ancestry in southwest USA) HapMap population, and the majority of our samples were predicted to be primarily of European ancestry similar to the CEU reference population. A pair-wise r^2 threshold of ≥ 0.8 was used because it gives the minimum number of tags required while

Table 4 - Gene List

Candidate Genes and Gene Regions		
ABCA1	F2	MYD88/ACAA1
ABCG5/8	F5	NCAN
ACE	FADS1/2/3	NOS3
AGT	FBLN5	NR1H2
AGTR/FGB	FOLH1	NR1H3/MADD
ALOX15	GALNT2	PBX4
ALOX5	GCKR	PCSK9
ALOX5AP	GCNT4/HMGCR/POLK	PLTP
APOA1/A4/A5/APOC3/BUD13/ZNF259	ICAM-1	PON1
APOB	IL10	PTGES
APOE/TOMM40/APOC1/C2/C4	IL6	PTGIS
ATG4C	iNOS(NOS2)	PTGS1
BAZ1B/BCL7B/TBL2/MLXIPL	ITGA2	PTGS2
CBS	LCAT	RETN
CCL2	LDLR	SCARB1
CCR7	LIPC	SELP
CD14	LIPG	SELPF5
CD36	LPA	SERPINE1
CDKN2A	LPL	SQLE
CELSR2/MYBPHL/PSRC1/SORT1	LTA/TNF	SREBF2
CETP	MMP3	TLR4
CILP2	MTHFR	TRIB1
CRP	MTR	USF1
DOCK7/ANGPTL3	MTRR	VCAM1
EFEMP1	MVK/MMAB	ZNF217

maintaining almost complete power relative to that of the reference panel[115]. After the lists of Tag SNPs for each gene were generated the sequences were analyzed using faFrag as described above. Any additional mutations in the +/- 60 bp window around the SNP were replaced with the degenerate nucleotide code and resubmitted to the ADT software for validation. The Tag SNPs were then checked for error codes, SNP scores and designability ranks. SNPs whose designability rank had dropped below 1.0 were removed and the Tagger algorithm was rerun to adjust for the loss of these SNPs. The sequence analysis was repeated and the SNP list was

resubmitted to ADT. The primary goal during this process was to maximize the allelic coverage of each gene region by using the minimum number of SNPs with the highest likelihood of success on the GoldenGate array.

All three SNP groups were combined to complete the final 1,536 SNP panel and then were submitted to Illumina's ADT for validation. If there were any further errors identified by the ADT software the processes for each SNP group, reviewed above, were repeated. Mainly this consisted of re-Tagging certain gene regions to account for the addition of the clinically relevant SNPs. In some cases we were able to identify additional 'tag' SNPs in high LD with some of the clinically relevant SNPs that were unable to be included on the array.

3.4.3 GoldenGate Genotyping

The Illumina GoldenGate Assay was used to genotype 1,914 samples from the MACS. The GoldenGate technology allows for the design of a custom panel of 1,536 SNPs. WGA DNA samples were submitted to the Genomics and Proteomics Core Laboratories (GPCL) at the University of Pittsburgh in 96 well plates. Each plate contained 4 randomly selected known samples for replicate analysis and 4 population controls consisting of one family of both parents and two children, used to control for Mendelian consistency of the genotype data.

The first step in the Illumina assay is to activate the genomic DNA for binding to paramagnetic particles. The next step involves the primer hybridization of the assay oligonucleotides to the genomic DNA bound to these particles. Illumina designs three oligonucleotides for each SNP. Two allele-specific oligos (ASO) bind directly upstream of the SNP and one locus-specific oligo (LSO) that binds farther downstream of the SNP. Following the hybridization step a product is formed that contains a single ASO and a LSO; this acts as the

template for the next step in the assay. A PCR reaction is performed on this template, where dye-labeled primers specific to the ASO are added to the reaction. The PCR product contains allele-specific dyes to the SNP and a unique address sequence produced from the LSO. The address sequence targets the single stranded product to the bead specific to that SNP located on the array. The single-stranded dye-labeled DNAs hybridize to their complement bead located on the array matrix. We received raw fluorescent data files and raw image files that were analyzed using the Illumina BeadStudio software package. Genotypes were called using the standard clustering algorithm in BeadStudio. Each SNP's cluster graph was examined individually and genotype calls were adjusted according to visual inspection.

3.5 QUALITY CONTROL

3.5.1 Samples

Samples were excluded if the replication error rate of the repeated samples was > 10% or if the sample call rate was < 90%. We also used the 10% GenCall (GC) score, which is the 10^{th} percentile of GC scores in the given distribution of sample GC scores. The GC score is a proprietary Illumina quality control metric that ranges from 0 to 1. GenCall scores are calculated for each genotype and generally decrease the further a sample is from the center of cluster it is associated with. Illumina suggests an initial GC cut-off of 0.15; we used a cut-off of 0.25 to screen our genotypes. Through graphical examination we excluded any samples below a threshold of 0.40 10% GC because samples below this threshold were associated with low sample call rates and were identified as low quality samples. (Figure 4A) Manual examination of

sample graphs was also performed in the BeadStudio with samples having abnormal distributions being excluded. These abnormal distributions were the result of improper hybridization and the samples were repeated.

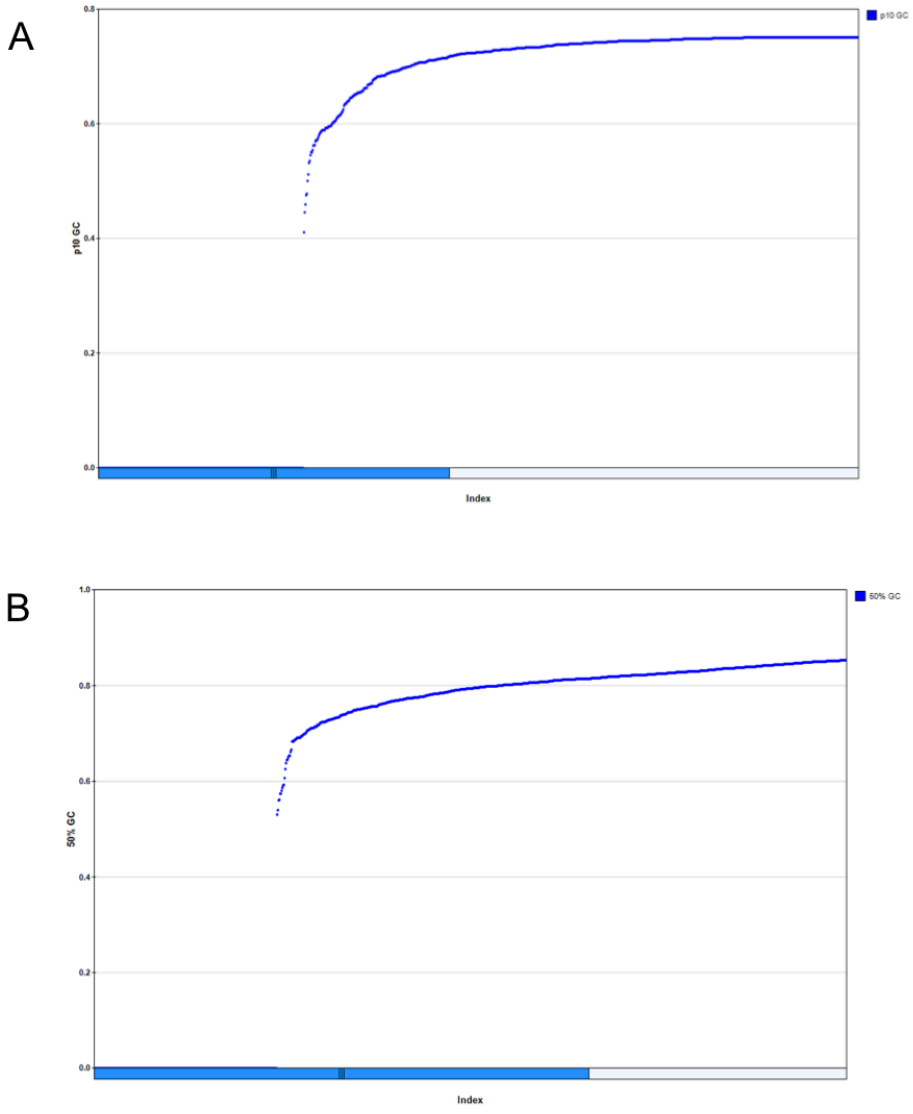


Figure 4 - GenCall Distributions

(A) 10% GC scores for each sample tested on the GoldenGate array. Index represents the ordered sample 10% GC scores. (B) 50% GC scores for each SNP tested on the GoldenGate array. Index represents the ordered SNP 50% GC scores.

3.5.2 SNPs

Loci were excluded if the minor allele frequency (MAF) was < 0.01 in our data set and if the locus call rate was < 80%. When performing the association analysis a MAF of 0.05 was used as a cutoff and was specific to the subpopulation being tested. We used the BeadStudio quality metric 50% GC, 50^{th} percentile of locus GC scores, to exclude loci with a score < 0.50 by graphical examination. (Figure 4B) Manual examination of SNP graphs was performed in BeadStudio and loci exhibiting evidence of copy number variation (CNV), having more than three distinct cluster or abnormal clustering, were removed. Figure 5 shows examples of SNP genotype graphs that shows evidence of possible CNV (A) compared to a SNP that shows good genotype clustering (B).

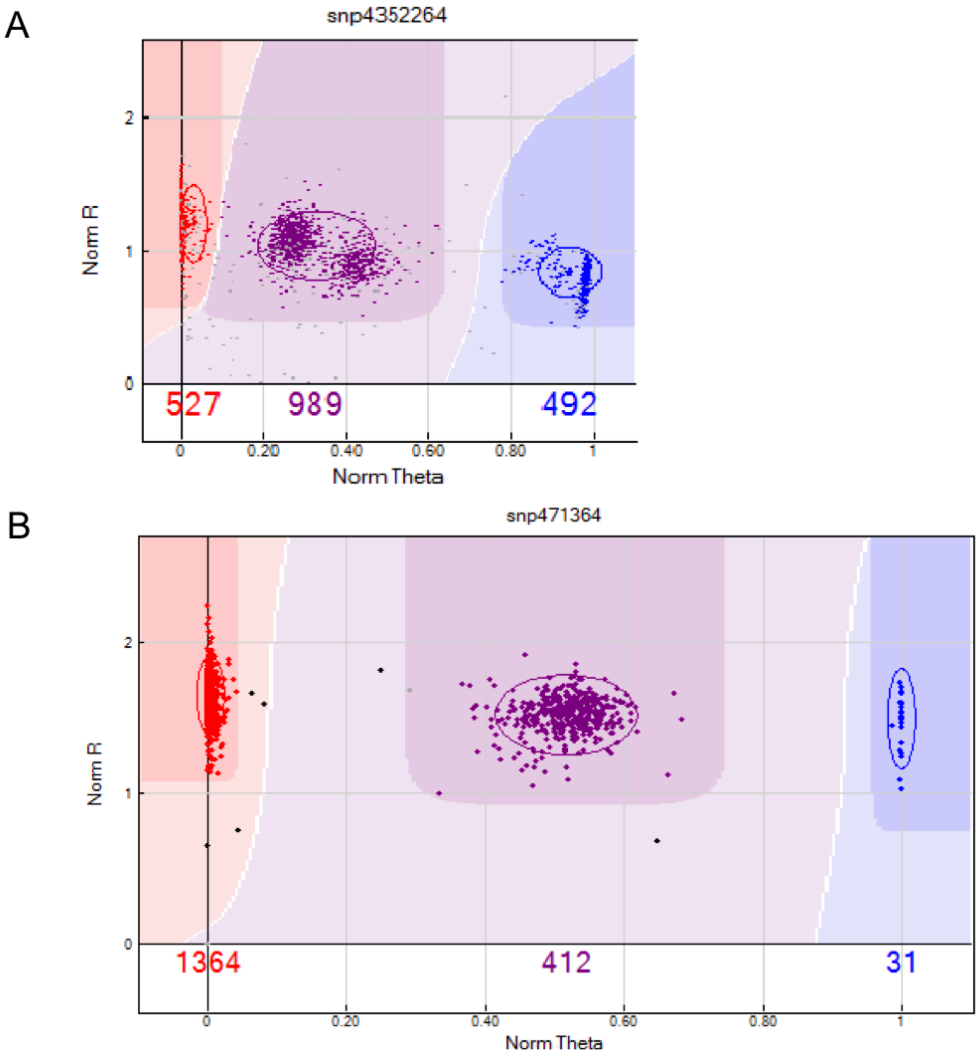


Figure 5 - Example SNP Clustering Plots

3.6 GENETIC ANALYSIS

3.6.1 Population Stratification

Using the 113 AIMs that passed our QC process we performed a multidimensional scaling (MDS) procedure to cluster individuals based on pairwise identity-by-state (IBS) distances. This

was done using the Multidimensional Scaling Population Stratification (MDS) procedure in PLINK[116]. We included reference populations downloaded from the HapMap database to identify the biogeographical ancestral origin of each cluster in our population. (Figure 6) In order to define clear boundaries between the biogeographical ancestral clusters, we performed a k-means clustering algorithm on the first two components of the MDS procedure. We specified 3 possible cluster positions in the k-means procedure. We used these cluster positions to define two main groups, European ancestry (EA) and African/European ancestry (AEA), which were used for our analyses. The k-means procedure was performed in the R software package[117].

3.6.2 Quantitative Trait Loci (QTL) Analysis

For single locus association the mean lipid phenotypes were regressed on each SNP using multivariate linear regression. Each SNP was tested using dummy variables (0,1,2) that represented the dosage of the minor allele assuming an additive genetic model. The first two components of the MDS procedure were included as covariates to control for population stratification. For the multi locus association tests haplotypes were constructed in PLINK using the haplotype blocks estimation procedure, which is based on pairwise linkage disequilibrium (LD). We then used these haplotypes in a linear regression test for association with mean lipid levels using the Haplotype-Based Association Test with the GLM procedure in PLINK.

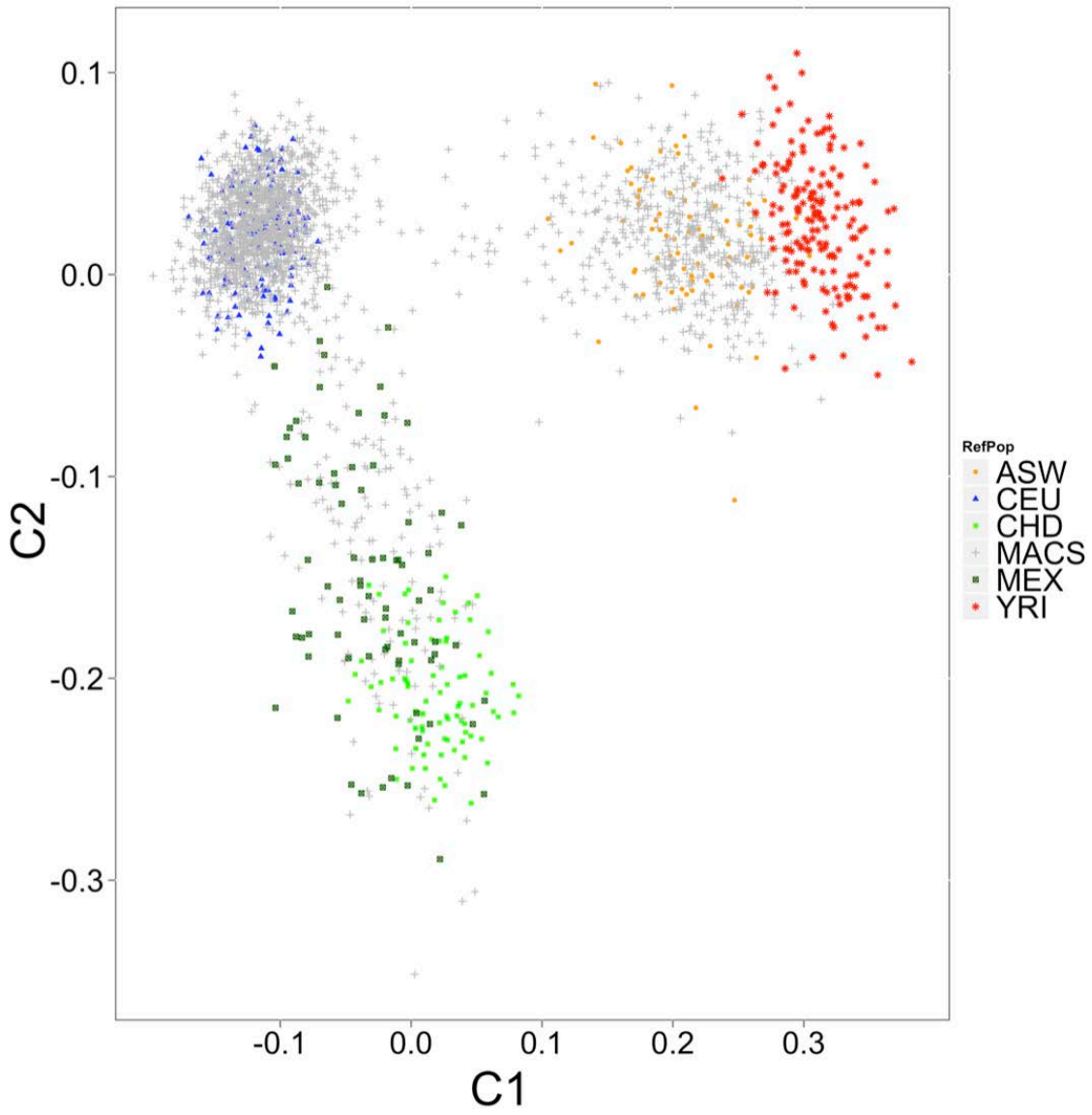


Figure 6 - HapMap Reference Population Clustering

Biogeographical ancestry in the MACS compared to HapMap reference populations. Scatter plot of the first two components from the MDS on the combined MACS and HapMap samples. Colors represent individuals from MACS or HapMap reference populations. Individuals from the MACS study are shown in grey. Individuals taken from the HapMap: African ancestry in the southwest USA (ASE) are shown in orange; Utah residents with Northern and Western European ancestry (CEU) from the CEPH collection are shown in blue; Chinese in metropolitan Denver, Colorado (CHD) are shown in light green; Mexican ancestry in Los Angles, California (MEX) are shown in dark green; and Yoruba (YRI) in Ibadan, Nigeria are shown in red

3.7 STATISTICAL ANALYSIS

Empirical p-values were calculated using both adaptive (aPerm) and max(T) (mPerm) permutation procedures in PLINK. Empirical p-values were calculated using $1x10^6$ permutations for the single locus association and $1x10^4$ permutations for the haplotype based association. Briefly, the aPerm is a test of each SNP's asymptotic p-value against the permutated p-value distribution. The mPerm is a test of each SNP's asymptotic p-value against the distribution of the maximum permutated p-values from each permutation. The Bonferroni correction for multiple independent tests was performed using PLINK. The genomic inflation factor was obtained from the standard PLINK output when running association tests. QQ-plots of each association's p-values were calculated and plotted in R.

$$Mean\ Lipid\ Phenotype = \beta_1 s + \beta_2 c_1 + \beta_3 c_2 + \beta_4 h + \beta_5 s*h$$

Each SNP, *s*, was tested with the first two components of the MDS, *c1 + c2*, to control for population stratification. The patient's exposure group is represented by *h*, and the interaction between each SNP and exposure group is represented by the final term *s*h*. This term tests if there is an interaction of HIV infection plus HAART treatment on each SNP's effect on mean lipid levels compared to the HIV/HAART (-/-) group. Classical covariates associated with lipid levels and CVD were not included in our model because it has been shown that a reduced model is best suited to detect genetic effects[118]. Analysis of Variance (ANOVA) was performed on mean lipid levels by biogeographical ancestry and HIV/HAART status, including an interaction term for BGA and HIV/HAART. This was performed using the [aov] procedure in R. Pairwise comparison of the means for all groups was carried out using the Tukey's Honest Significant Difference (TukeyHSD) procedure in R. The TukeyHSD test was used because it is a single-step procedure that compares all possible pairwise comparisons and controls for these multiple-

comparisons. It also works well in conjunction with the ANOVA model and handles unequal samples sizes better than some other statistical tests. It should be noted that the sample sizes for each group were unequal and this may affect the parameter estimates of the ANOVA and TukeyHSD.

4.0 RESULTS

4.1 PRELIMINARY GENOTYPE DATA AND QUALITY CONTROL

4.1.1 Preliminary Genotyping Using Fluorescence Polarization

Initially we had intended to use FP as the primary genotyping methodology for our study. As genomic DNA was extracted from the PBMC cell pellets we performed FP genotyping on a subset of samples for a small number of polymorphisms that had previously been shown to be associated with RCT[91]. The summary of the genotype and allele frequencies obtained from these experiments is shown in Table 5. Genotypes were called by plotting the degree of polarization (mP value) for the dyes against each other and defining clusters of sample that had high mP values for one or both of the dyes. (Figure 7) These experiments showed we could successfully extract genomic DNA from the PBMC cell pellets and that these DNAs were viable for PCR based genotyping.

Using FP is a cost effective and efficient methodology when investigating a limited number of SNPs in a small number of individuals. Our initial hypothesis focused on the possible genetic susceptibilities to HAART-associated dyslipidemia present in genes related to RCT based on the work of Morabia et al., which associated polymorphisms in 11 genes involved in

RCT with extreme serum lipid levels[91]. Even though it has been shown that polymorphisms in

Figure 7 - Genotype Clustering (FP) rs2294313

RCT genes play an important role in determining serum lipid levels, there remains a large number of genes associated with dyslipidemia and CVD that we were unable to investigate efficiently using FP genotyping. Given; there are approximately 80-100 genes that have been associated with dyslipidemia and CVD, our sample population of 1,945 individuals, and our ability to produce 1,000 genotypes a week using FP, it would have taken just over 57 years to effectively test all the known SNPs and genes associated with serum lipids and CVD. Fortunately, technology exists that can test thousands to millions of SNPs at the same time on a large number of samples. For further genotyping and association studies, we transitioned our

study to a high-throughput genotyping platform utilizing a custom designed Illumina GoldenGate genotyping assay.

Table 5 - Fluorescent Polarization Genotyping Summary

Summary Data from Fluorescent Polarization Genotyping

SNP	Genotype Frequencies			Allele Frequencies	
rs2294213 (PLTP 1b+26)	GG	GC	CC	p(G)	p(C)
n=640	(0.85)	(0.14)	(0.01)	(0.92)	(0.08)
rs2777801 (ABCA1 32b+30)	AA	AC	CC	p(A)	p(C)
n=442	(0.72)	(0.24)	(0.03)	(0.85)	(0.15)
rs328 (LPL S447X)	GG	GC	CC	p(G)	p(C)
n=459	(0.83)	(0.16)	(0.02)	(0.92)	(0.08)
rs5888 (SR-B1 A350A)	GG	GA	AA	p(G)	p(A)
n=181	(0.43)	(0.42)	(0.16)	(0.64)	(0.36)

4.1.2 Quality Control

From the total 1,945 samples, 1,914 passed our quality control (QC) filters, described in methods, yielding an overall sample success rate of 98% with a mean sample call rate of 95%. All of the 108 replicate samples displaying proper hybridization had duplicate concordance rates of >90%. A total of 1,281 loci passed our QC filters giving an overall locus success rate of 83% with a mean locus call rate of 99%. Specifically, 214 of the 267 clinical SNPs (80%), 954 of the 1,145 tag SNPs (83%) and 113 of the 124 AIMs (91%) passed QC procedures. In order to maximize the number of previously published clinically relevant SNPs present on our array we accepted lower designability rank during the design stage, accepting a score as low as 0.5 in some cases for these probes. This leniency accounts for the higher failure rate in the clinical SNP

group. Because of our diligence during the GoldenGate design phase, we have achieved a high rate of success for the loci being tested and produced a high quality set of data from which to perform the genetic association studies.

4.2 POPULATION STRATIFICATION

Population stratification introduces bias in genetic association studies increasing the probability of both Type I and Type II errors[119]. The MACS is a multi-ethnic cohort study, thus the population substructure is important to define before we conduct our analyses. We included 124 ancestral informative markers (AIMs) in our array to aid us in defining the ancestral populations present in the MACS. These are SNPs that have been shown to provide an accurate stratification of the component biogeographical ancestral source groups in an admixed population[120]. After genotyping our samples 113 AIMs passed our quality control metrics and were used in the substructure analysis. Pairwise identity-by-state (IBS) distances for the 1,914 samples that successfully genotyped were calculated using these AIMs. Then multi-dimensional scaling analysis (MDS) was performed using the IBS distances.

We first looked at the population structure using self-reported race categories which was suggested by Driver et al.,[76]. The study enrollment questionnaire used by the MACS has eight possible racial categories: White-non-Hispanic, White-Hispanic, Black-non-Hispanic, Black-Hispanic, American Indian or Alaskan Native, Asian or Pacific Islander, Other, and Other-Hispanic. We restricted our analysis to the two largest groups, the White-non-Hispanic (n=1,260) and the Black-non-Hispanic (n=485), which together account for 1,745 of the 1,945 samples originally studied. We stratified the self-reported racial groups into three subgroups to simplify

the visualization: White-non-Hispanic (WnH); Black-non-Hispanic (BnH); and Other (Oth), accounting for a further 200 individuals in all. The first two components of the MDS are shown in Figure 8A and show good stratification of our population with what appears to be three primary ancestral clusters. To clarify the ancestral population composition of each cluster we used the genotype data that are available for these AIMs in five reference populations from the HapMap database (Figure 6). Using these reference populations we confirmed the WnH cluster as co-clustering primarily with individuals of European descent, the BnH cluster as co-clustering with individuals of mixed Sub-Saharan African and European descent, and the Oth cluster as co-clustering with individuals of mixed European and Asian descent. A number of individuals give a self-reported race that does not agree with their biogeographical ancestry (Table 6) and this misclassification is a potential problem in our association studies. To resolve this, we performed a K-means clustering procedure on the first two components of the MDS. This allowed us to define three biogeographical ancestral groups based on their genetic relatedness. The results from these procedures are shown in Figure 8B and summarized in Table 6. We will use these genetically defined biogeographical ancestral groups for all our following analyses, focusing on individuals of European (EA) and African/European (AEA) descent.

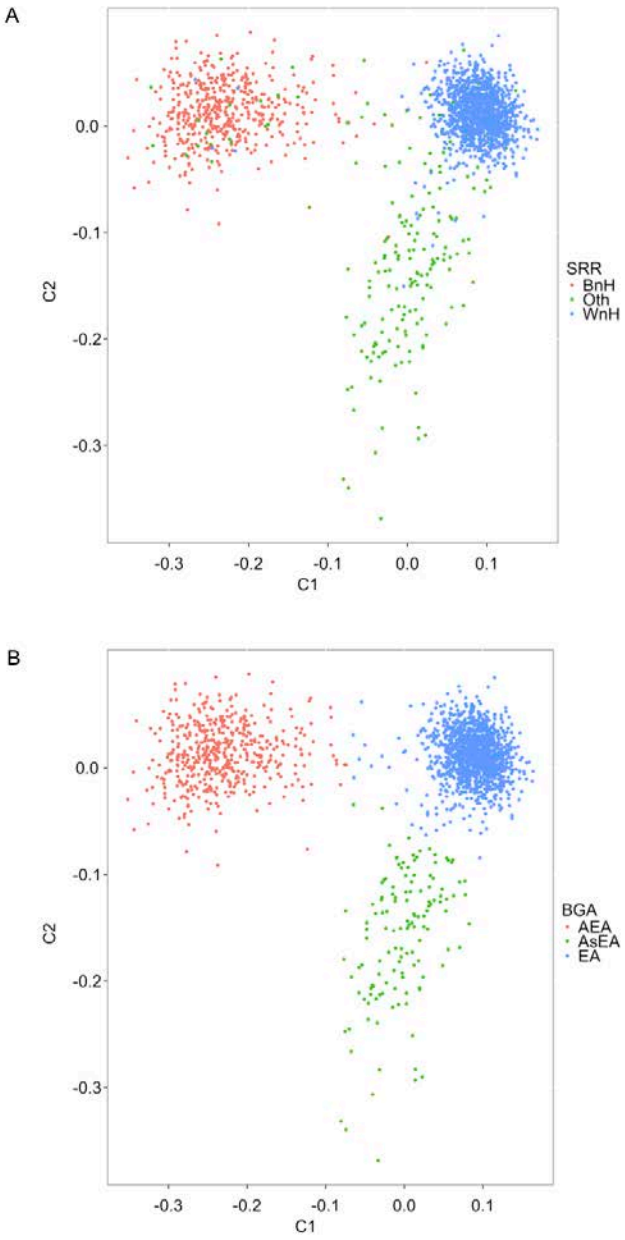


Figure 8 - Population Substructure in the MACS

Population sub-structure in the MACS. **(A)** Scatter plot of the first two components from the MDS procedure with dots representing individuals from the MACS. Colors represent self-reported race (SRR). Individuals who self-report as Black-non-Hispanic (BnH) are shown in red, individuals who self-report as White-non-Hispanic (WnH) are shown in blue and individuals who self-report as one of the White-Hispanic, Black-Hispanic, American Indian or Alaskan Native, Asian or Pacific Islander, Other, or Other-Hispanic (Oth) are shown in green. **(B)** Scatter plot of the first two components from the MDS procedure with dots representing individuals in our study population from the MACS. Here, colors represent defined biogeographical ancestry (BGA) after k-means clustering analysis was performed on the first two components of the MDS. Individuals with an ancestry derived from populations of African and European descent (AEA) are shown in red. Individuals with primarily European ancestry (EA) are shown in blue, and individuals with a complex ancestry derived from Asian and European descent (AsEA) are shown in green.

Table 6 - MACS Demographic Data

MACS Demographic Data

				Mean (Range)				
	HIV/HAART Status (N)			(-/-)		(+/+)		
Self Reported Race	(-/-)	(+/-)	(+/+)	Age	BMI	Age	BMI	Yrs on HAART
White-non-Hispanic	682	35	543	51.6 (20.1 - 84.3)	26.9 (17.0 - 56.6)	50.1 (24.4 - 76.8)	25.0 (15.6 - 41.3)	10.6 (0.3 - 14.6)
Black-non-Hispanic	200	37	248	44.8 (22.9 - 71.2)	28.1 (18.7 - 61.3)	44.1 (22.2 - 70.1)	25.9 (16.8 - 52.3)	9.1 (0.3 - 14.4)
Other	81	8	111	43.5 (19.7 - 70.2)	27.1 (17.1 - 46.6)	40.3 (21.7 - 66.8)	25.7 (16.5 - 43.5)	9.4 (0.3 - 13.9)
Total	963	80	902					
Biogeographical Ancestry								
European	713	38	540	51.4 (19.7 - 84.3)	27.0 (17.0 - 61.3)	49.8 (24.1 - 76.8)	25.0 (15.6 - 41.3)	10.5 (0.3 - 14.6)
African/European	194	36	258	44.5 (22.3 - 71.2)	28.1 (18.7 - 46.2)	44.1 (22.1 - 70.1)	25.8 (16.8 - 52.3)	9.2 (0.4 - 14.4)
Asian/European	46	6	83	42.0 (21.6 - 74.0)	26.4 (20.5 - 35.6)	39.4 (21.7 - 76.8)	25.7 (16.5 - 36.5)	9.1 (0.3 - 13.9)
Total	953	80	881					

4.3 SERUM LIPIDS IN HIV/HAART PATIENTS DIFFER BY BIOGEOGRAPHICAL ANCESTRY

To assess the extent of dyslipidemia in our cohort we investigated four serum lipid phenotypes: TCHOL, LDL-C, HDL-C, and TRIG. We were able to use longitudinal lipid phenotype data obtained for every individual enrolled in our study to define a physiological baseline for use in our genetic association studies. To do this, we took means of every individual's visit excluding those at which patients reported taking cholesterol-lowering medications, such as statins. Using the mean value from multiple visits for a particular analyte has been shown to be accurate and in some cases superior to a single measurement[121]. We will refer to these mean lipid levels throughout the paper as mTCHOL, mLDL-C, mHDL-C, and mTRIG.

Three possible treatment groups are present in each biogeographical ancestral population: HIV negative / HAART negative controls (-/-), HIV positive / HAART naïve individuals (+/-),

and HIV positive patients receiving HAART (+/+). The means of each group were calculated

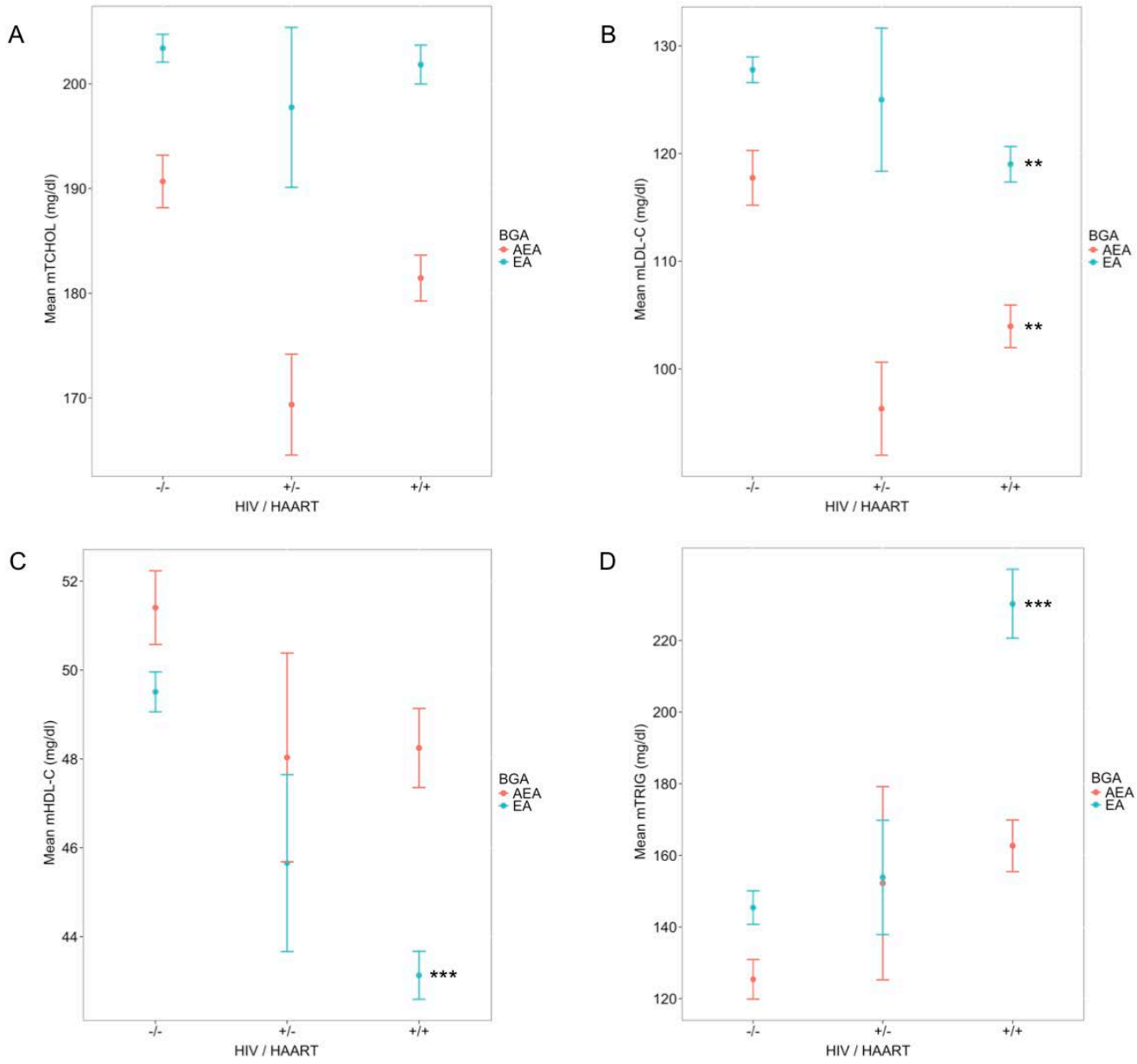


Figure 9 - Mean Serum Lipids Vary by Biogeographical Ancestry and HIV/HAART Status

Mean serum lipids vary by biogeographical ancestry (BGA) and HIV/HAART status. (A) Mean mTCHOL by HIV/HAART status. (B) Mean mLDL-C by HIV/HAART status. (C) Mean mHDL-C by HIV/HAART status. (D) Mean mTRIG by HIV/HAART status. The means of the European ancestral (EA) group are shown in blue and the means of the African/European ancestral group (AEA) are shown in red. Bars indicate standard error of the mean. Asterisks denote a significant pairwise comparison (TukeyHSD) of the mean serum lipid level of HIV/HAART (+/+) to HIV/HAART (-/-) within an ancestral group. [** = < 5.00 $x10^{-4}$, *** = < 5.00 $x10^{-7}$]

and plotted by HIV/HAART status and biogeographical ancestry (Figure 9). We performed an ANOVA as a formal test of the effects of biogeographical ancestry and HIV/HAART status. Additionally pair-wise comparisons of groups were performed using Tukey's Honest Significant Difference test (TukeyHSD). (Table 7) We have indicated the p-values for the comparison of HIV/HAART (+/+) to HIV/HAART (-/-) groups in Table 7, but all possible pair-wise comparisons were performed. We have concentrated our analysis on the HIV/HAART (-/-) and (+/+) groups for two reasons. Firstly, the HIV positive yet HAART naïve group is composed of a small number of individuals in comparison to the (-/-) and (+/+) groups, and secondly because they represent a unique population. They are either very recently infected with HIV, or have genetic features that allow for the control of HIV infection, and these features themselves may influence their dyslipidemic response to HIV infection.

Table 7 - Lipid Phenotypes by Biogeographical Ancestry and HIV/HAART Status

Lipid Phenotypes by Biogeographical Ancestry and HIV/HAART status

	Mean - mg/dl (SD)				*Pairwise Comparisons of the Mean*						
	EA		AEA		EA vs. AEA		(-/-) vs. (+/+)		ANOVA		
	(-/-)	(+/+)	(-/-)	(+/+)	in (-/-)	in (+/+)	in EA	in AEA	HIV/HAART	BGA	H/H x BGA
mTCHOL	203.4 (33.9)	201.8 (39.4)	190.7 (34.6)	181.4 (33.8)	$2.70x10^{-4}$	$< 1.00x10^{-7}$	0.98	0.09	$3.62x10^{-5}$	$<2.20x10^{-16}$	0.072
mLDL-C	127.8 (29.9)	118.9 (33.4)	117.7 (34.5)	103.9 (29.6)	$2.20x10^{-3}$	$2.00x10^{-7}$	$1.90x10^{-4}$	$1.80x10^{-4}$	$1.18x10^{-12}$	$1.45x10^{-13}$	0.045
mHDL-C	49.5 (11.5)	43.3 (11.5)	51.4 (11.4)	48.2 (13.8)	0.39	$1.50x10^{-6}$	$< 1.00x10^{-7}$	0.07	$3.96x10^{-15}$	$1.88x10^{-7}$	0.059
mTRIG	145.4 (117.1)	230.3 (199.2)	125.4 (75.0)	162.6 (108.7)	0.54	$1.00x10^{-7}$	$< 1.00x10^{-7}$	0.08	$< 2.20x10^{-16}$	$2.16x10^{-7}$	0.007

The mean mTCHOL levels for each ancestral group appear to vary similarly between the groups, with no significant difference between the HIV/HAART (-/-) and (+/+) groups seen for either population (Figure 9A). Biogeographical ancestry ($P = <2.20x10^{-16}$) and HIV/HAART status ($P = 3.62x10^{-5}$) had significant effects on mTCHOL but the interaction of the two ($P = 0.072$) was not significant and the pair-wise comparisons of HIV/HAART (+/+) to (-/-) controls

were not significant for both ancestral groups (AEA P = 0.08 | EA P = 0.98). (Table 7) This indicates that while HIV infection and HAART treatment appears to have an affect on mTCHOL levels, it doesn't appear to alter significantly the mean level in HIV/HAART (+/+) patients when compared to (-/-) controls.

Although no significant differences are seen in overall cholesterol levels, closer inspection of cholesterol subtypes does reveal some significant differences between ancestral groups and between HIV/HAART status groups. For example, mean mLDL-C levels vary differently between HIV/HAART status groups depending on genetic ancestry. (Figure 9B) The results of the ANOVA analysis show that HIV/HAART status ($P = 1.18x10^{-12}$) and biogeographical ancestry ($P = 1.45x10^{-13}$) have significant effects on the mean mLDL-C level. Additionally there is a significant interaction between HIV/HAART status and biogeographical ancestry (Table 7). The mean mLDL-C levels in the EA population are similar in the HIV/HAART (-/-) and (+/-) groups but are significantly lower in the (+/+) group. This lower mean mLDL-C level in the EA HIV/HAART (+/+) group is highly significant ($P = 1.96x10^{-4}$) when compared to the (-/-) men. The mean mLDL-C level in the AEA population is much lower in the HIV/HAART (+/-) group than in the (-/-) control group. The mLDL-C level in the AEA HIV/HAART (+/+) group is significantly lower than in the (-/-) group ($P = 1.85\ x10^{-4}$). These results show that, for both ancestral populations, the mean mLDL-C level is much lower in HIV positive men receiving HAART when compared to HIV/HAART negative controls.

The mean mHDL-C levels for both ancestral populations are lower in the HIV/HAART (+/-) group when compared with their (-/-) controls. (Figure 9C) The mean mHDL-C level in the AEA HIV/HAART (+/+) population was not significantly different in comparison to AEA (-/-) controls. In contrast, the mHDL-C level in the EA population was significantly lower in

HIV/HAART (+/+) patients than in (-/-) controls ($P < 1.00x10^{-7}$). (Table 7) The ANOVA results show that HIV/HAART status ($P = 3.96x10^{-15}$) and biogeographical ancestry ($P = 1.88x10^{-7}$) have significant effects on mHDL-C levels. The interaction between the two is not significant with a p-value of 0.059. These results show that the mean mHDL-C levels don't significantly differ for men of mixed African and European ancestry when infected with HIV and receiving HAART. Yet, for individuals of European ancestry the mean mHDL-C level of HIV/HAART (+/+) men is significantly lower when compared to the (-/-) controls, indicating that the impact of HIV infection and HAART use on mHDL-C levels differs between ancestral groups.

The mean mTRIG levels for both ancestral populations seem to be similar for HIV/HAART (+/-) men in comparison with (-/-) controls. In the EA population the mean mTRIG levels are almost 2-fold higher in the HIV/HAART (+/+) men ($P < 0.00x10^{-6}$) when compared to the (-/-) controls. (Figure 9D) In the AEA population the mean mTRIG levels appear to be slightly higher in the HIV/HAART (+/+) group in comparison with the HIV/HAART (-/-) controls but this difference is not statistically significant ($P = 0.08$). The results of the ANOVA show that the mean mTRIG levels are significantly affected by both biogeographical ancestry ($P = 2.16x10^{-7}$) and HIV/HAART status ($P = 2.20x10^{-16}$). There is evidence of an interaction between these two factors ($P = 7.00x10^{-3}$) and their effect on mean mTRIG levels. These results show that HIV infection and HAART do not seem to have significant affects on the mean mTRIG levels in men of mixed African and European ancestry, but that there appears to be a substantial effect of HIV infection and HAART therapy on the mean mTRIG levels in men of primarily European ancestry.

Our results show an important contribution of biogeographical ancestry on dyslipidemia beyond the known effects of HIV and HAART. The dyslipidemia profile in HIV/HAART (+/+)

individuals of primarily European ancestry is characterized by lower mLDL-C and mHDL-C levels, higher mTRIG levels but by no difference in mTCHOL. The lipid profile in HIV/HAART (+/+) men of mixed African and European ancestry is different, characterized by lower mLDL-C but no statistically significant changes in the other three lipid phenotypes.

4.4 GENETIC ASSOCIATIONS WITH SERUM LIPIDS

4.4.1 Single Locus Associations in HIV/HAART (+/+)

The results of the single locus association tests for lipid phenotype in HIV/HAART (+/+) men are shown in Figure 10. Each SNP was regressed against the mean lipid phenotype using multivariate linear regression and an additive genetic model with the first two components from the MDS included as covariates to control for any residual population stratification. Details of each association, including the number of individuals tested, genomic inflation factor and QQ-plots are shown in Figure 11 and summarized in Table 8. The QQ-plots show marked departure from an ordered normal distribution of p-values under the null hypothesis, most noticeable in (Figure 11-E) the QQ-plot for mHDL-C in the European ancestral group. This departure is most likely due our candidate gene study design, where we expect to have an enrichment of SNPs with positive associated p-values because we selected loci based on previous positive associations with serum lipid levels in the general population. Surprisingly there seems to be a substantial genomic inflation present in our study (Table 8) even though we went to great lengths to control for population substructure, the usual suspect when observing such high genomic inflation factors. The genomic inflation is most markedly increased in the association tests performed on

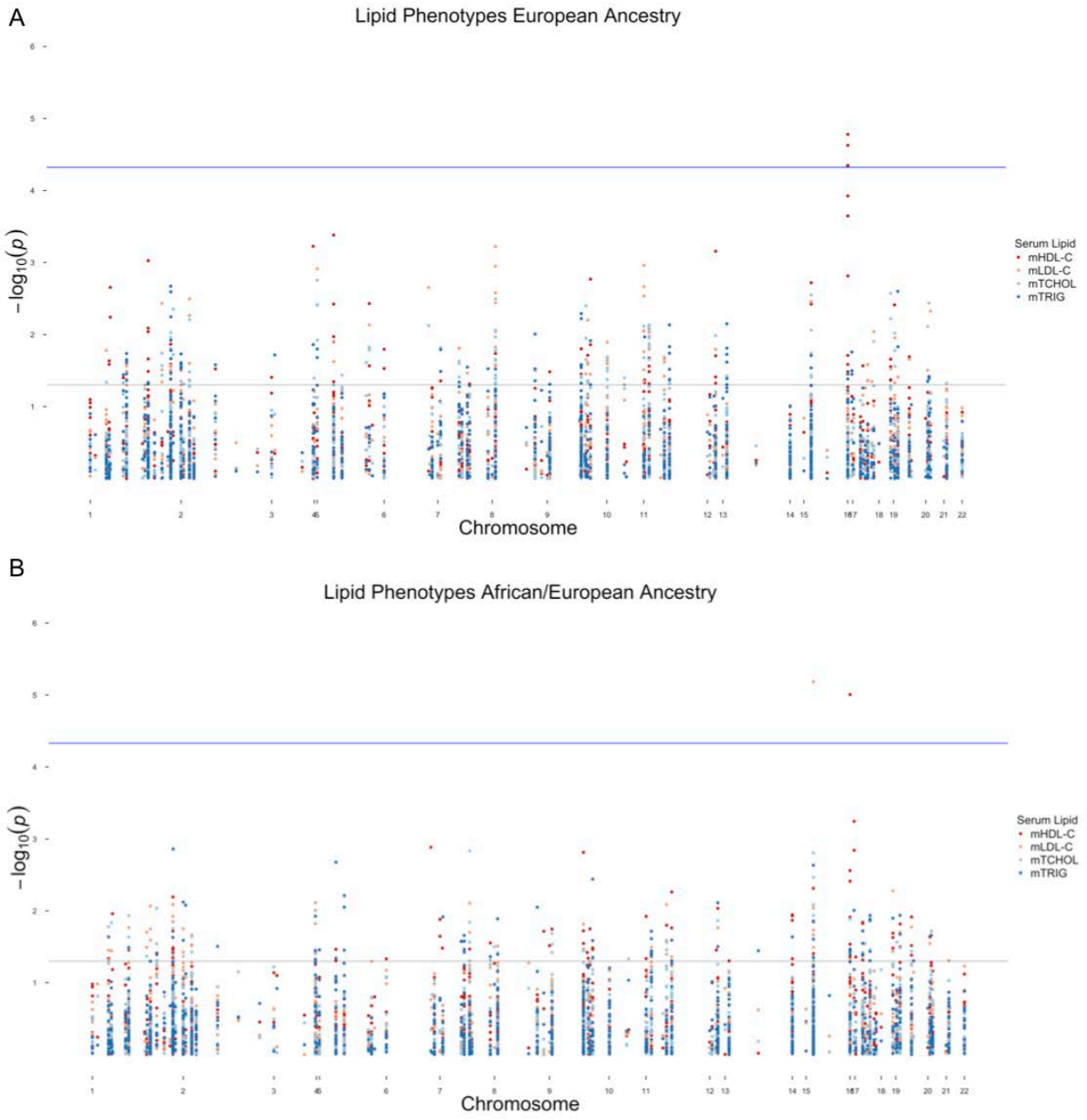


Figure 10 - Association Results for Serum Lipids in HIV/HAART (+/+) Patients

Results from the genetic association with the serum lipid phenotypes by ancestral population in HIV/HAART (+/+). (A) Manhattan plot of the –log10 association p-values by genomic position in the European ancestral group. (B) Manhattan plot of –log10 association p-values by genomic position in the African and European ancestral group. Association results from four serum lipid phenotypes are shown; red = mHDL, tan = mLDL, light blue = mTCHOL, and dark blue = mTRIG. The grey horizontal line represents a p-value of 0.05, blue horizontal lines represent a genomic significant p-value corrected for multiple independent tests; European ancestry p-value of 4.77×10^{-5}, African/European ancestry p-value of 4.66×10^{-5}.

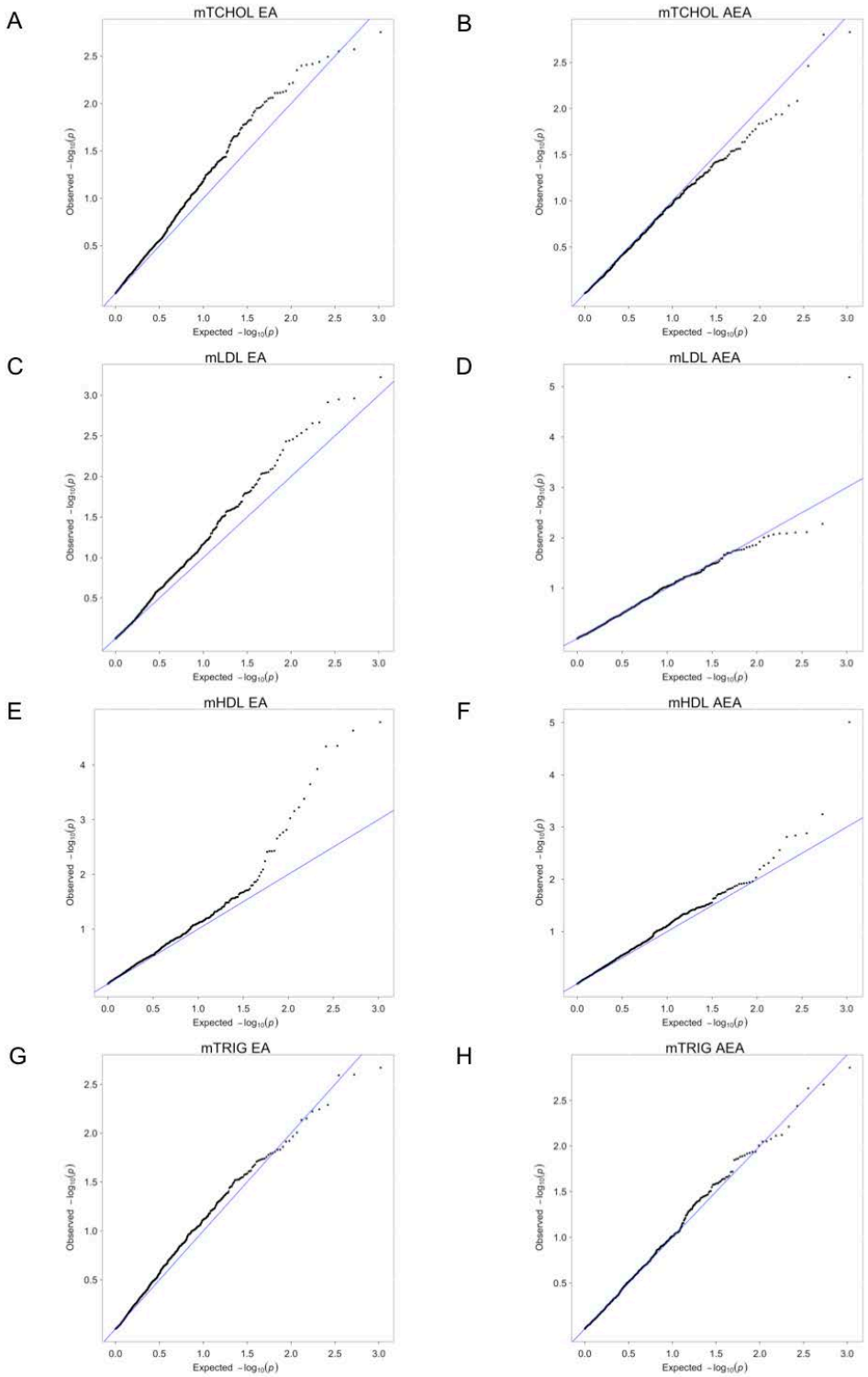


Figure 11 - QQ-Plots of HIV/HAART (+/+) Association P-values

Quantile-Quantile plots from the single locus genetic association studies. QQ-plots of expected vs. observed p-values from the genetic association with: mTCHOL (A, B); mLDL-C (C, D); mHDL-C (E, F); and mTRIG (G, H) in European (EA) and African/European ancestral (AEA) populations. The blue lines represent a line with a slope of one and an intercept of zero.

Table 8 - Association Diagnostics (+/+)

Association Diagnostics

	European Ancestry	African/European Ancestry
mTCHOL		
Patients	451	238
Loci	1050	1071
Genomic Inf.	1.261	1
mLDL		
Patients	410	224
Loci	1050	1071
Genomic Inf.	1.252	1
mHDL		
Patients	451	238
Loci	1050	1071
Genomic Inf.	1.235	1.159
mTRIG		
Patients	426	227
Loci	1050	1071
Genomic Inf.	1.201	1.159

the population of predominately European ancestry, even though this population should be the most ancestrally homogeneous of the two we tested. Taken together the residual deviations of the QQ-plots and the large inflation factors may be suggesting undefined familial relationships, unknown sample duplications, or poorly calibrated test statistics. We do not feel these results are indicative of a systematic bias in our data as is usually introduced by uncontrolled populations stratification but rather these results are reflecting characteristics of p-values that might be expected when performing a candidate gene association study.

In the European population 1,050 SNPs were tested and four loci (rs1532624, rs1532625, rs711752, and rs708272) had significant effects on mHDL-C levels at a Bonferroni-corrected critical p-value of 4.77×10^{-5} with unadjusted p-values of 1.66×10^{-5}, 2.36×10^{-5}, 4.48×10^{-5}, and

$4.59x10^{-5}$ respectively. All four SNPs are located in the cholesteryl ester-transfer protein (*CETP*) gene region on chromosome 16 (Figure 12-A). In the African/European population 1,071 SNPs were tested and two loci, rs261334 ($P = 6.53x10^{-6}$) and rs4783961 ($P = 9.83x10^{-6}$), had significant effects at the Bonferroni-corrected critical p-value of $4.66x10^{-5}$ respectively on mLDL-C and mHDL-C levels. The rs261334 SNP is located in the hepatic lipase (*LIPC*) gene region on chromosome 15 (Figure 13) and rs4783961 is located in the *CETP* gene region on chromosome 16 (Figure 12-B), but is not one of the SNPs in *CETP* shown above to affect mHDL-C levels in the European population. We also tested for the possible interaction of HAART with each lipid phenotype by adding an interaction term to our model and testing the control population HIV/HAART (-/-) and (+/+) groups. This interaction term tests for a difference in association with a phenotype between two groups and can suggest a gene *X* environment (HIV/HAART) interaction (Table 9). Of the six positive loci associated with lipid phenotypes, two showed significant differences in effect compared to the control population, rs261334 ($P^{inter} = 9.92x10^{-7}$) and rs4783961 ($P^{inter} = 6.78x10^{-4}$). Taken together these results show that polymorphisms in the *LIPC* and *CETP* gene regions that are associated with serum mLDL-C and mHDL-C cholesterol respectively in HIV positive men receiving HAART. Importantly, these associations seem to differ by biogeographical ancestry and may be affecting these serum lipids in ways different than in HIV negative HAART naïve individuals.

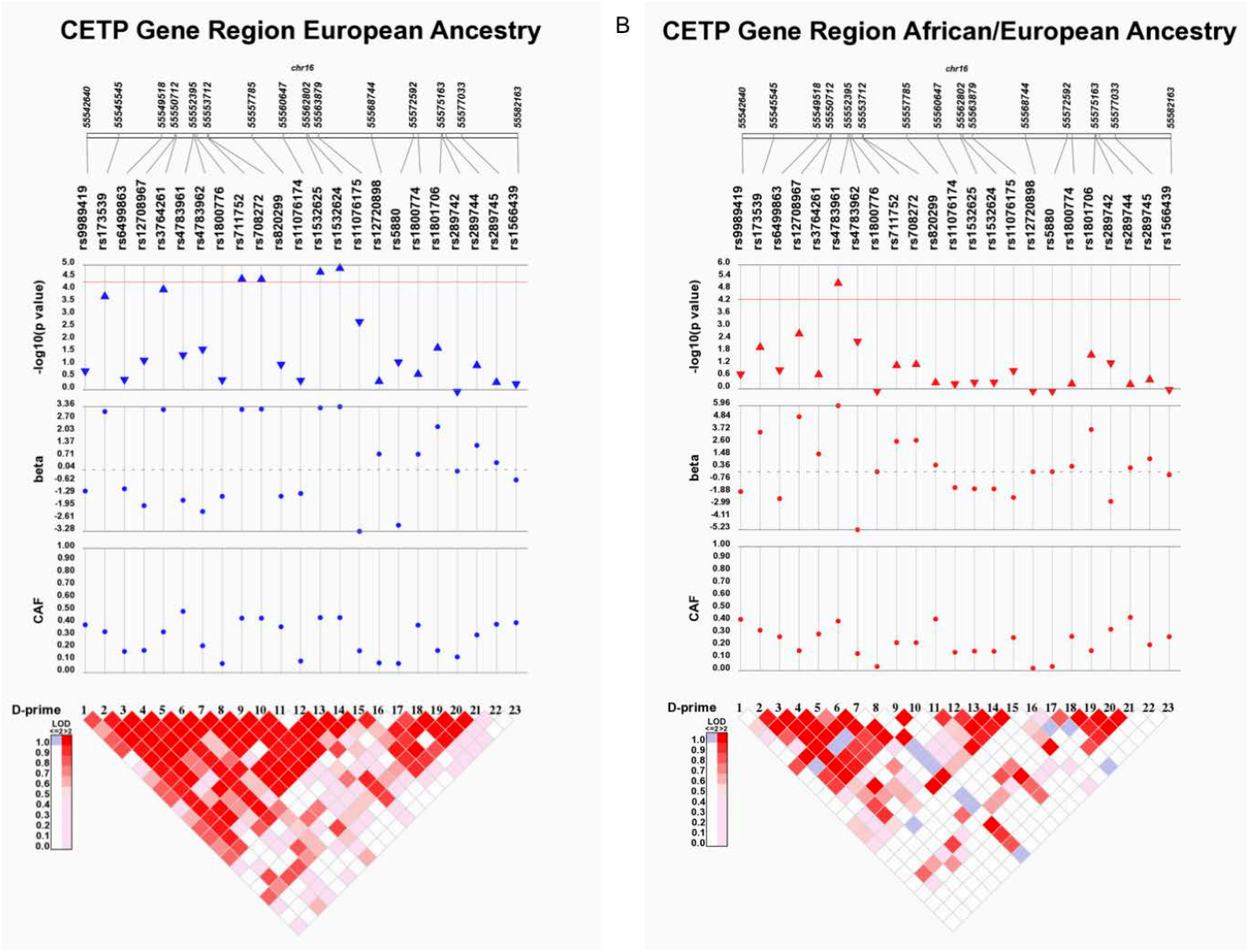


Figure 12 - Association Results for mHDL-C in the *CETP* Gene Region

Positive association results for mHDL-C in the *CETP* gene region on chromosome 16 for HIV/HAART (+/+). (Top) Genomic position (bp) of each SNP tested in the *CETP* gene region on chromosome 16. (Middle) The –log10 p-values with arrows indicated direction of effect of each SNP and red lines indicated genomic significance (AEA = 4.66×10^{-5}, EA = 4.77×10^{-5}), beta is the coefficient from each SNP's multiple linear regression equation and represents effective size, and CAF or coding allele frequency is the minor allele frequency of each SNP. (Bottom) Linkage disequilibrium plots generated from our genotype data for each ancestral group; (A) Association results in the European ancestral group. (B) Association results in the African / European ancestral group. Plots were generated in Synthesis-View[122].

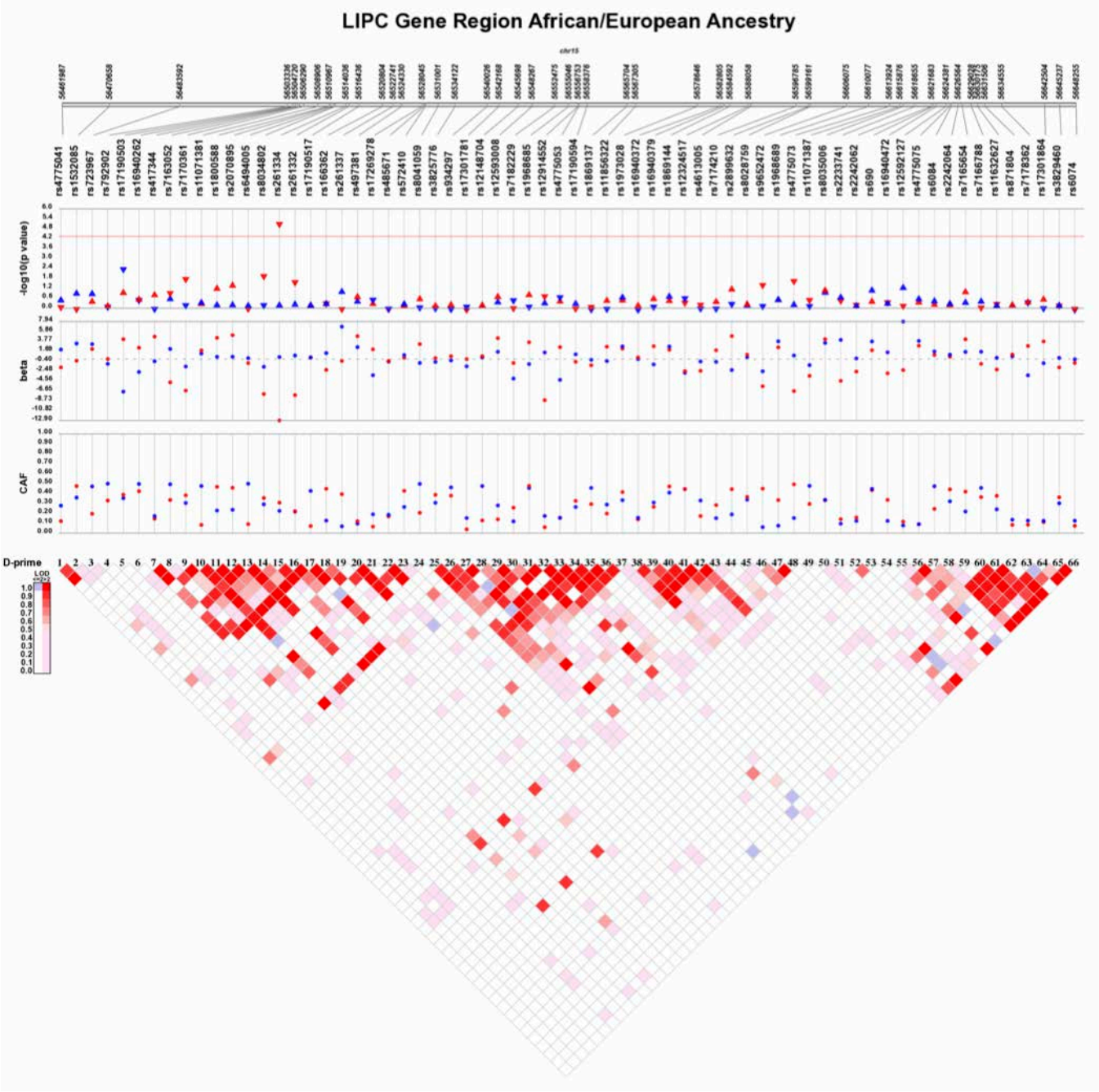


Figure 13 - Association Results for mLDL-C in the *LIPC* Gene Region

Association results for mLDL-C in the *LIPC* gene region on chromosome 15 for HIV/HAART (+/+). (Top) Genomic position (bp) of each SNP tested in the LIPC gene region on chromosome 15. (Middle) Red indicates AEA population and blue is the EA population. The –log10 p-values with arrows indicated direction of effect of each SNP and the red horizontal line indicates genomic significance at 4.66 x10^{-5}. Beta represents is the coefficient from each SNP's multiple linear regression equation and indicates effective size, and CAF or coding allele frequency is the minor allele frequency of each SNP. (Bottom) Linkage disequilibrium plot across the *LIPC* gene region calculated using genotype data from the African/European population. Plot was generated in Synthesis-View[122].

Table 9 - Results from Genetic Association Studies HIV/HAART (+/+)

Results from Single locus and Multi-locus Associations In HIV/HAART (+/+)

Single Locus Association

Phenotype - BGA	SNP	Chr	Gene/Region	MAF	Beta	P - Unadj	P - BONF	P - Inter	P - HWE	P - aPerm	P - mPerm
mLDL - African/European Ancestry	rs261334	15	LIPC	0.31	-12.90	$6.53x10^{-6}$	$6.99x10^{-3}$	$9.92x10^{-7}$	0.47	$2.29x10^{-5}$	0.009
mHDL - European Ancestry	rs1532624	16	CETP	0.44	3.36	$1.66x10^{-5}$	$1.74x10^{-2}$	0.12	0.86	$2.09x10^{-5}$	0.021
	rs1532625	16	CETP	0.44	3.30	$2.36x10^{-5}$	$2.48x10^{-2}$	0.25	0.79	$2.41x10^{-5}$	0.029
	rs711752	16	CETP	0.44	3.21	$4.48x10^{-5}$	$4.70x10^{-2}$	0.12	0.60	$2.41x10^{-5}$	0.051
	rs708272	16	CETP	0.44	3.23	$4.59x10^{-5}$	$4.82x10^{-2}$	0.16	0.54	$3.10x10^{-5}$	0.051
mHDL - African/European Ancestry	rs4783961	16	CETP	0.40	5.96	$9.83x10^{-6}$	$1.05x10^{-2}$	$6.78x10^{-4}$	0.24	$2.29x10^{-5}$	0.009

Haplotype Association	Chr	BP	SNPs	Haplotype	Beta	P - Unadj	P - aPerm	P -mPerm
mHDL - European Ancestry	16	55,560,647 - 55,562,980	rs11076174, rs1532625, rs1532624	AAA	3.30	$2.36x10^{-5}$	$1.00x10^{-4}$	0.028
	16	55,553,712 - 55,557,785	rs711752, rs708272, rs820299	AAA	3.21	$4.77x10^{-5}$	$1.00x10^{-4}$	0.053

4.4.2 Multilocus Haplotype Associations in HIV/HAART (+/+)

We next tested for haplotypes that may be associated with serum lipid levels in our populations by constructing haplotypes specific to each biogeographical ancestral population on the basis of the LD structure present in our genotype data. In the European population, 1,050 Tag SNPs were used to construct haplotypes in 230 genomic regions based on LD, which produced 875 discrete haplotypes that were then used for analysis. In the African/European population 1,071 Tag SNPs were used to construct haplotypes in 224 genomic regions yielding 798 discrete haplotypes. Two haplotypes were associated with mHDL-C in HIV/HAART (+/+) at the Bonferroni corrected p-value of $5.71x10^{-5}$ in the European ancestral group (Table 9). Both haplotypes are independent three SNP haplotypes located in the *CETP* gene region on chromosome 16. The most significant is an AAA haplotype for the three SNPs [rs11076174, rs1532625, and rs1532624 | $P = 2.36x10^{-5}$] spanning a 2.3Kb region of the *CETP* gene. The SNPs in this haplotype are all located in intronic

regions of *CETP* and span over five coding regions. The second is an AAA haplotype consisting of three SNPs [rs711752, rs708272, and rs820299 | P = $4.77x10^{-5}$] spanning a 4.1Kb region located 2.9Kb from the other significant haplotype in the *CETP* gene, but with no LD between these haplotypes. Similarly to the previous haplotype, these SNPs are located in intronic regions of *CETP* but this haplotype spans a region in the first intron that has been shown to possess a number of histone modification sites, or marks. These histone marks are often associated with epigenetic modifications that modulate the transcriptional activity of the gene. There were no haplotypes from the African/European ancestral group with significant p-values at the Bonferroni corrected critical limit. These results show two haplotypes located in the *CETP* gene that have significant effects on serum mHDL-C cholesterol levels in HIV positive patients receiving HAART of European ancestry, identifying this gene as critically important in the development of the HAART-associated dyslipidemia phenotype in this population.

4.4.3 Single Locus Associations in HIV/HAART (-/-) Controls

Lastly, we tested for genetic associations with serum lipid levels in the HIV/HAART (-/-) control population. The results from the single locus associations are shown in Figure 14. 14-A shows the association results in the European ancestral group. A total of 1,054 SNPs were tested in this population with a total of 8 SNPs being statistically significant at a Bonferroni-corrected critical p-value of $4.74x10^{-5}$. Four SNPs; rs3764261, rs173539, rs11076175, and rs711752 were associated with serum mHDL-C levels with unadjusted p-values of $1.52x10^{-6}$, $2.05x10^{-6}$, $4.13x10^{-6}$ and $3.96x10^{-5}$ respectively (Table 10). All four of these SNPs are located in the *CETP* gene region on chromosome 16 (Figure 15). The remaining four SNPs; rs2072560, rs964184, rs662799, and rs9326246 were associated with mTRIG levels having unadjusted p-values of

$2.36x10^{-6}$, $3.5x10^{-6}$, $4.64x10^{-6}$, and $6.07x10^{-6}$ respectively. These four SNPs are located in the *BUD13* homolog (*BUD13*) / zinc finger protein 259 (*ZNF259*) / apolipoprotein A-V (*APOA5*) gene region on chromosome 11 (Figure 15). Presumably, these SNPs are reflecting variations in the *APOA5* gene, which plays an important role in plasma TRIG levels, but these genes are located too closely together to definitively rule out the role of the other two. The African/European ancestral group had no genomically significant associations with any of the serum lipid phenotypes we tested. (Figure 14-B) These results show that SNPs in *CETP* and *BUD13/ZNF259/APOA5* gene regions are associated with serum mHDL-C and mTRIG levels, respectively, in HIV negative HAART naïve men of European Ancestry. In contrast, there were no associations with serum lipid levels in the HIV negative HAART naïve men of mixed African and European ancestry.

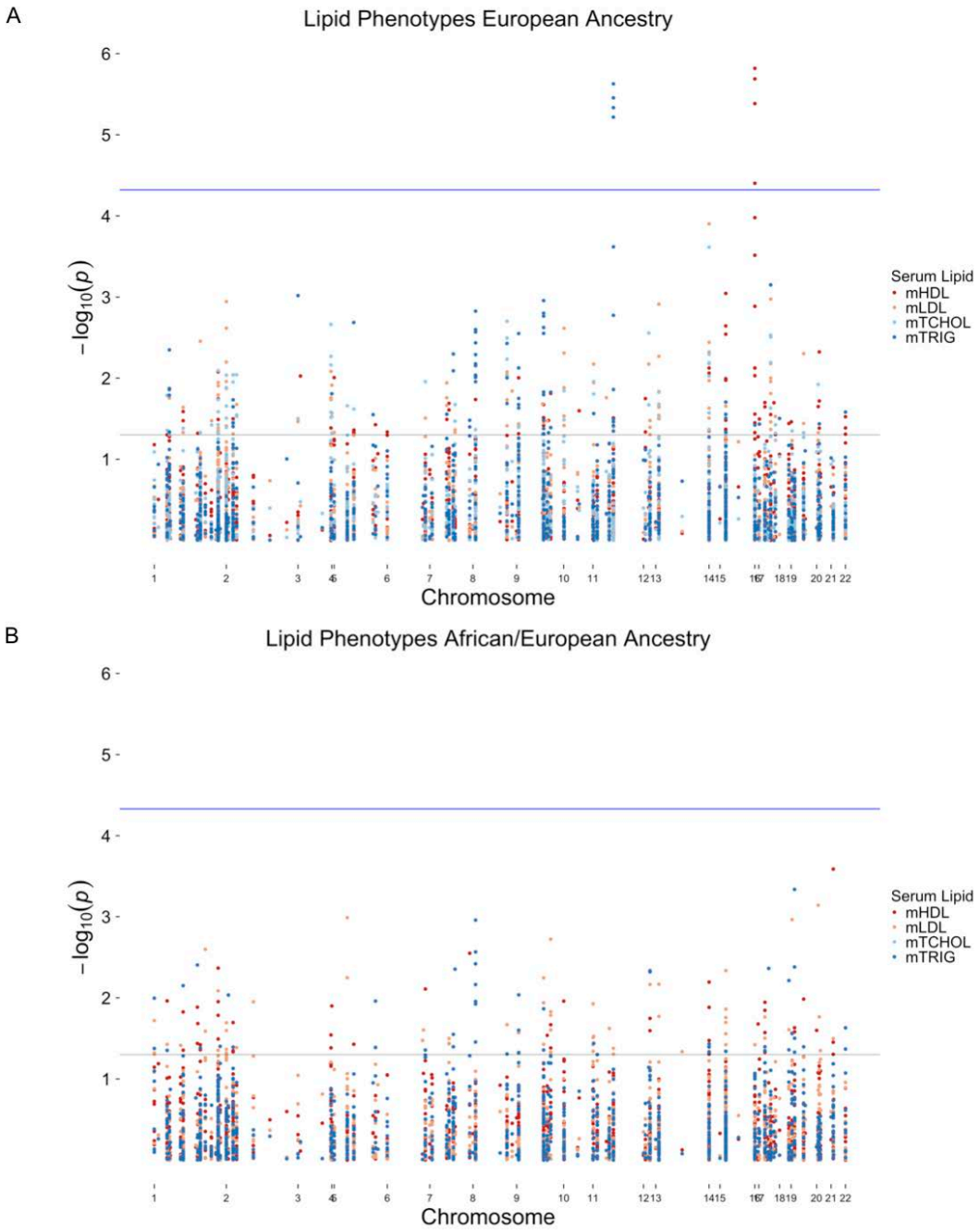


Figure 14 - Association Results for Serum Lipids in HIV/HAART (-/-)

Results from the genetic association with the serum lipid phenotypes by ancestral population in HIV/HAART (-/-). (A) Manhattan plot of the –log10 association p-values by genomic position in the European ancestral group. (B) Manhattan plot of –log10 association p-values by genomic position in the African and European ancestral group. Association results from four serum lipid phenotypes are shown; red = mHDL-C, tan = mLDL-C, light blue = mTCHOL, and dark blue = mTRIG. The grey horizontal line represents a p-value of 0.05, blue horizontal lines represent a genomic significant p-value corrected for multiple independent tests; European ancestry p-value of 4.77×10^{-5}, African/European ancestry p-value of 4.66×10^{-5}.

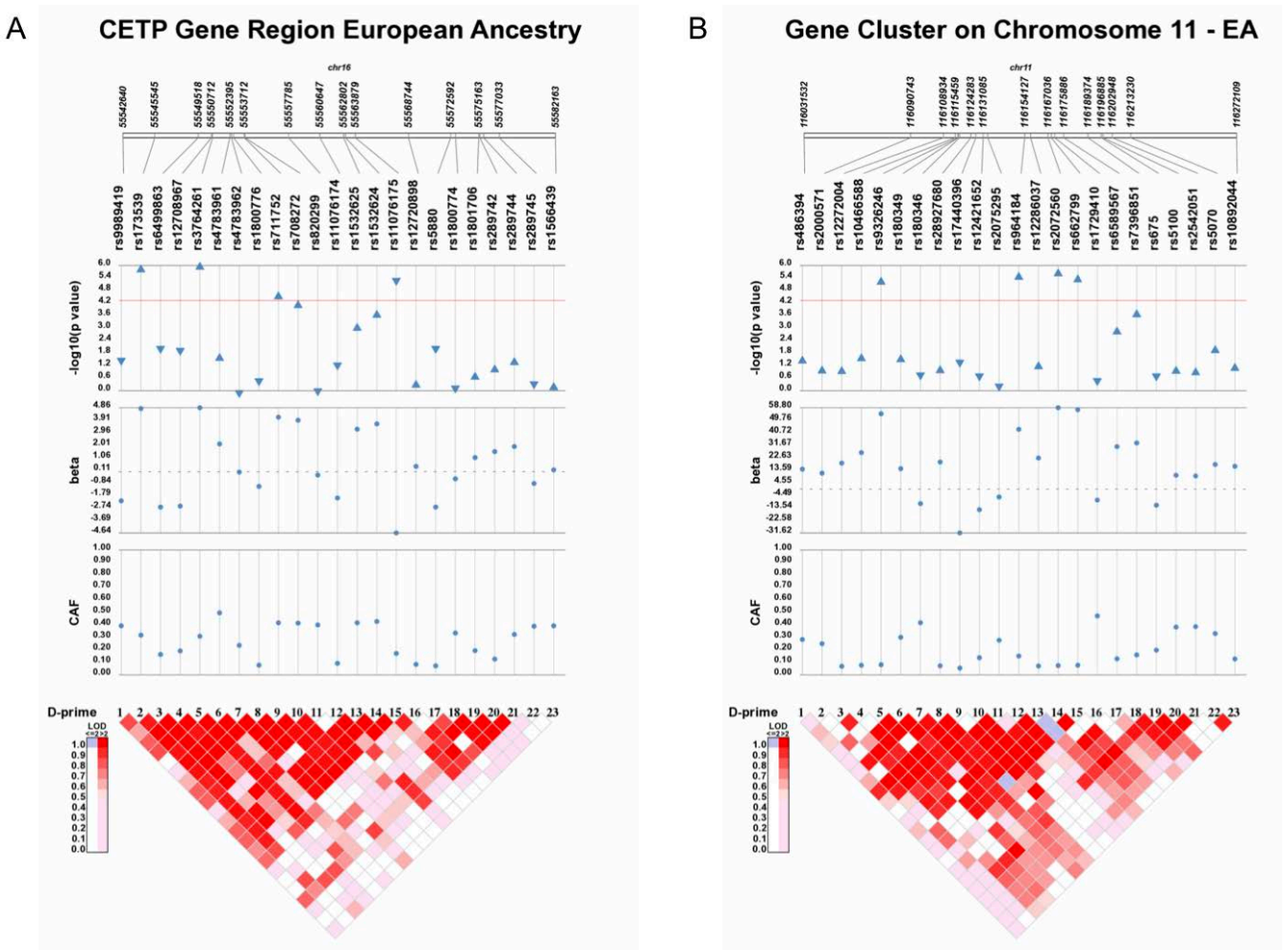


Figure 15 - Associated Gene Regions in HIV/HAART (-/-) Patients of European Ancestry

Positive association results for HIV/HAART (-/-) of European ancestry (Top) Genomic position (bp) of each SNP tested in the *CETP* gene region on chromosome 16. (Middle) The –log10 p-values with arrows indicated direction of effect of each SNP and the red line indicated genomic significance (EA = 4.77×10^{-5}), beta is the coefficient from each SNP's multiple linear regression equation and represents effective size, and CAF or coding allele frequency is the minor allele frequency of each SNP. (Bottom) Linkage disequilibrium plots generated from our genotype data; (A) Association results with serum mHDL-C levels in the *CETP* gene region on chromosome 16 for the European ancestral group. (B) Association results with serum mTRIG levels on chromosome 11 for the European ancestral group. Plots were generated in Synthesis-View[122]

Table 10 - Single Locus Association HIV/HAART (-/-)

Results from Single Locus Associations HIV/HAART (-/-)

Phenotype - BGA	SNP	Chr	Gene/Region	MAF	Beta	P - Unadj	P - BONF	P - HWE
mHDL - European Ancestry	rs3764261	16	CETP	0.31	3.33	1.52×10^{-6}	1.597×10^{-3}	0.93
	rs173539	16	CETP	0.32	3.27	2.05×10^{-6}	2.159×10^{-3}	0.60
	rs11076175	16	CETP	0.17	-3.91	4.13×10^{-6}	4.357×10^{-3}	0.89
	rs711752	16	CETP	0.42	2.74	3.96×10^{-5}	4.174×10^{-2}	0.14
mTRIG - European Ancestry	rs2072560	11	BUD13/ZNF259/APOA5	0.07	58.80	2.36×10^{-6}	2.49×10^{-3}	0.79
	rs964184	11	BUD13/ZNF259/APOA5	0.15	43.21	3.51×10^{-6}	3.70×10^{-3}	0.46
	rs662799	11	BUD13/ZNF259/APOA5	0.08	57.40	4.64×10^{-6}	4.89×10^{-3}	0.79
	rs9326246	11	BUD13/ZNF259/APOA5	0.08	54.44	6.07×10^{-6}	6.40×10^{-3}	1.00

5.0 DISCUSSION

Exposure to highly active antiretroviral therapy can have a number of metabolic side effects. The dyslipidemia observed in these patients is well-documented but varies in prevalence in any given population. We hypothesized that this variation was due to the genetic predisposition of individuals interacting with HAART and/or HIV infection. To test this hypothesis we designed a custom Illumina Golden Gate array consisting of candidate loci that had been associated with dyslipidemia and/or cardiovascular disease in the general population. We used this array to genotype 1,945 individuals enrolled in the Multicenter AIDS Cohort Study and performed genetic associations on four serum lipid phenotypes obtained from the MACS data.

5.1 DYSLIPIDEMIA PHENOTYPES BY TREATMENT STATUS AND BGA

Stratification of our population by biogeographical ancestry and subsequent examination of serum lipid levels, showed an intriguing result; the mean levels of some of the phenotypes tested varied by treatment status dependent on biogeographical ancestry. People from a particular ancestral population have different metabolic responses to HIV infection and HAART. Serum lipid levels are known to be different depending on race in the general population. Specifically, Whites have been shown to have higher TRIG levels than Blacks with Blacks having higher HDL-C levels and no difference in LDL-C or TCHOL compared to Whites[42,43]. In our control

population we observed that the EA and AEA populations had no significant difference in mTRIG or mHDL-C levels, yet we observed higher levels of mLDL-C and mTCHOL in the EA group compared to the AEA population. (Table 7) Indeed, when we formally tested this observation we found, that for every lipid phenotype we tested, biogeographical ancestry was a significant factor affecting mean serum lipid levels. The ANOVA results also showed that HIV/HAART status had significant effects on mean serum lipid levels for all phenotypes tested. The interaction term between HIV/HAART status and biogeographical ancestry was significant in two phenotypes, mLDL-C and mTRIG, but not significant for mHDL-C and mTCHOL. This suggests that there may be an interaction between HIV/HAART and biogeographical ancestry in affecting the mean serum levels of mLDL-C and mTRIG. We assume this interaction is due, at least in part, to shared genetic variants in a BGA group, yet other shared social, cultural, or environmental variants may also be contributing to this observed interaction. HIV/HAART (+/+) individuals of European ancestry had higher mean levels of mTRIG, mLDL-C, and mTCHOL when compared to the African/European population. HIV/HAART (+/+) individuals of African and European ancestry had higher mHDL-C levels when compared to the European population. These results support previous findings that HIV infection and HAART treatment results in dyslipidemia as well as offering a new understanding of the extent to which the degree of metabolic change depends on an individual's biogeographical ancestry, presumably due to shared genetic susceptibility. Additionally these results help us characterize the different dyslipidemia profiles at the population level for each ancestral group. Specifically we see that HIV positive men of primarily European ancestry receiving HAART have lower mLDL-C and mHDL-C with higher mTRIG levels when compared to HIV/HAART (-/-) controls. HIV positive men of mixed African and European ancestry receiving HAART have a different

dyslipidemia profile consisting of only lower mLDL-C levels when compared to HIV/HAART (-/-) controls. These results are encouraging because they hint at validating our initial hypothesis that dyslipidemia in HIV positive patients receiving HAART is in part due to genetic susceptibility, in this case biogeographical ancestry.

5.2 GENETIC ASSOCIATIONS WITH SERUM LIPIDS

We found a number of interesting and perhaps novel loci associated with mLDL-C and mHDL-C levels in HIV positive patients receiving HAART. The only locus we found to be associated with mLDL-C levels for either population was rs261334. This SNP is located in the first intron of *LIPC* on chromosome 15. The *LIPC* gene product, hepatic lipase, hydrolyzes phospholipids and triglycerides carried by lipoproteins in the serum as well as being a bridging factor for receptor-mediated lipoprotein uptake in the liver. Polymorphisms in the promoter region of *LIPC* have previously been shown to be associated with cardiovascular disease and lipid metabolism[123-125]. Interestingly, the effects of hepatic lipase are usually associated with HDL-C and LDL-C particle size and density rather than serum levels of these lipoproteins, but in our study we have found that the rs261334 polymorphism is associated with lower serum mLDL-C levels in HIV/HAART (+/+) men of African and European ancestry compared to HIV/HAART (-/-) controls. This association between hepatic lipase and serum mLDL-C is novel and may be suggestive of some altered role or function due to HIV infection and/or HAART treatment. This is suggestively supported by the positive interaction of the rs261334 SNP in the HIV/HAART (+/+) population compared to the HIV/HAART (-/-) controls. The rs261334 polymorphism is in linkage disequilibrium with a number of SNPs in the 5' and promoter regions of *LIPC* gene and may not

itself be the causative allele. The five other SNPs we found to be significantly associated with mHDL-C levels in both ancestral populations were in the *CETP* region on chromosome 16. In the men of European ancestry four SNPs (rs711752, rs708272, rs1532625, and rs1532624) and two haplotypes containing these four SNPs were found to be associated with increased mHDL-C levels. The *CETP* gene product transfers cholesterol esters from HDL-C to other lipoproteins, such as VLDL and LDL-C, in exchange for TRIG. *CETP* has been shown to be important in lipid metabolism relating to cardiovascular disease and can have pro- or anti-atherogenic effects. One of our associated SNPs, rs708272, is also known as TaqIB and is one of the oldest-known and most extensively studied polymorphisms in this gene[126-128]. The TaqIB variant does not alter the amino acid sequence of the *CETP* protein but is in strong LD with another SNP rs1800775, also known as -C629A, which is located in the promoter region of *CETP* and found to affect promoter activity directly. The LD in this gene makes it difficult to pinpoint a specific disease marker. In the European population all four of our significantly associated SNPs are in strong LD with each other. Interestingly, in the African/European population these four SNPs failed to be associated with mHDL-C. Another SNP in *CETP*, rs4783961, was found to be significantly associated with mHDL-C in men of African and European ancestry. The LD structure in African/Europeans is weaker than in Europeans but it seems from the LD plot generated from our genotype data that rs4783961 polymorphism tags a similar region in *CETP* as those in the European population.

Of the four lipid phenotypes we tested, only mLDL-C and mHDL-C showed genetic associations. This could be due to a number of possibilities. There might be additional loci affecting these serum lipid levels that were not included on our array, or we may have not had sufficient power to detect the presumable small effect size present at each locus, which is

common in complex traits. Since we stratified our population by biogeographical ancestry and HIV/HAART status, the number of subjects in each test group did diminish. Despite this potential loss in power, we feel this stratification method best controls the bias introduced when performing genetic associations on admixed populations. Another possible reason for the lack of other associations may be that HAART causes such severe disturbances in cholesterol metabolism that any known genetic susceptibility to dyslipidemia or CVD has little effect.

The MACS population consists of both HIV infected and HIV uninfected men. This provides a control population from which we can perform genetic associations with serum lipids and compare these results to those obtained from the HIV/HAART (+/+) population. Similarly to the HIV/HAART (+/+) population, we found associations in the *CETP* gene region with serum mHDL-C levels in men of European ancestry. Three of the SNPs; rs3764261, rs173539 and rs711752, are located in the promoter region of *CETP* and are in strong LD with the each other as well as the known clinically relevant SNPs rs708272 and rs1800775. One SNP, rs11076175, is located in the intron region of *CETP*. Although rs11076175 appears to be in high LD with the three other positively associated SNPs, its effect on serum mHDL-C is in the opposite direction to the three SNPs located in the promoter region. This may suggest a separate region of interest that has an association with decreased levels of mHDL-C in the HIV/HAART (-/-) men of European ancestry. These associations with decreased mHDL-C levels in the HIV/HAART (-/-) controls do not persist in the (+/+) test population. The association with increased serum levels of mHDL-C for these *CETP* polymorphisms appears similar across the HIV/HAART treatment groups and does not explain the observed lower mHDL-C levels seen in the HIV/HAART (+/+) compared to the (-/-) controls.

We also observed an association with increased serum mTRIG in a region on chromosome 11 in HIV/HAART (-/-) men of European ancestry. This region contains a cluster of six genes; *BUD13* homolog (*BUD13*), zinc finger protein 259 (*ZNF259*), apolipoprotein A-V (*APOA5*), apolipoprotein A-IV (*APOA4*), apolipoprotein C-III (*APOC3*), and apolipoprotein A-I (*APOA1*). Four SNPs in this region were significantly associated with increase mTRIG levels; rs2072560, rs964184, rs662799, rs9326246 and were located in or around the *BUD13/ZNF259/APOA5* genes. Specifically, rs2072560 is located in an intron of *APOA5*, rs964184 is downstream proximal to the 3'-UTR of *ZNF259*, rs662799 is located in the 5' promoter region of *APOA5*, and rs9326246 is downstream of *BUD13*. There is strong LD between these SNPs yet we hypothesize that *APOA5* is the probably candidate explaining this association with serum mTRIG levels. It has been shown that polymorphisms in *APOA5* have strong associations with TRIG concentrations and CVD risk in the general population[129,130]. In fact, one of our SNPs rs662799, is also known as T-1131>C and is one of the most widely studied polymorphisms in this gene. The fact that we identified this gene as being associated with serum mTRIG levels in HIV/HAART (-/-) men of European ancestry is not surprising. What is surprising, that even though the HIV/HAART (+/+) population had almost two fold higher levels of serum mTRIG levels, this gene cluster was not found to be significantly associated with mTRIG levels in the (+/+) men of European ancestry. This is evidence that the genes or mechanisms involved in the elevated mTRIG levels in HIV/HAART (+/+) patients may be novel and different than what has been identified in the general population.

There are no statistically significant associations with any of the four serum lipid levels in HIV/HAART (-/-) men of African and European ancestry. This surprising result, along with the possible reasons for this lack of association, will be discussed in the following section.

5.3 UNANSWERED QUESTIONS AND LIMITATIONS

While we have identified a number of SNPs associated with serum lipid levels in our HIV/HAART positive population, it is disconcerting we did not find more positive associations. In designing our study we selected for, what is generally thought to be, the most important reproducible genetic determinants associated with dyslipidemia and CVD in the general population. This candidate study approach should have led to a number of positive associations being identified, that fact that we failed to do so means we need to address the potential weaknesses and limitations present in our study.

5.3.1 Over-Correction in Candidate Gene Association Studies

The problem of false positive results in genetic association studies due to multiple independent comparisons is widely acknowledged. The current methods to correct for multiple comparisons include using a Bonferroni corrected p-value or a False Discovery Rate procedure. These procedures are often used in GWAS studies because there usually is no initial hypothesis to which genes may be associated with a given phenotype. A GWAS tests for thousands to millions of SNPs evenly distributed across the entire human genome to capture the maximum genetic variation present in a population.

In our study we chose to use a Bonferroni correction because we initially wanted to be as conservative as possible to limit our Type I error. Given our candidate loci approach this kind of

correction may be overly conservative because, unlike a GWAS, we had a defined set of hypotheses. We selected 267 SNPs that had been previously associated with dyslipidemia and CVD in the general population, we also selected 1,146 tagSNPs for 75 candidate genes we identified. By definition, these tagSNPs were selected based on their genetic correlation to other SNPs located in the gene but are generally independent of other tagSNPs. Even though these tagSNPs are mostly independent of each other they still may be in high LD with each other depending on the genetic correlation structure present in a gene. Using the Bonferroni correction for tests that are highly dependent, tagSNPs within a gene may be in high LD with other tagSNPs, is inappropriate and results in an overly conservative critical p-value. A better method for correction may be not to divide α by the *n* number of tests being performed but rather use an *n* for the number of independent loci being tested. Of course we are assuming these genes are independent of each other, biologically speaking, these gene products may be involved in shared physiological or cellular process and may impact the function and/or expression of each other directly. This would then violate the assumption of independence and thus it may be appropriate to use an even smaller *n* in correcting α for truly independent tests. Currently, there is no accepted method for assessing the true independence of loci being tested in a candidate gene study so we used a Bonferroni correction based on the number of tests performed with the understanding we are increasing the Type II error (false negatives) in our study.

5.3.2 The Problem of Power

The power of a study is the probability to reject the null hypothesis when the null hypothesis is false. The power and Type II error are directly related; in that power is equivalent to 1-β, where β is the Type II error. In our study we used a Bonferroni correction, which increases the Type II

error in favor of decreasing the study's Type I error. Thus, we are intrinsically decreasing the power of our study because we want to limit the probability of falsely rejecting the null hypothesis. Statistical power is also dependent on other factors, such as the magnitude of effect being investigated and the sample size used to detect this effect. Generally, when trying to detect a small effect you need an increasingly large sample. Dyslipidemia is a complex human trait. Serum lipid levels are affected by a number of different genes with varying effects sizes. Given the lack of positive associations in our study, we may have been underpowered to detect the effect of genes we tested for, either because of our stringent statistical correction or smaller than needed sample size. Ideally we would like to genotype SNPs with the strongest associations in the remaining MACS population to valid our results and to increase the power to detect genes with important but small effects.

5.4 FUTURE WORK

5.4.1 Investigating Gene Expression

Our association results identified a number of SNPs associated with altered serum lipid levels in HIV/HAART positive individuals. The SNPs we identified are in strong LD with known variants in the promoter regions of *CETP* and *LIPC*. This suggests that the genetic variation we found to be associated with serum lipid levels is related to the expression and thus the protein levels of the genes we have identified, primarily through alterations in promoter activity. We hypothesize the mechanism by which these genes affect serum lipids are as follows and outlined in Figure 16: SNPs associated with decreased *CETP* expression could results in the observed association with

increased HDL-C and SNPs associated with increased expression of *LIPC* could ultimately result in the observed association of decreased LDL-C, through an intermediate step involving HDL-C. The fact that, familial CETP deficiency is characterized by increased levels of HDL-C supports our first mechanistic hypothesis that decreased levels of CETP may explain the association with increased HDL-C. In contrast, LIPC (hepatic lipase) is more often implicated in receptor mediated HDL-C uptake at the liver and modulations in LDL-C particle size and density. The association with decreased LDL-C may be a novel insight to LIPC's role in lipid metabolism, especially in our unique population.

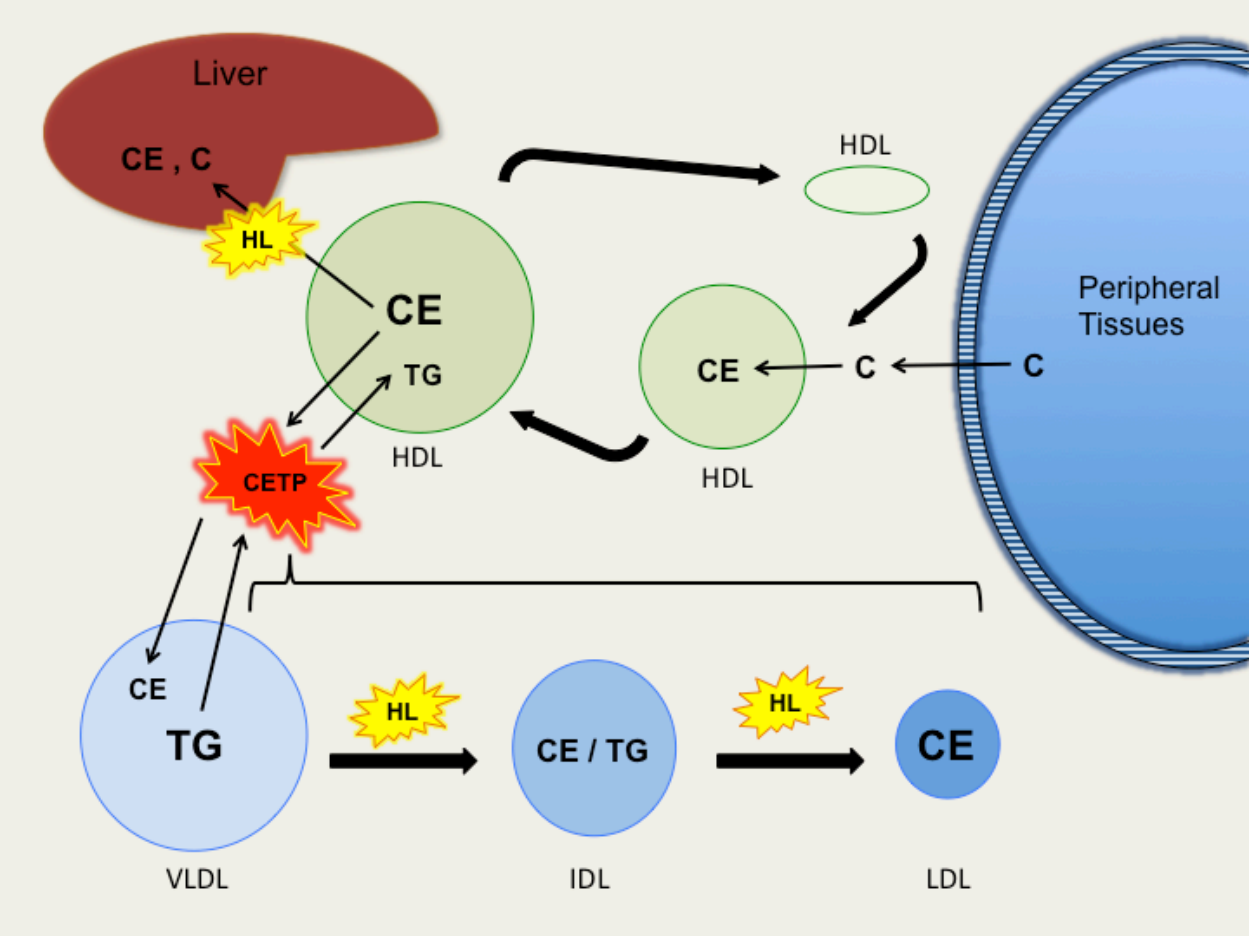


Figure 16 - Possible Mechanisms of CETP and LIPC

To formally investigate these potential mechanisms of action we have two options that are currently available to us, which focus on expression levels of *CETP* and *LIPC*. The first technique we could employ would be to use a luciferase-reporter gene assay. In this experiment the candidate promoter region containing the SNPs of interest are cloned into reporter vectors containing gene/s downstream of the promoter then when transcribed, produce a luminescent protein. The luminescence of the protein can be read and compared to a wild-type control. If the mutant promoter alters the transcriptional activity, its luminescence will be different (higher or lower) compared to the wild-type. The second method is one our lab is already performing, RNA transcriptome studies. In these experiments RNA is extracted from patient PBMCs and run on a RNA microarray. The microarray reports relative expression levels of over 30k known RNA transcripts. We can take this raw data and compare people who possess the SNPs of interest to people who do not possess these SNPs. In this way we can see if there is a relative difference in expression of *CETP* and *LIPC* associated with the proposed altered promoter function.

5.4.2 Genetic Associations on Additional MACS Phenotypes

The main outcomes of our study were serum lipid levels that we associated with dyslipidemia in HIV infected individuals receiving HAART. Even though we focused on dyslipidemia in this study, we ultimately are interested in the relationship of these lipids to CVD risk in our population. Serum lipids are not the only biological phenotypes we can investigate; there are a number of CVD risk surrogates, which are indicators of subclinical atherosclerosis. A recent study was done by Kingsley et al.[71], in which they investigated coronary artery calcification (CAC) in HIV/HAART positive patients in the Pittsburgh Men's Study (PMS), a subset of the

MACS. We would like to use our genotype data for the PMS participants to investigate the possible association of lipid metabolism and inflammatory genes to CAC development in HIV/HAART positive patients.

As discussed in the introduction, cholesterol availability is critical to the HIV lifecycle. HIV employs a number of strategies to modulate and sequester cellular cholesterol during its infection. We are interested in the possibility that people with different SNPs in cholesterol metabolism genes may respond to HIV infection in different ways. Specifically we are interested to see if our SNPs are associated with HIV disease progression, including; viral load, years to first reported AIDS symptoms, and rapid versus long-term non progression.

5.5 PUBLIC HEALTH IMPACT AND SIGNIFICANCE

We are the first to show that dyslipidemia in HIV positive men receiving HAART differs on the basis of biogeographical ancestry. We also report the first characterization of the dyslipidemia profiles in these patients at the population level. We hypothesized that the dyslipidemia seen in this population was due to a preexisting genetic susceptibility. Our results show that for two of the phenotypes tested a genetic association is present and these associations appear different in each ancestral population. Interestingly, only the polymorphism in *LIPC* in the African/European population agrees with the observed changes in serum mLDL-C levels of HIV positive men receiving HAART. These results advance our understanding of the mechanism of HAART-associated dyslipidemia and may lead to improved treatment of these patients. For example, our data show that HAART-associated dyslipidemia varies by BGA. In fact, HIV/HAART (+/+) men of African and European ancestry seem to develop an anti-atherogenic lipid profile characterized

by lower levels of LDL-C, which was the only serum lipid we tested that was significantly different than (-/-) controls for this population. This implies that any prophylaxis treatment with statins for HIV/HAART (+/+) men of mixed African and European ancestry would be unnecessary. Obviously it is important for physicians to treat high LDL-C levels with statins on a case by case basis, but our data show that LDL-C levels of HIV/HAART (+/+) men was lower on average when compared to (-/-) controls suggesting that HAART may not actually be increasing CVD risk mediated by increasing LDL-C levels. Rather, this observed increase previously associated with HAART is a restoration to health of HIV infected individuals.

APPENDIX

SNP INFORMATION

Table 11 - SNP Information

SNP Name	Chr	Coordinate	Clinical	AIM	Tag	References (Included in Bibliography)
rs2986742	1	6472963		Y		Kosoy et al. - 2009
rs4845882	1	11765754			Y	
rs1537516	1	11770448			Y	
rs4846049	1	11772952			Y	
rs1476413	1	11774887			Y	
rs12121543	1	11777258			Y	
rs1801133	1	11778965			Y	
rs7533315	1	11783270			Y	
rs17037397	1	11784750			Y	
rs17367504	1	11785365			Y	
rs2066472	1	11785558			Y	
rs6541030	1	12530765		Y		Kosoy et al. - 2009
rs647325	1	18043473		Y		Kosoy et al. - 2009
rs10903129	1	25641524	Y			Aulchenko et al. - 2009
rs4908343	1	27804285		Y		Kosoy et al. - 2009
rs1325502	1	42132857		Y		Kosoy et al. - 2009
rs11206510	1	55268627	Y			Willer et al. - 2008, Kathiresan et al. - 2008, Aulchenko et al. - 2009
rs7525649	1	55271744			Y	
rs2495488	1	55272794			Y	
rs2479409	1	55277238			Y	
rs4927193	1	55282460			Y	
rs499718	1	55285137			Y	
rs2479412	1	55286661			Y	
rs11206514	1	55288592			Y	
rs572512	1	55289932			Y	
rs7552841	1	55291340			Y	
rs557435	1	55293452			Y	
rs533375	1	55295949			Y	

rs615563	1	55298884			Y	
rs10465832	1	55301395			Y	
rs505151	1	55301775			Y	
rs17111555	1	55302190			Y	
rs662145	1	55302416			Y	
rs13312	1	55305330			Y	
rs12130799	1	55435960		Y		Kosoy et al. - 2009
rs1167998	1	62704220	Y			Aulchenko et al. - 2009
rs10493326	1	62725961			Y	
rs1168026	1	62743584			Y	
rs17123694	1	62753688			Y	
rs13375691	1	62768961			Y	
rs1748195	1	62822181	Y			Willer et al. - 2008
rs11207998	1	62864059			Y	
rs10889353	1	62890784	Y			Aulchenko et al. - 2009, Kathiresan et al. - 2009
rs912540	1	62907715			Y	
rs12130333	1	62964365	Y			Kathiresan et al. - 2009
rs12080049	1	63012419			Y	
rs17123828	1	63020039			Y	
rs17316381	1	63046891			Y	
rs17388604	1	63103662			Y	
rs1413242	1	63104452			Y	
rs1892534	1	65878532	Y			
rs12753193	1	65942267	Y			Sabatti et al. - 2009
rs3118378	1	68622275		Y		Kosoy et al. - 2009
rs4584412	1	100954164			Y	
rs1409419	1	100955984			Y	
rs3176860	1	100959807			Y	
rs3176861	1	100959909			Y	
rs3917009	1	100961998			Y	
rs3917012	1	100968247			Y	
rs3176870	1	100969771			Y	
rs3176871	1	100969877			Y	
rs3917014	1	100969972			Y	
rs3917016	1	100973735			Y	
rs3176877	1	100975983			Y	
rs3176878	1	100976286			Y	
rs3783621	1	100977083			Y	
rs10493936	1	100985923			Y	
rs3737576	1	101482151		Y		Kosoy et al. - 2009
rs4623734	1	109584728			Y	
rs4246519	1	109585078			Y	
rs17035415	1	109589016			Y	
rs585362	1	109591318			Y	
rs437444	1	109597131			Y	
rs10858082	1	109600244			Y	
rs626387	1	109600326			Y	
rs6689614	1	109608622			Y	
rs6657811	1	109608806	Y			Sandhu et al. - 2008

rs17035665	1	109615242			Y	
rs4970834	1	109616403	Y			Sandhu et al. - 2008
rs611917	1	109616775	Y			Sandhu et al. - 2008
rs12740374	1	109619113	Y			Kathiresan et al. - 2009
rs660240	1	109619361	Y			Sandhu et al. - 2008
rs646776	1	109620053	Y			Kathiresan et al. - 2008, Sabatti et al. - 2009, Aulchenko et al. - 2009
rs17035949	1	109622442			Y	
rs602633	1	109623034	Y			Willer et al. - 2008
rs599839	1	109623689	Y			Willer et al. - 2008, Kathiresan et al. - 2008
rs657420	1	109627659			Y	
rs17584208	1	109634710			Y	
rs17645143	1	109637280			Y	
rs3850615	1	109641261			Y	
rs7515901	1	109641419			Y	
rs443345	1	109650777			Y	
rs464218	1	109657829			Y	
rs11102972	1	109682244			Y	
rs7536292	1	109696216			Y	
rs17646665	1	109713574			Y	
rs7554936	1	149389113		Y		Kosoy et al. - 2009
rs2228145	1	151240043			Y	
rs4129267	1	152692888	Y			
rs2794520	1	157945440	Y			Sabatti et al. - 2009
rs3093077	1	157946260			Y	
rs2808630	1	157947492			Y	
rs1205	1	157948857			Y	
rs1130864	1	157949715			Y	
rs1800947	1	157950062	Y			Sabatti et al. - 2009
rs3737787	1	159276147			Y	
rs2774276	1	159278340			Y	
rs2073653	1	159279384			Y	
rs2516837	1	159281351			Y	
rs2774279	1	159284180			Y	
rs1040404	1	166426514		Y		Kosoy et al. - 2009
rs12120904	1	167744666			Y	
rs10919181	1	167744885			Y	
rs2420369	1	167752483			Y	
rs9332649	1	167757803			Y	
rs9332629	1	167764238			Y	
rs9332623	1	167764984			Y	
rs1557572	1	167775200			Y	
rs6016	1	167778744			Y	
rs9332595	1	167780979			Y	
rs6035	1	167788473			Y	
rs12120605	1	167789178			Y	
rs6427198	1	167789873			Y	
rs6427199	1	167790161			Y	
rs9332579	1	167791936			Y	
rs1894697	1	167793225			Y	

rs6427203	1	167795655			Y	
rs2298905	1	167805466			Y	
rs10489184	1	167815664			Y	
rs6703462	1	167817214			Y	
rs3753305	1	167820682			Y	
rs3917854	1	167825659			Y	
rs6128	1	167829528			Y	
rs1569471	1	167830754			Y	
rs3766122	1	167833451			Y	
rs3917793	1	167834169			Y	
rs760694	1	167835322			Y	
rs2420378	1	167836542			Y	
rs3917779	1	167837472			Y	
rs3917744	1	167844614			Y	
rs2076074	1	167845210			Y	
rs3917740	1	167845890			Y	
rs3917739	1	167846002			Y	
rs2235305	1	167846269			Y	
rs2235302	1	167846914			Y	
rs6131	1	167847509			Y	
rs3917727	1	167847882			Y	
rs3917709	1	167851169			Y	
rs2244529	1	167853656			Y	
rs3917698	1	167854741			Y	
rs3917688	1	167857704			Y	
rs3917676	1	167859135			Y	
rs3917657	1	167865366			Y	
rs3917651	1	167866686			Y	
rs6691334	1	167869007			Y	
rs12401978	1	167871409			Y	
rs1569476	1	167875541			Y	
rs1053381	1	169311944	Y			
rs4648304	1	184907476			Y	
rs689470	1	184907681			Y	
rs5275	1	184909681			Y	
rs2745557	1	184915844			Y	
rs689466	1	184917374			Y	
rs12042763	1	184918499			Y	
rs2143417	1	184920410			Y	
rs3024505	1	205006527			Y	
rs1554286	1	205010856			Y	
rs1518111	1	205011268			Y	
rs1800893	1	205013790			Y	
rs1800890	1	205015988			Y	
rs10494879	1	205018827			Y	
rs4844614	1	205941798	Y			Sabatti et al. - 2009
rs4951629	1	210853506		Y		Kosoy et al. - 2009
rs17465637	1	220890152	Y			Samani et al. - 2007
rs3009715	1	228262336			Y	

rs1321258	1	228268651		Y	
rs7531967	1	228273972		Y	
rs2883025	1	228274995		Y	
rs6684432	1	228276413		Y	
rs6541267	1	228277868		Y	
rs6673349	1	228278021		Y	
rs11122275	1	228279156		Y	
rs910820	1	228280942		Y	
rs16850871	1	228286093		Y	
rs16850878	1	228289864		Y	
rs10495293	1	228293369		Y	
rs12029443	1	228293831		Y	
rs7536663	1	228295422		Y	
rs12035297	1	228305444		Y	
rs7527527	1	228307777		Y	
rs1474839	1	228308023		Y	
rs10489614	1	228313992		Y	
rs7550262	1	228315362		Y	
rs2144296	1	228330458		Y	
rs7538126	1	228330685		Y	
rs1973943	1	228338686		Y	
rs11122383	1	228339548		Y	
rs6690424	1	228345515		Y	
rs4846908	1	228351679		Y	
rs1998064	1	228353751		Y	
rs2144300	1	228361539	Y		Aulchenko et al. - 2009, Willer et al. - 2008
rs4846914	1	228362314	Y		Kathiresan et al. - 2008, Willer et al. - 2008
rs10127775	1	228362412		Y	
rs2281722	1	228363666		Y	
rs609526	1	228375529		Y	
rs1124110	1	228377595		Y	
rs678050	1	228380534		Y	
rs611229	1	228390690		Y	
rs612577	1	228390987		Y	
rs6677241	1	228404510		Y	
rs4846928	1	228408528		Y	
rs4846934	1	228414355		Y	
rs16851142	1	228418757		Y	
rs1980533	1	228435573		Y	
rs3811488	1	228438506		Y	
rs3811486	1	228438684		Y	
rs11122474	1	228445305		Y	
rs6688678	1	228447910		Y	
rs3789630	1	228454121		Y	
rs901675	1	228456625		Y	
rs11588358	1	228475681		Y	
rs17761459	1	228476185		Y	
rs7533085	1	228476247		Y	
rs16851310	1	228476840		Y	

rs6698963	1	228481648			Y	
rs2273970	1	228481771			Y	
rs2273968	1	228481867			Y	
rs13728	1	228482841			Y	
rs1043908	1	228483917			Y	
rs1043944	1	228484183			Y	
rs7022	1	228484287			Y	
rs9431818	1	228484622			Y	
rs12757317	1	228485931			Y	
rs4846948	1	228486993			Y	
rs6541309	1	228488126			Y	
rs6687864	1	228488345			Y	
rs17762687	1	228488654			Y	
rs10864734	1	228488833			Y	
rs11122478	1	228490104			Y	
rs1998486	1	228491370			Y	
rs7536290	1	228903325			Y	
rs10864771	1	228904295			Y	
rs7079	1	228904954			Y	
rs2478523	1	228908132			Y	
rs3789664	1	228910047			Y	
rs2493131	1	228910128			Y	
rs2478545	1	228910744			Y	
rs6687360	1	228911615			Y	
rs699	1	228912417	Y			Willer et al. - 2008, Tsai et al. - 2009
rs4762	1	228912600	Y			Tsai et al. - 2009
rs11568028	1	228913867			Y	
rs2004776	1	228915325			Y	
rs3889728	1	228915454			Y	
rs1078499	1	228915719			Y	
rs7539020	1	228915813			Y	
rs2493134	1	228915982			Y	
rs3789678	1	228916105			Y	
rs2148582	1	228916422			Y	
rs11122580	1	228917667			Y	
rs2493137	1	228918739			Y	
rs7549009	1	228921694			Y	
rs16831987	1	235020410			Y	
rs12097647	1	235060752			Y	
rs7541539	1	235079167			Y	
rs10802564	1	235088700			Y	
rs10495387	1	235103880			Y	
rs10925257	1	235112783			Y	
rs12049580	1	235123532			Y	
rs6679990	1	235131853			Y	
rs316873	1	240409127		Y		Kosoy et al. - 2009
rs798443	2	7885726		Y		Kosoy et al. - 2009
rs7421394	2	14673800		Y		Kosoy et al. - 2009
rs6754295	2	21059688	Y			Aulchenko et al. - 2009

rs7557067	2	21061717	Y			Kathiresan et al. - 2009
rs13392272	2	21070995			Y	
rs4371387	2	21074638			Y	
rs1042034	2	21078786	Y			Knoblauch et al. - 2004
rs1801702	2	21078990			Y	
rs1801703	2	21079417	Y			Knoblauch et al. - 2004
rs1800479	2	21080888	Y			Knoblauch et al. - 2004
rs1801701	2	21082332			Y	
rs693	2	21085700	Y			Aulchenko et al. - 2009, Kathiresan et al. - 2008, Knoblauch et al. - 2004, Sabatti et al. - 2009, Willer et al. - 2008
rs533617	2	21087477			Y	
rs12713771	2	21089070			Y	
rs673548	2	21091049	Y			Aulchenko et al. - 2009, Knoblauch et al. - 2004, Sabatti et al. - 2009
rs12713844	2	21091918			Y	
rs11126598	2	21093869			Y	
rs12713911	2	21094043			Y	
rs12713956	2	21095010			Y	
rs10199768	2	21097505			Y	
rs11676704	2	21097863			Y	
rs12720791	2	21098122			Y	
rs7589300	2	21098265			Y	
rs1801700	2	21099318			Y	
rs679899	2	21104419	Y			Knoblauch et al. - 2004
rs12720832	2	21110218			Y	
rs520354	2	21113117	Y			Knoblauch et al. - 2004
rs1800478	2	21113641			Y	
rs12720796	2	21115503			Y	
rs531819	2	21117144			Y	
rs1367117	2	21117405	Y			Knoblauch et al. - 2004
rs512535	2	21121287	Y			Knoblauch et al. - 2004
rs17398765	2	21124256			Y	
rs1713222	2	21124828	Y			Sandhu et al. - 2008
rs7575840	2	21126995	Y			Kathiresan et al. - 2008
rs11893021	2	21127083			Y	
rs11892073	2	21129906			Y	
rs515135	2	21139562	Y			Kathiresan et al. - 2009, Willer et al. - 2008
rs562338	2	21141826	Y			Willer et al. - 2008
rs754523	2	21165196	Y			Willer et al. - 2008
rs4591370	2	21237247	Y			Sandhu et al. - 2008
rs538928	2	21242524	Y			Sandhu et al. - 2008
rs488507	2	21247194	Y			Sandhu et al. - 2008
rs506585	2	21250687	Y			Sandhu et al. - 2008
rs478442	2	21252721	Y			Sandhu et al. - 2008
rs704795	2	27569998			Y	
rs1260326	2	27584444	Y			Kathiresan et al. - 2009, Willer et al. - 2008, Sabatti et al. - 2009
rs780094	2	27594741	Y			Aulchenko et al. - 2009, Willer et al. - 2008, Knoblauch et al. - 2008
rs4665987	2	27609329			Y	
rs11127129	2	27930212	Y			Willer et al. - 2008
rs4666200	2	29391915		Y		Kosoy et al. - 2009

rs4670767	2	37794900		Y		Kosoy et al. - 2009
rs10180615	2	43892201			Y	
rs4148195	2	43893137			Y	
rs2278356	2	43893379			Y	
rs4148191	2	43896408			Y	
rs4953019	2	43896897			Y	
rs10205816	2	43897759			Y	
rs4148189	2	43901034			Y	
rs10439467	2	43901850			Y	
rs1864814	2	43902095			Y	
rs4245786	2	43902624			Y	
rs4073237	2	43903376			Y	
rs4148188	2	43904191			Y	
rs4289236	2	43907627			Y	
rs4148185	2	43909826			Y	
rs4953020	2	43916400			Y	
rs6756629	2	43918594	Y			Aulchenko et al. - 2009
rs11887534	2	43919751	Y			Kathiresan et al. - 2009
rs4148202	2	43921323			Y	
rs4148208	2	43923898			Y	
rs4299376	2	43926080			Y	
rs6544713	2	43927385	Y			Kathiresan et al. - 2009
rs4953023	2	43927504	Y			Kathiresan et al. - 2009
rs4148214	2	43932508			Y	
rs6709904	2	43933828			Y	
rs6755809	2	43949840			Y	
rs4148217	2	43952937			Y	
rs12468591	2	43953519			Y	
rs6756676	2	43953673			Y	
rs4245794	2	43954353			Y	
rs4953028	2	43955331			Y	
rs4148221	2	43955995			Y	
rs4148222	2	43956138			Y	
rs2954804	2	43961246			Y	
rs2954805	2	43962662			Y	
rs2954806	2	43962785			Y	
rs4953029	2	43963011			Y	
rs4352264	2	55941781			Y	
rs11694213	2	55942171			Y	
rs1346787	2	55946116			Y	
rs17278665	2	55949498			Y	
rs7563085	2	55956674			Y	
rs1346786	2	55961837			Y	
rs3791675	2	55964813			Y	
rs3791673	2	55967042			Y	
rs1430193	2	55974357			Y	
rs3791668	2	55975497			Y	
rs3791661	2	55983374			Y	
rs1430197	2	55988331			Y	

rs960993	2	55989732			Y	
rs3791656	2	55995226			Y	
rs2868440	2	56012150			Y	
rs13400937	2	79718431		Y		Kosoy et al. - 2009
rs260690	2	108946170		Y		Kosoy et al. - 2009
rs3917368	2	113299253			Y	
rs2853550	2	113303592			Y	
rs1143634	2	113306861			Y	
rs1143630	2	113308126			Y	
rs1143627	2	113310858			Y	
rs13032029	2	113316886			Y	
rs928940	2	113593966			Y	
rs1794066	2	113602821			Y	
rs579543	2	113606102			Y	
rs315951	2	113607057			Y	
rs397211	2	113608612			Y	
rs315948	2	113609727			Y	
rs11885498	2	113616648			Y	
rs315958	2	113617322			Y	
rs10496971	2	145486413		Y		Kosoy et al. - 2009
rs560887	2	169471394	Y			Sabatti et al. - 2009
rs2627037	2	179314783		Y		Kosoy et al. - 2009
rs1569175	2	200730199		Y		Kosoy et al. - 2009
rs2943634	2	226776324	Y			Samani et al. - 2007
rs10510228	3	2183832		Y		Kosoy et al. - 2009
rs4955316	3	30390616		Y		Kosoy et al. - 2009
rs5875	3	38139378			Y	
rs156265	3	38145814			Y	
rs2239621	3	38148737			Y	
rs7625290	3	38153294			Y	
rs172111	3	38162998			Y	
rs9809104	3	39121433		Y		Kosoy et al. - 2009
rs1799864	3	46374212	Y			
rs6548616	3	79482265		Y		Kosoy et al. - 2009
rs5868	3	120413481	Y			
rs12695382	3	120430861	Y			Willer et al. - 2008
rs12629908	3	122005406		Y		Kosoy et al. - 2009
rs1055419	3	126796272	Y			
rs9845457	3	137397166		Y		Kosoy et al. - 2009
rs734873	3	149233045		Y		Kosoy et al. - 2009
rs409742	3	149895055			Y	
rs10935724	3	149903943			Y	
rs718858	3	149918202			Y	
rs1492099	3	149920193			Y	
rs12721241	3	149924310			Y	
rs389566	3	149929072			Y	
rs385338	3	149931846			Y	
rs12721298	3	149934037			Y	
rs6801836	3	149938227			Y	

rs5182	3	149942085			Y	
rs275645	3	149947144			Y	
rs275644	3	149948230			Y	
rs275643	3	149948294			Y	
rs275642	3	149948381			Y	
rs427832	3	149949053			Y	
rs2261950	3	149950955			Y	
rs2030763	3	181447421		Y		Kosoy et al. - 2009
rs1513181	3	190057690		Y		Kosoy et al. - 2009
rs17132047	4	1697184	Y			
rs735794	4	2809236			Y	
rs9715869	4	2809457			Y	
rs12503220	4	2819940			Y	
rs16843487	4	2831761			Y	
rs17833172	4	2838312			Y	
rs2032470	4	2843633			Y	
rs7689864	4	2845194			Y	
rs2097081	4	2850907			Y	
rs3755885	4	2857739			Y	
rs3775068	4	2858239			Y	
rs2237004	4	2867072			Y	
rs9291090	4	5441538		Y		Kosoy et al. - 2009
rs10007810	4	41249121		Y		Kosoy et al. - 2009
rs1369093	4	73464055		Y		Kosoy et al. - 2009
rs385194	4	85528102		Y		Kosoy et al. - 2009
rs7657799	4	105594872		Y		Kosoy et al. - 2009
rs2702414	4	179636517		Y		Kosoy et al. - 2009
rs316598	5	2417626		Y		Kosoy et al. - 2009
rs870347	5	6898035		Y		Kosoy et al. - 2009
rs1801394	5	7923973			Y	
rs7730643	5	7928963			Y	
rs3776467	5	7929315			Y	
rs326122	5	7929611			Y	
rs7703033	5	7932950			Y	
rs162031	5	7933287			Y	
rs162033	5	7933835			Y	
rs3815743	5	7940109			Y	
rs2287780	5	7942304			Y	
rs1802059	5	7950319			Y	
rs8659	5	7953833			Y	
rs7715062	5	7959907			Y	
rs327588	5	7961359			Y	
rs162270	5	7962907			Y	
rs37369	5	35072872		Y		Kosoy et al. - 2009
rs6451722	5	43747135		Y		Kosoy et al. - 2009
rs33445	5	52317671			Y	
rs27377	5	52325134			Y	
rs6867926	5	52326815			Y	
rs35234	5	52327036			Y	

rs35235	5	52327994			Y	
rs35236	5	52328312			Y	
rs10071255	5	52336624			Y	
rs16880636	5	52336848			Y	
rs6859355	5	52337473			Y	
rs153141	5	52340753			Y	
rs1421940	5	52341770			Y	
rs3212418	5	52351929			Y	
rs7735277	5	52354673			Y	
rs246406	5	52354833			Y	
rs7705328	5	52359325			Y	
rs3212433	5	52372901			Y	
rs3212435	5	52373083			Y	
rs3212476	5	52379113			Y	
rs3212541	5	52393232			Y	
rs3212545	5	52394794			Y	
rs2303122	5	52415034			Y	
rs3212627	5	52416680			Y	
rs1109527	5	52422907			Y	
rs12518145	5	52427607			Y	
rs4492095	5	74353035			Y	
rs12655792	5	74367747			Y	
rs7708899	5	74368733			Y	
rs3761739	5	74667257			Y	
rs2303152	5	74677463			Y	
rs11742194	5	74682634			Y	
rs12654264	5	74684359	Y			Kathiresan et al. - 2008, two papers this year
rs10063134	5	74685405			Y	
rs3846662	5	74686840	Y			Aulchenko et al. - 2009
rs5908	5	74687955			Y	
rs7717396	5	74690535			Y	
rs3846663	5	74691482	Y			Kathiresan et al. - 2009
rs10474435	5	74693036			Y	
rs4703670	5	74696112			Y	
rs16872523	5	74701432			Y	
rs5744612	5	74885225			Y	
rs5744665	5	74911800			Y	
rs4703674	5	74918437			Y	
rs5744707	5	74926374			Y	
rs5744712	5	74927758			Y	
rs2287711	5	74934508			Y	
rs12657828	5	79121482		Y		Kosoy et al. - 2009
rs7721577	5	139985970			Y	
rs2569190	5	139993100			Y	
rs5744451	5	139994067			Y	
rs2569193	5	139995679			Y	
rs11959615	5	148181509			Y	
rs17778257	5	148184770			Y	
rs1042718	5	148187110			Y	

rs1042719	5	148187640			Y	
rs7702861	5	148192051			Y	
rs12652757	5	148194422			Y	
rs17108817	5	148196095			Y	
rs11957757	5	148196380			Y	
rs6556352	5	155404292		Y		Kosoy et al. - 2009
rs1501908	5	156330747	Y			Kathiresan et al. - 2009
rs1500127	5	165672560		Y		Kosoy et al. - 2009
rs6422347	5	177795689		Y		Kosoy et al. - 2009
rs1040045	6	4692158		Y		Kosoy et al. - 2009
rs2504853	6	12643097		Y		Kosoy et al. - 2009
rs7745461	6	22019595		Y		Kosoy et al. - 2009
rs915654	6	31646476			Y	
rs2239704	6	31648120			Y	
rs3093542	6	31648672			Y	
rs1041981	6	31648763			Y	
rs1799964	6	31650287			Y	
rs1800630	6	31650455			Y	
rs1800629	6	31651010			Y	
rs3093662	6	31652168			Y	
rs3093671	6	31654959			Y	
rs769178	6	31655493			Y	
rs2254287	6	33251926	Y			Willer et al. - 2008
rs2397060	6	51719429		Y		Kosoy et al. - 2009
rs192655	6	90574999		Y		Kosoy et al. - 2009
rs7646	6	151464443	Y			
rs4869817	6	154508510	Y			
rs3127596	6	160873025			Y	
rs7449650	6	160877104			Y	
rs9457933	6	160877759			Y	
rs3798220	6	160881127	Y			Kathiresan et al. - 2009
rs10755578	6	160889728			Y	
rs7761293	6	160890953			Y	
rs9364559	6	160896138			Y	
rs6415084	6	160900320			Y	
rs3798221	6	160918138			Y	
rs7771801	6	160928105			Y	
rs7770628	6	160938164			Y	
rs13202636	6	160949718			Y	
rs9355814	6	160951243			Y	
rs7759633	6	160990980			Y	
rs1367210	6	161002789			Y	
rs9346833	6	161004632			Y	
rs783149	6	161008908			Y	
rs1406888	6	161011583			Y	
rs4458655	6	163141782		Y		Kosoy et al. - 2009
rs1871428	6	168408609		Y		Kosoy et al. - 2009
rs731257	7	12635776		Y		Kosoy et al. - 2009
rs10244051	7	15030358	Y			Sabatti et al. - 2009

rs2191348	7	15030780	Y			Sabatti et al. - 2009
rs12670798	7	21573877	Y			Aulchenko et al. - 2009
rs7801617	7	22724607			Y	
rs7805828	7	22725087			Y	
rs2056576	7	22727727			Y	
rs12700386	7	22729534			Y	
rs1800796	7	22732771			Y	
rs2069835	7	22734396			Y	
rs2069840	7	22735097			Y	
rs2069852	7	22738785			Y	
rs10242595	7	22740756			Y	
rs32314	7	32145649		Y		Kosoy et al. - 2009
rs2330442	7	42346596		Y		Kosoy et al. - 2009
rs2237281	7	72483987			Y	
rs2240466	7	72494205	Y			Aulchenko et al. - 2009
rs2237279	7	72499785			Y	
rs714052	7	72502805	Y			Kathiresan et al. - 2009
rs12056034	7	72516581			Y	
rs2074754	7	72529690	Y			
rs2353082	7	72551104			Y	
rs17145721	7	72572016			Y	
rs17145738	7	72620810	Y			Kathiresan et al. - 2008, Willer et al. - 2008
rs14415	7	72622716			Y	
rs11974409	7	72627326			Y	
rs11760752	7	72660998			Y	
rs17145813	7	72684694			Y	
rs13230514	7	72686031			Y	
rs4717865	7	73092135		Y		Kosoy et al. - 2009
rs10499859	7	80096746	Y			Love-Gregory et al. - 2008
rs13438282	7	80100160	Y			Love-Gregory et al. - 2008
rs9784998	7	80100937	Y			Love-Gregory et al. - 2008
rs1049654	7	80113391	Y			Love-Gregory et al. - 2008
rs3211810	7	80114953	Y			Love-Gregory et al. - 2008
rs3211842	7	80120572	Y			Love-Gregory et al. - 2008
rs3211849	7	80121259	Y			Love-Gregory et al. - 2008
rs3211850	7	80121352	Y			Love-Gregory et al. - 2008
rs1054516	7	80122878	Y			Love-Gregory et al. - 2008
rs3173798	7	80123786	Y			Love-Gregory et al. - 2008
rs3211868	7	80124958	Y			Love-Gregory et al. - 2008
rs3211870	7	80125145	Y			Love-Gregory et al. - 2008
rs1358337	7	80126321	Y			Love-Gregory et al. - 2008
rs3211909	7	80132051	Y			Love-Gregory et al. - 2008
rs3211913	7	80132540	Y			Love-Gregory et al. - 2008
rs3173804	7	80137786	Y			Love-Gregory et al. - 2008
rs3211938	7	80138385	Y			Love-Gregory et al. - 2008
rs7755	7	80144207	Y			Love-Gregory et al. - 2008
rs13246513	7	80144687	Y			Love-Gregory et al. - 2008
rs10954737	7	83370983		Y		Kosoy et al. - 2009
rs854549	7	94764521			Y	

rs854551	7	94765613		Y	
rs854552	7	94765860		Y	
rs3917572	7	94767126		Y	
rs854555	7	94768327		Y	
rs3917550	7	94772509		Y	
rs662	7	94775382		Y	
rs3917538	7	94775829		Y	
rs3917527	7	94778194		Y	
rs2074354	7	94778523		Y	
rs2299257	7	94780701		Y	
rs854560	7	94784020		Y	
rs3917498	7	94784191		Y	
rs854561	7	94784953		Y	
rs3917490	7	94786777		Y	
rs2049649	7	94787265		Y	
rs2299260	7	94787473		Y	
rs2299261	7	94787599		Y	
rs854568	7	94787737		Y	
rs2299262	7	94787864		Y	
rs854569	7	94787991		Y	
rs2237583	7	94788113		Y	
rs3917481	7	94788701		Y	
rs854571	7	94792555		Y	
rs854572	7	94792632		Y	
rs854573	7	94792799		Y	
rs17166818	7	94796564		Y	
rs13228784	7	94798761		Y	
rs705308	7	97533299	Y		Kosoy et al. - 2009
rs4727479	7	100552387		Y	
rs6956010	7	100553755		Y	
rs2227631	7	100556258		Y	
rs2227667	7	100561469		Y	
rs2227672	7	100562406		Y	
rs2227692	7	100565964		Y	
rs1050955	7	100569180		Y	
rs757716	7	100574122		Y	
rs7803075	7	130392606	Y		Kosoy et al. - 2009
rs10236187	7	139093846	Y		Kosoy et al. - 2009
rs2373962	7	150311910		Y	
rs2373961	7	150312143		Y	
rs10277237	7	150314277		Y	
rs1799983	7	150327044		Y	
rs3918186	7	150333365		Y	
rs3918188	7	150333714		Y	
rs1808593	7	150339235		Y	
rs7830	7	150340504		Y	
rs2373929	7	150345745		Y	
rs6464211	7	151504786	Y		Kosoy et al. - 2009
rs10108270	8	4178201	Y		Kosoy et al. - 2009

rs7819412	8	11082571	Y			Kathiresan et al. - 2009
rs3943253	8	13403871		Y		Kosoy et al. - 2009
rs17091742	8	19840410			Y	
rs1800590	8	19840951	Y			Knoblauch et al. - 2004
rs1534649	8	19843921			Y	
rs1031045	8	19845392	Y			Knoblauch et al. - 2004
rs3779788	8	19847373			Y	
rs343	8	19855067			Y	
rs249	8	19855286			Y	
rs253	8	19855697	Y			Knoblauch et al. - 2004
rs264	8	19857460			Y	
rs268	8	19857809	Y			Knoblauch et al. - 2004
rs270	8	19857956	Y			Knoblauch et al. - 2004, Morabia et al. - 2003
rs285	8	19859469	Y			Knoblauch et al. - 2004, O'Donnell et al. - 2007
rs320	8	19863357	Y			Bauerfeind et al. - 2006, Knoblauch et al. - 2004
rs325	8	19863608			Y	
rs328	8	19864004	Y			Bauerfeind et al. - 2006, Kathiresan et al. - 2008, Knoblauch et al. - 2004, Morabia et al. - 2003
rs3289	8	19867472	Y			Knoblauch et al. - 2004
rs11570892	8	19867897			Y	
rs13702	8	19868772			Y	
rs2197089	8	19870653	Y			Willer et al. - 2008
rs10096633	8	19875201	Y			Aulchenko et al. - 2009, Sabatti et al. - 2009
rs17091872	8	19876257			Y	
rs10097668	8	19878009			Y	
rs12678919	8	19888502	Y			Kathiresan et al. - 2009, Willer et al. - 2008
rs10503669	8	19891970	Y			Willer et al. - 2008
rs17489268	8	19896325	Y			Samani et al. - 2007
rs17411031	8	19896590	Y			Samani et al. - 2007
rs2083637	8	19909455	Y			Aulchenko et al. - 2009
rs894210	8	19910123	Y			Kathiresan et al. - 2009
rs765547	8	19910554	Y			
rs6993414	8	19947198	Y			Willer et al. - 2008
rs6586891	8	19958878	Y			Willer et al. - 2008
rs1471939	8	28997224		Y		Kosoy et al. - 2009
rs12544346	8	86611868		Y		Kosoy et al. - 2009
rs9642976	8	103185175	Y			
rs7844723	8	122977684		Y		Kosoy et al. - 2009
rs13268726	8	126082315			Y	
rs3750236	8	126093234			Y	
rs9297708	8	126100653			Y	
rs10103274	8	126100936			Y	
rs9297709	8	126102377			Y	
rs4559257	8	126501167			Y	
rs4871593	8	126506812			Y	
rs2385110	8	126508511			Y	
rs10956245	8	126509542			Y	
rs4871594	8	126511247			Y	
rs2980874	8	126513970			Y	
rs2235108	8	126517971			Y	

rs17405319	8	126518588			Y	
rs13273254	8	126526735			Y	
rs17321515	8	126555591	Y			Aulchenko et al. - 2009, Kathiresan et al. - 2008, Willer et al. - 2008
rs2954029	8	126560154	Y			Kathiresan et al. - 2009, Willer et al. - 2008
rs6987702	8	126573908	Y			Aulchenko et al. - 2009
rs2001907	8	140310363		Y		Kosoy et al. - 2009
rs4736319	8	143981882			Y	
rs6471580	8	143983703			Y	
rs4736358	8	143984312			Y	
rs4536	8	143992763			Y	
rs6414	8	143993436			Y	
rs1408801	9	12662320		Y		Kosoy et al. - 2009
rs471364	9	15279578	Y			Kathiresan et al. - 2009
rs7041637	9	21951866			Y	
rs3731239	9	21964218			Y	
rs2811709	9	21970151			Y	
rs3731211	9	21976847			Y	
rs3731208	9	21977155			Y	
rs575427	9	22001477			Y	
rs545226	9	22002422			Y	
rs7044859	9	22008781	Y			Samani et al. - 2007
rs518394	9	22009673			Y	
rs496892	9	22014351	Y			
rs10965215	9	22019445			Y	
rs564398	9	22019547			Y	
rs7865618	9	22021005	Y			Samani et al. - 2007
rs17694493	9	22031998			Y	
rs12352425	9	22032086			Y	
rs13290048	9	22034804			Y	
rs11790231	9	22043591			Y	
rs1011970	9	22052134			Y	
rs10811650	9	22057593			Y	
rs4977756	9	22058652			Y	
rs16905599	9	22059144			Y	
rs12555547	9	22062040			Y	
rs1412832	9	22067543			Y	
rs16905613	9	22070363			Y	
rs10116277	9	22071397	Y			Helgadottir et al. - 2008
rs10965227	9	22071796			Y	
rs1547705	9	22072375			Y	
rs1333040	9	22073404	Y			Helgadottir et al. - 2007, 2008
rs10757274	9	22086055	Y			Helgadottir et al. - 2008
rs2383206	9	22105026	Y			McPherson et al. - 2007
rs2383207	9	22105959	Y			Helgadottir et al. - 2007, 2008
rs10757278	9	22114477	Y			Helgadottir et al. - 2007, 2008
rs1333049	9	22115503	Y			Helgadottir et al. - 2008, Willer et al. - 2008
rs10811661	9	22124094	Y			Helgadottir et al. - 2008
rs10511828	9	28618500		Y		Kosoy et al. - 2009
rs3793451	9	70849100		Y		Kosoy et al. - 2009

rs10519	9	98911550	Y		Levy et al. - 2007, O'Donnell et al. - 2007
rs1323432	9	103402758	Y		Willer et al. - 2008
rs2482430	9	106579173		Y	
rs2482432	9	106582993		Y	
rs4149338	9	106585724		Y	
rs1331924	9	106587443		Y	
rs2740485	9	106588314		Y	
rs4149337	9	106588639		Y	
rs4149336	9	106590460		Y	
rs2740484	9	106591001		Y	
rs4149335	9	106591547		Y	
rs2740481	9	106596238		Y	
rs2740479	9	106603258		Y	
rs1999431	9	106604497		Y	
rs2066716	9	106608526	Y		Bauerfeind et al. - 2006, Knoblauch et al. - 2004
rs2297409	9	106616067		Y	
rs2254884	9	106621570		Y	
rs4149313	9	106626574	Y		Bauerfeind et al. - 2006, Knoblauch et al. - 2004
rs2066717	9	106631299		Y	
rs2515629	9	106634185		Y	
rs2297398	9	106638705		Y	
rs6479282	9	106641333		Y	
rs12235875	9	106641446		Y	
rs2274873	9	106642499		Y	
rs2472448	9	106644328		Y	
rs914544	9	106647033		Y	
rs4149290	9	106651648		Y	
rs2253174	9	106660153		Y	
rs2230806	9	106660688	Y		Bauerfeind et al. - 2006, Knoblauch et al. - 2004
rs2230805	9	106663850		Y	
rs4149281	9	106666210		Y	
rs4743764	9	106668925		Y	
rs1929842	9	106673444		Y	
rs1929841	9	106673546		Y	
rs2000069	9	106675690		Y	
rs1999429	9	106680607	Y		Knoblauch et al. - 2004
rs3904997	9	106685830		Y	
rs11789603	9	106686840		Y	
rs4149268	9	106687041	Y		Willer et al. - 2008
rs3890182	9	106687476	Y		Kathiresan et al. - 2008, 2008
rs3847300	9	106688251		Y	
rs3847303	9	106688473		Y	
rs3905000	9	106696891	Y		Aulchenko et al. - 2009
rs3847305	9	106697074		Y	
rs2740491	9	106698206		Y	
rs1883025	9	106704122	Y		Kathiresan et al. - 2009
rs3758294	9	106704636		Y	
rs2740487	9	106704782		Y	
rs1800978	9	106705799	Y		Knoblauch et al. - 2004

rs2740486	9	106706334			Y	
rs3905001	9	106707885			Y	
rs2575879	9	106708646	Y			Bauerfeind et al. - 2006, Knoblauch et al. - 2004
rs2777794	9	106708707			Y	
rs4149265	9	106712319			Y	
rs2437818	9	106715243			Y	
rs4149263	9	106717110			Y	
rs4149261	9	106718099			Y	
rs2515618	9	106718618			Y	
rs10991410	9	106719021			Y	
rs10512335	9	106720460			Y	
rs2515616	9	106721816			Y	
rs2791952	9	106722338			Y	
rs2515614	9	106724139			Y	
rs10991413	9	106724240			Y	
rs10991414	9	106725410			Y	
rs2472377	9	106726925			Y	
rs2472507	9	106727014			Y	
rs10991416	9	106727519			Y	
rs1800977	9	106730271	Y			Knoblauch et al. - 2004
rs2422493	9	106730816	Y			Bauerfeind et al. - 2006, Knoblauch et al. - 2004
rs2472493	9	106735669			Y	
rs4452883	9	116819395			Y	
rs3789875	9	116835109			Y	
rs12347433	9	116837418			Y	
rs10513300	9	119170027		Y		Kosoy et al. - 2009
rs10759930	9	119501442			Y	
rs2770150	9	119502960			Y	
rs1927907	9	119512585			Y	
rs5030717	9	119513655			Y	
rs2149356	9	119514020			Y	
rs1927906	9	119519936			Y	
rs7037117	9	119523484			Y	
rs913930	9	119523830			Y	
rs7045953	9	119525616			Y	
rs10759934	9	119528817			Y	
rs4837960	9	124163792			Y	
rs16911652	9	124168568			Y	
rs1213266	9	124176705			Y	
rs10306135	9	124177516			Y	
rs3842788	9	124180027			Y	
rs10306148	9	124183205			Y	
rs4273915	9	124185150			Y	
rs3842798	9	124185564			Y	
rs1771782	9	124189312			Y	
rs10306202	9	124199342			Y	
rs12001465	9	131536497			Y	
rs4636306	9	131536838			Y	
rs11790822	9	131537732			Y	

rs2302821	9	131541702			Y	
rs11788591	9	131542622			Y	
rs10739757	9	131547731			Y	
rs10448290	9	131550187			Y	
rs10988493	9	131551777			Y	
rs7873087	9	131555776			Y	
rs10988497	9	131559219			Y	
rs3844048	9	131562090			Y	
rs6478915	9	131564271			Y	
rs2073821	9	134922943		Y		Kosoy et al. - 2009
rs501120	10	44073873	Y			
rs3780894	10	45194666			Y	
rs2228064	10	45198056			Y	
rs4948671	10	45200062			Y	
rs7918542	10	45216257			Y	
rs7099684	10	45217161			Y	
rs7917687	10	45217867			Y	
rs3780900	10	45218919			Y	
rs2115819	10	45221095			Y	
rs10900213	10	45224720			Y	
rs11239524	10	45232777			Y	
rs3780901	10	45237382			Y	
rs3780902	10	45239063			Y	
rs17157784	10	45240820			Y	
rs3780906	10	45241988			Y	
rs2279435	10	45243534			Y	
rs1565097	10	45244262			Y	
rs1487562	10	45248828			Y	
rs901681	10	45254535			Y	
rs2291427	10	45256230			Y	
rs7896431	10	45257429			Y	
rs10751383	10	45257998			Y	
rs2288619	10	45259824			Y	
rs3793791	10	50511710		Y		Kosoy et al. - 2009
rs4746136	10	74971000		Y		Kosoy et al. - 2009
rs11185790	10	91362513	Y			Sabatti et al. - 2009
rs1075374	10	91366955	Y			Sabatti et al. - 2009
rs4244285	10	96531606	Y			
rs1135654	10	96721871	Y			
rs4918842	10	115306802		Y		Kosoy et al. - 2009
rs4880436	10	134500093		Y		Kosoy et al. - 2009
rs10839880	11	7806892		Y		Kosoy et al. - 2009
rs1837606	11	15794713		Y		Kosoy et al. - 2009
rs2946788	11	23967106		Y		Kosoy et al. - 2009
rs2070850	11	46698071			Y	
rs2070852	11	46701501			Y	
rs3136456	11	46702437			Y	
rs2167079	11	47226831	Y			Sabatti et al. - 2009
rs3758668	11	47227351			Y	

rs10838681	11	47231640			Y	
rs7120118	11	47242866	Y			Sabatti et al. - 2009
rs326217	11	47259851			Y	
rs3816724	11	47268621			Y	
rs10838687	11	47269468			Y	
rs10838689	11	47291100			Y	
rs10838692	11	47301676			Y	
rs11039183	11	47303516			Y	
rs3729953	11	47315573			Y	
rs7395662	11	48475469	Y			Aulchenko et al. - 2009
rs16906158	11	49127350			Y	
rs2299650	11	49130583			Y	
rs202719	11	49149266			Y	
rs202700	11	49173275			Y	
rs202676	11	49184196			Y	
rs639572	11	49190997			Y	
rs12099184	11	49195312			Y	
rs174537	11	61309256	Y			Sabatti et al. - 2009
rs102275	11	61314379	Y			Kathiresan et al. - 2009, Sabatti et al. - 2009
rs174546	11	61326406	Y			Sabatti et al. - 2009
rs174547	11	61327359	Y			Kathiresan et al. - 2009
rs174556	11	61337211	Y			Sabatti et al. - 2009
rs174570	11	61353788	Y			Aulchenko et al. - 2009
rs1535	11	61354548	Y			Sabatti et al. - 2009
rs2845573	11	61358484			Y	
rs174575	11	61358579			Y	
rs2524299	11	61361358			Y	
rs174578	11	61362075			Y	
rs174602	11	61380990			Y	
rs498793	11	61381281			Y	
rs526126	11	61381461			Y	
rs174605	11	61383497			Y	
rs174616	11	61385698			Y	
rs1000778	11	61411881			Y	
rs174455	11	61412693			Y	
rs174468	11	61420267			Y	
rs13966	11	61421568			Y	
rs11227699	11	66655068		Y		Kosoy et al. - 2009
rs1447352	11	92362409	Y			Sabatti et al. - 2009
rs17293642	11	102160138			Y	
rs559518	11	102161289			Y	
rs11225422	11	102161528			Y	
rs2408489	11	102162896			Y	
rs7125062	11	102168713			Y	
rs7125320	11	102168918			Y	
rs11225426	11	102169801			Y	
rs996999	11	102172273			Y	
rs10488	11	102173232			Y	
rs475007	11	102174522			Y	

rs1144393	11	102174619			Y	
rs1155764	11	102177659			Y	
rs484915	11	102178458			Y	
rs7926920	11	102203934			Y	
rs12277638	11	102206435			Y	
rs650108	11	102213997			Y	
rs646910	11	102214732			Y	
rs522616	11	102220258			Y	
rs613804	11	102225631			Y	
rs626750	11	102226155			Y	
rs486394	11	116031532	Y			Willer et al. - 2008
rs2000571	11	116090743	Y			Willer et al. - 2008
rs12272004	11	116108934	Y			Aulchenko et al. - 2009
rs10466588	11	116115459			Y	
rs9326246	11	116116943			Y	
rs180349	11	116117037			Y	
rs17092646	11	116117174			Y	
rs180346	11	116117869			Y	
rs28927680	11	116124283	Y			Kathiresan et al. - 2008
rs17440396	11	116126980			Y	
rs12421652	11	116131085			Y	
rs2075295	11	116133611			Y	
rs964184	11	116154127	Y			Kathiresan et al. - 2009, Willer et al. - 2008
rs12286037	11	116157417	Y			Willer et al. - 2008
rs4417316	11	116157511			Y	
rs2072560	11	116167036	Y			Knoblauch et al. - 2004
rs3135506	11	116167617	Y			Hubayek et al. - 2009, Knoblauch et al. - 2004
rs651821	11	116167789			Y	
rs662799	11	116168917	Y			Hubayek et al. - 2009, Knoblauch et al. - 2004, Willer et al. - 2008
rs17120035	11	116169061			Y	
rs1729410	11	116170871			Y	
rs633389	11	116172547			Y	
rs6589567	11	116175886			Y	
rs2266789	11	116180027	Y			Knoblauch et al. - 2004
rs7396851	11	116189374	Y			
rs675	11	116196885	Y			Aulchenko et al. - 2009, Knoblauch et al. 2004
rs5100	11	116197904			Y	
rs5092	11	116198674	Y			Knoblauch et al. - 2004
rs5091	11	116199215			Y	
rs2542051	11	116202948	Y			Knoblauch et al. - 2004
rs2854116	11	116205379	Y			Knoblauch et al. - 2004
rs4520	11	116206745	Y			Knoblauch et al. - 2004
rs5128	11	116208850	Y			Hubayek et al. - 2009, Knoblauch et al. - 2004
rs5081	11	116211556	Y			Knoblauch et al. - 2004
rs5070	11	116213230	Y			Bauerfeind et al. - 2006, Knoblauch et al. - 2004
rs1799837	11	116213468	Y			Knoblauch et al. - 2004
rs10892044	11	116272109	Y			Kathiresan et al. - 2009
rs948028	11	120149657		Y		Kosoy et al. - 2009
rs2416791	12	11592755		Y		Kosoy et al. - 2009

rs1513056	12	17299059		Y		Kosoy et al. - 2009
rs214678	12	45963217		Y		Kosoy et al. - 2009
rs772262	12	54450001		Y		Kosoy et al. - 2009
rs10778213	12	102019281			Y	
rs2070586	12	107801849		Y		Kosoy et al. - 2009
rs7953327	12	108466918			Y	
rs888194	12	108471301			Y	
rs2241201	12	108476515			Y	
rs877710	12	108478359			Y	
rs11067253	12	108482361			Y	
rs754180	12	108488451			Y	
rs17183465	12	108514133			Y	
rs1503767	12	117373871		Y		Kosoy et al. - 2009
rs2650000	12	119873345	Y			Kathiresan et al. - 2009, Sabatti et al. - 2009
rs7310409	12	119909244			Y	
rs1169300	12	119915608			Y	
rs11057811	12	123823319			Y	
rs838878	12	123826598			Y	
rs701106	12	123831589			Y	
rs838894	12	123835449			Y	
rs2272310	12	123837631			Y	
rs9919713	12	123838716			Y	
rs5888	12	123850701	Y			Morabia et al. - 2003
rs989892	12	123853169			Y	
rs838900	12	123859550			Y	
rs11057820	12	123862917			Y	
rs10846739	12	123863377			Y	
rs4765615	12	123865783			Y	
rs2343394	12	123866504			Y	
rs1902569	12	123869957			Y	
rs10846745	12	123878565			Y	
rs11057841	12	123882696			Y	
rs12581963	12	123883078			Y	
rs10744181	12	123883183			Y	
rs10773107	12	123883344			Y	
rs10744182	12	123883535			Y	
rs7954519	12	123886108			Y	
rs4765623	12	123886803			Y	
rs12582221	12	123889193			Y	
rs7306660	12	123893337			Y	
rs11057852	12	123894779			Y	
rs10773111	12	123898908			Y	
rs11615630	12	123907823			Y	
rs7139401	12	123911401			Y	
rs4379922	12	123917069			Y	
rs10846760	12	123917725			Y	
rs6488950	12	123920056			Y	
rs6490506	13	19472373	Y			
rs9319336	13	26522356		Y		Kosoy et al. - 2009

rs4076128	13	30203143			Y	
rs4073261	13	30204646			Y	
rs4769055	13	30207830			Y	
rs4075131	13	30208919			Y	
rs9578196	13	30209557			Y	
rs10507391	13	30210096			Y	
rs12429692	13	30210178			Y	
rs4769873	13	30210689			Y	
rs4503649	13	30212174			Y	
rs3885907	13	30212455			Y	
rs10162089	13	30214738			Y	
rs12431114	13	30215452			Y	
rs4254165	13	30215878			Y	
rs3803277	13	30216308			Y	
rs9671124	13	30222253			Y	
rs9579648	13	30225032			Y	
rs9315048	13	30225840			Y	
rs9508833	13	30229154			Y	
rs4769875	13	30229915			Y	
rs4769058	13	30231431			Y	
rs9315051	13	30234177			Y	
rs3935644	13	30235640			Y	
rs4769060	13	30235877			Y	
rs9670198	13	30242453			Y	
rs9579653	13	30242712			Y	
rs4319601	13	30245122			Y	
rs4769062	13	30245221			Y	
rs4238139	13	30245589			Y	
rs7997709	13	33745737		Y		Kosoy et al. - 2009
rs9530435	13	74891888		Y		Kosoy et al. - 2009
rs9521430	13	109027997	Y			
rs9522149	13	110625168		Y		Kosoy et al. - 2009
rs1760921	14	19887971		Y		Kosoy et al. - 2009
rs2357442	14	51677717		Y		Kosoy et al. - 2009
rs1950993	14	57308440		Y		Kosoy et al. - 2009
rs8021730	14	66956534		Y		Kosoy et al. - 2009
rs946918	14	82542621		Y		Kosoy et al. - 2009
rs10142528	14	91400543			Y	
rs7155616	14	91402475			Y	
rs3825687	14	91403364			Y	
rs929609	14	91406742			Y	
rs929610	14	91406906			Y	
rs2430340	14	91407527			Y	
rs7143288	14	91407630			Y	
rs917907	14	91409400			Y	
rs2267989	14	91409603			Y	
rs10484030	14	91412338			Y	
rs2498847	14	91414053			Y	
rs2430343	14	91415418			Y	

rs966382	14	91421405			Y	
rs1006856	14	91423715			Y	
rs2430353	14	91424870			Y	
rs2267994	14	91426699			Y	
rs2243400	14	91428673			Y	
rs2267996	14	91429478			Y	
rs917908	14	91433539			Y	
rs2430363	14	91434804			Y	
rs741198	14	91437164			Y	
rs2300137	14	91439777			Y	
rs1861085	14	91440987			Y	
rs2160079	14	91444158			Y	
rs2160080	14	91444240			Y	
rs2249946	14	91453895			Y	
rs17732513	14	91456132			Y	
rs2430368	14	91457294			Y	
rs2284340	14	91458800			Y	
rs10149504	14	91463925			Y	
rs2430373	14	91464063			Y	
rs12586793	14	91471387			Y	
rs2474028	14	91473413			Y	
rs3783937	14	91477446			Y	
rs13379081	14	91478062			Y	
rs3814835	14	91484133			Y	
rs2474030	14	91485073			Y	
rs200354	14	98445074		Y		Kosoy et al. - 2009
rs3784230	14	104750100		Y		Kosoy et al. - 2009
rs12439433	15	34007327		Y		Kosoy et al. - 2009
rs2624265	15	36935941	Y			Sabatti et al. - 2009
rs4775041	15	56461987	Y			Aulchenko et al. - 2009, Willer et al. - 2008
rs10468017	15	56465804	Y			Kathiresan et al. - 2009, Willer et al. - 2008
rs1532085	15	56470658	Y			Aulchenko et al. - 2009, Sabatti et al. - 2009
rs723967	15	56483592	Y			Knoblauch et al. - 2004
rs792902	15	56503336			Y	
rs375372	15	56503907			Y	
rs17190503	15	56504314			Y	
rs16940262	15	56504720			Y	
rs1373657	15	56505054			Y	
rs417344	15	56505211			Y	
rs7163052	15	56505522			Y	
rs7170361	15	56506290			Y	
rs11071381	15	56508906			Y	
rs1800588	15	56510967	Y			Bauerfeind et al. - 2006, Knoblauch et al. - 2004, Kathiresan et al. - 2008, 2008
rs2070895	15	56511231	Y			Bauerfeind et al. - 2006, Kathiresan et al. - 2009, Morabia et al. - 2003
rs11631342	15	56511675			Y	
rs6494005	15	56511816			Y	
rs8034802	15	56512084			Y	
rs261334	15	56514036			Y	
rs261332	15	56514617	Y			Kaess et al. - 2011, Willer et al. - 2008

rs11636642	15	56514793		Y	
rs17190517	15	56516436		Y	
rs17301739	15	56517931		Y	
rs12914035	15	56518181		Y	
rs17301746	15	56518687		Y	
rs166362	15	56520804		Y	
rs261337	15	56522741		Y	
rs497381	15	56524330		Y	
rs17269278	15	56528045		Y	
rs485671	15	56528426		Y	
rs572410	15	56528676		Y	
rs8041059	15	56531001		Y	
rs3825776	15	56534122		Y	
rs934297	15	56540026		Y	
rs11631153	15	56541284		Y	
rs17301781	15	56542168		Y	
rs12902119	15	56543846		Y	
rs12148704	15	56545698		Y	
rs12593008	15	56548267		Y	
rs7182229	15	56552475		Y	
rs1968685	15	56555046		Y	
rs12914552	15	56556753		Y	
rs4775053	15	56556875		Y	
rs17190594	15	56558376		Y	
rs1869137	15	56565704		Y	
rs11856322	15	56567305		Y	
rs1973028	15	56578646		Y	
rs16940372	15	56582805		Y	
rs16940379	15	56583537		Y	
rs1869144	15	56584592	Y		Knoblauch et al. - 2004
rs1869145	15	56584791		Y	
rs12324517	15	56588058		Y	
rs4613005	15	56588526		Y	
rs7174210	15	56596785		Y	
rs2899632	15	56599161		Y	
rs8028759	15	56599491		Y	
rs9652472	15	56606075		Y	
rs12899928	15	56606843		Y	
rs1968689	15	56610077		Y	
rs4775073	15	56613924		Y	
rs17301822	15	56614915		Y	
rs11071387	15	56615876		Y	
rs8035006	15	56616507		Y	
rs1971544	15	56616743		Y	
rs2233741	15	56618655		Y	
rs6078	15	56621285	Y		Bauerfeind et al. - 2006, Knoblauch et al. - 2004
rs2242062	15	56621683		Y	
rs690	15	56622033	Y		Bauerfeind et al. - 2006, Knoblauch et al. - 2004
rs16940472	15	56622609		Y	

rs11852861	15	56623600			Y	
rs12592139	15	56624230			Y	
rs12592127	15	56624381			Y	
rs4775075	15	56625056			Y	
rs6082	15	56625249	Y			Bauerfeind et al. - 2006, Knoblauch et al. - 2004
rs6084	15	56625330	Y			Bauerfeind et al. - 2006, Knoblauch et al. - 2004, Morabia et al. - 2003
rs2242064	15	56626564			Y	
rs7165654	15	56627331			Y	
rs7166788	15	56629028			Y	
rs11632627	15	56630175			Y	
rs871804	15	56631506			Y	
rs7178362	15	56634555			Y	
rs7170227	15	56641879			Y	
rs17301864	15	56642504			Y	
rs3829460	15	56645237			Y	
rs6074	15	56648255	Y			Bauerfeind et al. - 2006, Knoblauch et al. - 2004
rs17228212	15	65245693	Y			Samani et al. - 2007
rs2899826	15	72521553		Y		Kosoy et al. - 2009
rs8035124	15	89906712		Y		Kosoy et al. - 2009
rs9806366	15	99620480	Y			
rs4984913	16	680467		Y		Kosoy et al. - 2009
rs4781011	16	10882812		Y		Kosoy et al. - 2009
rs2269793	16	19180409		Y		Kosoy et al. - 2009
rs12596776	16	55476849	Y			Willer et al. - 2008
rs9989419	16	55542640	Y			Kathiresan et al. - 2009
rs173539	16	55545545	Y			Kathiresan et al. - 2009
rs6499863	16	55549518			Y	
rs12708967	16	55550712			Y	
rs3764261	16	55550825	Y			Kaess et al. - 2011, Willer et al. - 2008, Thompson et al. - 2007, Sabatti et al. - 2009
rs4783961	16	55552395	Y			Bauerfeind et al. - 2006, Knoblauch et al. - 2004, Thompson et al. - 2007
rs4783962	16	55552539			Y	
rs1800776	16	55552735	Y			Bauerfeind et al. - 2006, Knoblauch et al. - 2004, Boekholdt et al. - 2003
rs711752	16	55553712	Y			Boekholdt et al. - 2003, Knoblauch et al. - 2004
rs708272	16	55553789	Y			Boekholdt et al. - 2003, Knoblauch et al. - 2004, Thompson et al. - 2007
rs1864163	16	55554734	Y			Willer et al. - 2008
rs820299	16	55557785			Y	
rs11076174	16	55560647			Y	
rs7205804	16	55562390	Y			Bauerfeind et al. - 2006, Morabia et al. - 2003
rs1532625	16	55562802	Y			Bauerfeind et al. - 2006, Knoblauch et al. - 2004, Boekholdt et al. - 2003
rs1532624	16	55562980	Y			Aulchenko et al. - 2009
rs11076175	16	55563879			Y	
rs289714	16	55564952	Y			Boekholdt et al. - 2003, Kathiresan et al. - 2009
rs289717	16	55566889			Y	
rs289718	16	55567433	Y			Boekholdt et al. - 2003, Knoblauch et al. - 2004
rs12720898	16	55568744			Y	
rs891144	16	55569437			Y	
rs5880	16	55572592	Y			Boekholdt et al. - 2003, Knoblauch et al. - 2004

rs1800774	16	55573046			Y	
rs5882	16	55573593	Y			Boekholdt et al. - 2003
rs1800777	16	55574820	Y			Boekholdt et al. - 2003
rs1801706	16	55575163	Y			Bauerfeind et al. - 2006, Knoblauch et al. - 2004, Boekholdt et al. - 2003
rs289742	16	55575263			Y	
rs289744	16	55575603			Y	
rs289745	16	55577033			Y	
rs1566439	16	55582163	Y			Willer et al. - 2008
rs818386	16	63964209		Y		Kosoy et al. - 2009
rs2271293	16	66459571	Y			Aulchenko et al. - 2009, Kathiresan et al. - 2009
rs5923	16	66531454	Y			Bauerfeind et al. - 2006, Knoblauch et al. - 2004
rs4986970	16	66533821	Y			Bauerfeind et al. - 2006, Knoblauch et al. - 2004
rs1109166	16	66534883	Y			Knoblauch et al. - 2004
rs255049	16	66570972	Y			Sabatti et al. - 2009
rs255052	16	66582496	Y			Willer et al. - 2008
rs2966849	16	83741183		Y		Kosoy et al. - 2009
rs1879488	17	1348363		Y		Kosoy et al. - 2009
rs9891572	17	2375258	Y			Sabatti et al. - 2009
rs4790210	17	4475587			Y	
rs916055	17	4481583			Y	
rs7217186	17	4486141			Y	
rs2664593	17	4491881			Y	
rs748694	17	4496938			Y	
rs3730340	17	11985108	Y			
rs4795050	17	23103032			Y	
rs4795051	17	23103624			Y	
rs11653716	17	23108659			Y	
rs7208775	17	23109235			Y	
rs9906835	17	23113501			Y	
rs2297516	17	23119857			Y	
rs9797244	17	23121258			Y	
rs2248814	17	23124448			Y	
rs11080344	17	23128638			Y	
rs4795067	17	23130802			Y	
rs3729508	17	23133157			Y	
rs944725	17	23133698			Y	
rs17722851	17	23134963			Y	
rs3794764	17	23135555			Y	
rs16966563	17	23140076			Y	
rs8072199	17	23140975			Y	
rs3794766	17	23146048			Y	
rs3730013	17	23150045			Y	
rs10459953	17	23151645			Y	
rs2779248	17	23151959			Y	
rs11080358	17	23154723			Y	
rs12150211	17	23158559			Y	
rs2734299	17	23159085			Y	
rs11867200	17	29600082			Y	
rs1860190	17	29600686			Y	

rs2530797	17	29610207			Y	
rs991804	17	29611838			Y	
rs8067673	17	35954741			Y	
rs3136695	17	35962217			Y	
rs2023906	17	35967212			Y	
rs588019	17	35968633			Y	
rs2290065	17	35969886			Y	
rs3136685	17	35973325			Y	
rs7218773	17	42679210			Y	
rs2056131	17	42688742			Y	
rs8077375	17	42691151			Y	
rs16941751	17	42691506			Y	
rs4525555	17	42692948			Y	
rs3892084	17	42695420			Y	
rs7219925	17	42696639			Y	
rs8074094	17	42703020			Y	
rs3851806	17	42705918			Y	
rs7209700	17	42706117			Y	
rs11868894	17	42708188			Y	
rs7217214	17	42710226			Y	
rs2292867	17	42712488			Y	
rs2292699	17	42717294			Y	
rs2292700	17	42717422			Y	
rs2292864	17	42722680			Y	
rs12603582	17	42732576			Y	
rs3744453	17	42737869			Y	
rs3809863	17	42740011			Y	
rs12600865	17	42748983			Y	
rs6504833	17	42751086			Y	
rs2033111	17	51143279		Y		Kosoy et al. - 2009
rs2071409	17	53703229			Y	
rs8082134	17	53707625			Y	
rs2243828	17	53713883			Y	
rs2107545	17	53715117			Y	
rs12452417	17	53722201			Y	
rs4459609	17	58902680			Y	
rs4305	17	58911961			Y	
rs4309	17	58913655			Y	
rs4311	17	58914495			Y	
rs4362	17	58927493			Y	
rs9894286	17	58929764			Y	
rs4461142	17	58931780			Y	
rs4267385	17	58937488			Y	
rs11652805	17	60417613		Y		Kosoy et al. - 2009
rs10512572	17	67023694		Y		Kosoy et al. - 2009
rs2125345	17	71293786		Y		Kosoy et al. - 2009
rs4798812	18	9410504		Y		Kosoy et al. - 2009
rs4797824	18	13874513	Y			
rs4800105	18	17905980		Y		Kosoy et al. - 2009

rs12962061	18	45332909			Y	
rs2097055	18	45349485			Y	
rs8093249	18	45351396			Y	
rs11875600	18	45353184			Y	
rs12970066	18	45361150			Y	
rs6507931	18	45367006			Y	
rs9958734	18	45372396			Y	
rs17798991	18	45375314			Y	
rs11664186	18	45375696			Y	
rs12606544	18	45377235			Y	
rs4939585	18	45377761			Y	
rs8099385	18	45380883			Y	
rs16950701	18	45382323			Y	
rs4939883	18	45421212	Y			Aulchenko et al. - 2009, Kathiresan et al. - 2009, Willer et al. - 2008
rs2156552	18	45435666	Y			Kathiresan et al. - 2008, Willer et al. - 2008
rs7238445	18	48035542		Y		Kosoy et al. - 2009
rs881728	18	57484088		Y		Kosoy et al. - 2009
rs4891825	18	66018643		Y		Kosoy et al. - 2009
rs874299	18	73185272		Y		Kosoy et al. - 2009
rs1862513	19	7639793			Y	
rs3745367	19	7640511			Y	
rs3745369	19	7641475			Y	
rs1423096	19	7645177			Y	
rs10401670	19	7648802			Y	
rs2967605	19	8375738	Y			Kathiresan et al. - 2009
rs5030340	19	10243281			Y	
rs5030390	19	10243537			Y	
rs5030391	19	10243637			Y	
rs281432	19	10251658			Y	
rs5030400	19	10256796			Y	
rs3093032	19	10257336			Y	
rs3093030	19	10258403			Y	
rs1529729	19	11024562	Y			Kathiresan et al. - 2008
rs6511720	19	11063306	Y			Aulchenko et al. - 2009, Kathiresan et al. - 2008,2009, Willer et al. - 2008
rs8110695	19	11067530			Y	
rs2228671	19	11071912	Y			Aulchenko et al. - 2009, Bauerfeind et al. - 2006, Knoblauch et al. - 2004, Morabia et al. - 2003, Sandhu et al. - 2008
rs2569556	19	11079965			Y	
rs5930	19	11085265	Y			Bauerfeind et al. - 2006, Knoblauch et al. - 2004, Morabia et al. 2003
rs2738446	19	11088326	Y			Kathiresan et al. - 2009
rs688	19	11088602	Y			Kathiresan et al. - 2008, Wellcome et al. - 2007
rs5925	19	11091881	Y			Bauerfeind et al. - 2006, Knoblauch et al. - 2004
rs5927	19	11094941	Y			Boekholdt et al. - 2003, Knoblauch et al. - 2004
rs2569540	19	11099239			Y	
rs2738459	19	11099473			Y	
rs2304182	19	11099877			Y	
rs6413503	19	11102904			Y	
rs3826810	19	11103133			Y	
rs2738464	19	11103307			Y	

rs2738465	19	11103496			Y	
rs1433099	19	11103658	Y			Bauerfeind et al. - 2006, Knoblauch et al. - 2004, Morabia et al. 2003
rs10412238	19	19179158			Y	
rs7250569	19	19185443			Y	
rs2228603	19	19190924	Y			Kathiresan et al. - 2008, Willer et al. - 2008
rs1859055	19	19202160			Y	
rs10424369	19	19206356			Y	
rs10413995	19	19214000	Y			
rs10402294	19	19215937			Y	
rs4808931	19	19219672			Y	
rs10407933	19	19225319			Y	
rs10422819	19	19504028			Y	
rs1036215	19	19518198			Y	
rs16996148	19	19519472	Y			Aulchenko et al. - 2009, Kathiresan et al. - 2008, Willer et al. - 2008
rs17216525	19	19523220	Y			Kathiresan et al. - 2009
rs4145657	19	19523832			Y	
rs892023	19	19524850			Y	
rs8103992	19	19526643			Y	
rs16996174	19	19569551			Y	
rs1476459	19	19581028			Y	
rs10500212	19	19584215			Y	
rs1054685	19	19595539			Y	
rs873870	19	19599554			Y	
rs2304130	19	19650528	Y			Aulchenko et al. - 2009
rs8113143	19	38344087		Y		Kosoy et al. - 2009
rs4803750	19	49939467	Y			Sandhu et al. - 2008
rs10402271	19	50021054	Y			Sandhu et al. - 2008, Willer et al. - 2008
rs283814	19	50081064			Y	
rs7254892	19	50081436			Y	
rs6857	19	50084094			Y	
rs157580	19	50087106	Y			Aulchenko et al. - 2009, Sabatti et al. - 2009
rs2075650	19	50087459	Y			Aulchenko et al. - 2009, Sandhu et al. - 2008
rs8106922	19	50093506			Y	
rs405509	19	50100676	Y			Bauerfeind et al. - 2006, Knoblauch et al. - 2004
rs440446	19	50101007	Y			Bauerfeind et al. - 2006, Knoblauch et al. - 2004
rs769449	19	50101842			Y	
rs7412	19	50103919	Y			Bauerfeind et al. - 2006, Knoblauch et al. - 2004
rs439401	19	50106291	Y			Aulchenko et al. - 2009
rs4420638	19	50114786	Y			Kathiresan et al. - 2008, 2008, 2009, Sandhu et al. - 2008, Saxena et al. - 2007, Wellcome et al. - 2007, Willer et al. - 2008
rs4803770	19	50119193			Y	
rs4803771	19	50119488			Y	
rs7259004	19	50124397			Y	
rs12721108	19	50139081			Y	
rs5167	19	50140305	Y			Knoblauch et al. - 2004
rs5168	19	50140395			Y	
rs2288911	19	50141124	Y			Knoblauch et al. - 2004
rs5120	19	50143460	Y			Knoblauch et al. - 2004

rs5127	19	50144534			Y	
rs892101	19	50147298	Y			Knoblauch et al. - 2004
rs10413089	19	50147428			Y	
rs204907	19	50153836			Y	
rs1405655	19	55574431			Y	
rs3219281	19	55578899			Y	
rs3745099	19	57593717		Y		Kosoy et al. - 2009
rs2532060	19	60306735		Y		Kosoy et al. - 2009
rs6104567	20	10143433		Y		Kosoy et al. - 2009
rs6102059	20	38662198	Y			Kathiresan et al. - 2009
rs1800961	20	42475778	Y			Kathiresan et al. - 2009
rs742034	20	43955412			Y	
rs7270170	20	43960190			Y	
rs1736493	20	43964041			Y	
rs11086985	20	43965662			Y	
rs6073952	20	43970339			Y	
rs378114	20	43971834			Y	
rs441346	20	43972899			Y	
rs2294212	20	43973971			Y	
rs4810479	20	43978455			Y	
rs7679	20	44009909	Y			Kathiresan et al. - 2009
rs6094237	20	44062697			Y	
rs6065912	20	44065774			Y	
rs4810482	20	44067957			Y	
rs3787268	20	44075138			Y	
rs2250889	20	44075813			Y	
rs6094238	20	44086514			Y	
rs6095543	20	47546707			Y	
rs1393342	20	47548642			Y	
rs6095546	20	47550362			Y	
rs5619	20	47554250			Y	
rs5602	20	47555385			Y	
rs6090996	20	47567189			Y	
rs508757	20	47572378			Y	
rs5628	20	47574089			Y	
rs11697964	20	47579678			Y	
rs501908	20	47589174			Y	
rs12625166	20	47600610			Y	
rs476496	20	47611191			Y	
rs927068	20	47611381			Y	
rs6019902	20	47611620			Y	
rs477627	20	47613465			Y	
rs693649	20	47621486			Y	
rs11905684	20	47624098			Y	
rs11905590	20	51611297			Y	
rs2766669	20	51611633			Y	
rs6126871	20	51615163			Y	
rs6126872	20	51615982			Y	
rs2766672	20	51620191			Y	

rs6091708	20	51622424			Y	
rs6063966	20	51626495			Y	
rs3748503	20	51626775			Y	
rs16998248	20	51631747			Y	
rs1075390	20	51642245			Y	
rs3907047	20	53434321		Y		Kosoy et al. - 2009
rs741754	21	36143344	Y			
rs2835370	21	36807495		Y		Kosoy et al. - 2009
rs8128028	21	43339612			Y	
rs2401193	21	43339836			Y	
rs11911976	21	43340549			Y	
rs11909493	21	43340760			Y	
rs466791	21	43343023			Y	
rs719037	21	43344056			Y	
rs11700748	21	43346131			Y	
rs1051319	21	43346936			Y	
rs3788050	21	43347732			Y	
rs2124458	21	43348749			Y	
rs11203172	21	43353184			Y	
rs4920037	21	43354960			Y	
rs1789953	21	43356005			Y	
rs234705	21	43356841			Y	
rs11701048	21	43364494			Y	
rs1788484	21	43370010			Y	
rs2850146	21	43373655			Y	
rs4818909	21	44588388	Y			
rs1296819	22	16456546		Y		Kosoy et al. - 2009
rs4821004	22	30696359		Y		Kosoy et al. - 2009
rs4822056	22	40549800			Y	
rs7285782	22	40554964			Y	
rs2267439	22	40567715			Y	
rs7287886	22	40567973			Y	
rs714015	22	40589044			Y	
rs4822063	22	40606688			Y	
rs4822064	22	40609452			Y	
rs13055841	22	40609599			Y	
rs2269658	22	40610564			Y	
rs9623468	22	40611240			Y	
rs17002737	22	40611958			Y	
rs2267443	22	40617400			Y	
rs12157312	22	40620123			Y	
rs2269660	22	40622869			Y	
rs17379759	22	40629481			Y	
rs5031002	X	66859350	Y			Sabatti et al. - 2009

BIBLIOGRAPHY

1. Centers for Disease Control and Prevention (CDC) Pneumocystis pneumonia - Los Angeles. *MMWR, Morb Mortal Wkly Rep.* Jun5;30:250-2 (1981).
2. Centers for Disease Control and Prevention (CDC) Follow-Up on Kaposi's Sarcoma and Pneumocystis Pneumonia. *MMWR, Morb Mortal Wkly Rep.* Aug 28;30:409-10 (1981).
3. Centers for Disease Control and Prevention (CDC) Kaposi's Sarcoma and Pneumocystis Pneumonia Among Homosexual Men - New York City and California. *MMWR, Morb Mortal Wkly Rep.* July 4;30:305-8 (1981).
4. Goldsmith, J.C. et al. Immunologic profiles of adults with congenital bleeding disorders. *AIDS research* **1**, 163-79 (1983).
5. Barré-Sinoussi, F. et al. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science (New York, N.Y.)* **220**, 868-71 (1983).
6. Gallo, R.C. et al. Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS). *Science (New York, N.Y.)* **220**, 865-7 (1983).
7. *Global Report: UNAIDS report on the global AIDS pandemic 2010*. *2010* (2010).at <http://www.unaids.org/globalreport/global_report.htm>
8. Trevisan, M., Liu, J., Bahsas, F.B. & Menotti, A. Syndrome X and mortality: a population-based study. Risk Factor and Life Expectancy Research Group. *American journal of epidemiology* **148**, 958-66 (1998).
9. Hirsch, V.M., Olmsted, R.A., Murphey-Corb, M., Purcell, R.H. & Johnson, P.R. An African primate lentivirus (SIVsm) closely related to HIV-2. *Nature* **339**, 389-92 (1989).
10. Vanden Haesevelde, M.M. et al. Sequence analysis of a highly divergent HIV-1-related lentivirus isolated from a wild captured chimpanzee. *Virology* **221**, 346-50 (1996).
11. Stevenson, M. HIV-1 pathogenesis. *Nature medicine* **9**, 853-60 (2003).
12. Connor, R.I., Sheridan, K.E., Ceradini, D., Choe, S. & Landau, N.R. Change in coreceptor use correlates with disease progression in HIV-1--infected individuals. *The Journal of experimental medicine* **185**, 621-8 (1997).
13. Stone, M. et al. A limited number of simian immunodeficiency virus (SIV) env variants are transmitted to rhesus macaques vaginally inoculated with SIVmac251. *Journal of virology* **84**, 7083-95 (2010).
14. Liu, J. et al. Low-dose mucosal simian immunodeficiency virus infection restricts early replication kinetics and transmitted virus variants in rhesus monkeys. *Journal of virology* **84**, 10406-12 (2010).
15. Brenchley, J.M. et al. CD4+ T cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract. *The Journal of experimental medicine* **200**, 749-59 (2004).
16. Montagnier, L. 25 years after HIV discovery: prospects for cure and vaccine. *Virology* **397**, 248-54 (2010).
17. Perelson, A.S., Neumann, A.U., Markowitz, M., Leonard, J.M. & Ho, D.D. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science (New York, N.Y.)* **271**, 1582-6 (1996).
18. Mellors, J.W. et al. Prognosis in HIV-1 infection predicted by the quantity of virus in plasma. *Science (New York, N.Y.)* **272**, 1167-70 (1996).
19. Larder, B.A., Darby, G. & Richman, D.D. HIV with reduced sensitivity to zidovudine (AZT) isolated during prolonged therapy. *Science (New York, N.Y.)* **243**, 1731-4 (1989).
20. Basavapathruni, A. & Anderson, K.S. Reverse transcription of the HIV-1 pandemic. *The FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **21**, 3795-808 (2007).

21. Mellors, J.W. et al. In vitro selection and molecular characterization of human immunodeficiency virus-1 resistant to non-nucleoside inhibitors of reverse transcriptase. *Molecular pharmacology* **41**, 446-51 (1992).
22. Davey, R.T. et al. HIV-1 and T cell dynamics after interruption of highly active antiretroviral therapy (HAART) in patients with a history of sustained viral suppression. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 15109-14 (1999).
23. Liao, Z., Cimakasky, L.M., Hampton, R., Nguyen, D.H. & Hildreth, J.E. Lipid rafts and HIV pathogenesis: host membrane cholesterol is required for infection by HIV type 1. *AIDS research and human retroviruses* **17**, 1009-19 (2001).
24. Luo, C., Wang, K., Liu, D.Q., Li, Y. & Zhao, Q.S. The functional roles of lipid rafts in T cell activation, immune diseases and HIV infection and prevention. *Cellular & molecular immunology* **5**, 1-7 (2008).
25. Pietiäinen, V.M., Marjomäki, V., Heino, J. & Hyypiä, T. Viral entry, lipid rafts and caveosomes. *Annals of medicine* **37**, 394-403 (2005).
26. del Real, G. et al. Statins inhibit HIV-1 infection by down-regulating Rho activity. *The Journal of experimental medicine* **200**, 541-7 (2004).
27. Moog, C., Aubertin, A.M., Kirn, A. & Luu, B. Oxysterols, but not cholesterol, inhibit human immunodeficiency virus replication in vitro. *Antiviral chemistry & chemotherapy* **9**, 491-6 (1998).
28. Zheng, Y.-H., Plemenitas, A., Fielding, C.J. & Peterlin, B.M. Nef increases the synthesis of and transports cholesterol to lipid rafts and HIV-1 progeny virions. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 8460-5 (2003).
29. Tayo, B.O. et al. Genetic Background of Patients from a University Medical Center in Manhattan: Implications for Personalized Medicine. *PLoS ONE* **6**, e19166 (2011).
30. Mujawar, Z. et al. Human immunodeficiency virus impairs reverse cholesterol transport from macrophages. *PLoS biology* **4**, e365 (2006).
31. van 't Wout, A.B. et al. Nef induces multiple genes involved in cholesterol synthesis and uptake in human immunodeficiency virus type 1-infected T cells. *Journal of virology* **79**, 10053-8 (2005).
32. Grunfeld, C. et al. Lipids, lipoproteins, triglyceride clearance, and cytokines in human immunodeficiency virus infection and the acquired immunodeficiency syndrome. *The Journal of clinical endocrinology and metabolism* **74**, 1045-52 (1992).
33. Safrin, S. & Grunfeld, C. Fat distribution and metabolic changes in patients with HIV infection. *AIDS (London, England)* **13**, 2493-505 (1999).
34. El-Sadr, W.M. et al. Effects of HIV disease on lipid, glucose and insulin levels: results from a large antiretroviral-naive cohort. *HIV medicine* **6**, 114-21 (2005).
35. Hellerstein, M.K. et al. Increased de novo hepatic lipogenesis in human immunodeficiency virus infection. *The Journal of clinical endocrinology and metabolism* **76**, 559-65 (1993).
36. Roger, V.L. et al. Heart disease and stroke statistics--2011 update: a report from the American Heart Association. *Circulation* **123**, e18-e209 (2011).
37. Wilson, P.W. et al. Prediction of coronary heart disease using risk factor categories. *Circulation* **97**, 1837-47 (1998).
38. Watts, G.F. & Karpe, F. Triglycerides and atherogenic dyslipidaemia: extending treatment beyond statins in the high-risk cardiovascular patient. *Heart (British Cardiac Society)* **97**, 350-6 (2011).
39. Enas, E.A. Hypertriglyceridemia and elevated lipoprotein(a) are major risk factors for coronary events in middle-aged men. *The American journal of cardiology* **78**, 859-60 (1996).
40. van Leuven, S.I. et al. Systemic inflammation as a risk factor for atherothrombosis. *Rheumatology (Oxford, England)* **47**, 3-7 (2008).
41. Metcalf, P.A. et al. African American-white differences in lipids, lipoproteins, and apolipoproteins, by educational attainment, among middle-aged adults: the Atherosclerosis Risk in Communities Study. *American journal of epidemiology* **148**, 750-60 (1998).
42. Kuller, L.H. Ethnic differences in atherosclerosis, cardiovascular disease and lipid metabolism. *Current opinion in lipidology* **15**, 109-13 (2004).
43. Johnson, J.L. et al. Gender and racial differences in lipoprotein subclass distributions: the STRRIDE study. *Atherosclerosis* **176**, 371-7 (2004).
44. Rader, D.J. & Daugherty, A. Translating molecular discoveries into new therapies for atherosclerosis. *Nature* **451**, 904-13 (2008).
45. Santos, R.D. & Nasir, K. Insights into atherosclerosis from invasive and non-invasive imaging studies: Should we treat subclinical atherosclerosis? *Atherosclerosis* **205**, 349-56 (2009).

46. Bonnet, F. et al. Increase of atherogenic plasma profile in HIV-infected patients treated with protease inhibitor-containing regimens. *Journal of acquired immune deficiency syndromes (1999)* **25**, 199-200 (2000).
47. Carr, A. et al. A syndrome of peripheral lipodystrophy, hyperlipidaemia and insulin resistance in patients receiving HIV protease inhibitors. *AIDS (London, England)* **12**, F51-8 (1998).
48. Périard, D. et al. Atherogenic dyslipidemia in HIV-infected individuals treated with protease inhibitors. The Swiss HIV Cohort Study. *Circulation* **100**, 700-5 (1999).
49. Riddler, S. a et al. Impact of HIV infection and HAART on serum lipids in men. *JAMA : the journal of the American Medical Association* **289**, 2978-82 (2003).
50. Ranganathan, S. & Kern, P.A. The HIV protease inhibitor saquinavir impairs lipid metabolism and glucose transport in cultured adipocytes. *The Journal of endocrinology* **172**, 155-62 (2002).
51. Brown, T.T. et al. Fat distribution and longitudinal anthropometric changes in HIV-infected men with and without clinical evidence of lipodystrophy and HIV-uninfected controls: a substudy of the Multicenter AIDS Cohort Study. *AIDS research and therapy* **6**, 8 (2009).
52. Galli, M., Ridolfo, A.L. & Gervasoni, C. Cardiovascular disease risk factors in HIV-infected patients in the HAART era. *Annals of the New York Academy of Sciences* **946**, 200-13 (2001).
53. Hadigan, C. et al. Metabolic abnormalities and cardiovascular disease risk factors in adults with human immunodeficiency virus infection and lipodystrophy. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **32**, 130-9 (2001).
54. Rickerts, V., Brodt, H., Staszewski, S. & Stille, W. Incidence of myocardial infarctions in HIV-infected patients between 1983 and 1998: the Frankfurt HIV-cohort study. *European journal of medical research* **5**, 329-33 (2000).
55. Jütte, A. et al. Increasing morbidity from myocardial infarction during HIV protease inhibitor treatment? *AIDS (London, England)* **13**, 1796-7 (1999).
56. Bozzette, S. a, Ake, C.F., Tam, H.K., Chang, S.W. & Louis, T.A. Cardiovascular and cerebrovascular events in patients treated for human immunodeficiency virus infection. *The New England journal of medicine* **348**, 702-10 (2003).
57. Henry, K. et al. Severe premature coronary artery disease with protease inhibitors Hormone-receptor status of breast cancer in Papua New Guinea. *The Lancet* **351**, 1998-1998 (1998).
58. Currier, J.S. et al. Coronary heart disease in HIV-infected individuals. *Journal of acquired immune deficiency syndromes (1999)* **33**, 506-12 (2003).
59. Riddler, S. a et al. Antiretroviral therapy is associated with an atherogenic lipoprotein phenotype among HIV-1-infected men in the Multicenter AIDS Cohort Study. *Journal of acquired immune deficiency syndromes (1999)* **48**, 281-8 (2008).
60. Behrens, G. et al. Impaired glucose tolerance, beta cell function and lipid metabolism in HIV patients under treatment with protease inhibitors. *AIDS (London, England)* **13**, F63-70 (1999).
61. Grunfeld, C. Dyslipidemia and its Treatment in HIV Infection. *Topics in HIV medicine : a publication of the International AIDS Society, USA* **18**, 112-8 (2011).
62. Lee, G.A. et al. The metabolic effects of lopinavir/ritonavir in HIV-negative men. *AIDS (London, England)* **18**, 641-9 (2004).
63. Purnell, J.Q. et al. Effect of ritonavir on lipids and post-heparin lipase activities in normal subjects. *AIDS (London, England)* **14**, 51-7 (2000).
64. Périard, D. et al. Atherogenic dyslipidemia in HIV-infected individuals treated with protease inhibitors. The Swiss HIV Cohort Study. *Circulation* **100**, 700-5 (1999).
65. Noor, M.A. et al. Metabolic effects of indinavir in healthy HIV-seronegative men. *AIDS (London, England)* **15**, F11-8 (2001).
66. Mallal, S. et al. Association between presence of HLA-B*5701, HLA-DR7, and HLA-DQ3 and hypersensitivity to HIV-1 reverse-transcriptase inhibitor abacavir. *Lancet* **359**, 727-32 (2002).
67. Klein, D., Hurley, L.B., Quesenberry, C.P. & Sidney, S. Do protease inhibitors increase the risk for coronary heart disease in patients with HIV-1 infection? *Journal of acquired immune deficiency syndromes (1999)* **30**, 471-7 (2002).
68. Triant, V.A., Lee, H., Hadigan, C. & Grinspoon, S.K. Increased acute myocardial infarction rates and cardiovascular risk factors among patients with human immunodeficiency virus disease. *The Journal of clinical endocrinology and metabolism* **92**, 2506-12 (2007).
69. Law, M.G. et al. The use of the Framingham equation to predict myocardial infarctions in HIV-infected patients: comparison with observed events in the D:A:D Study. *HIV medicine* **7**, 218-30 (2006).

70. Friis-Møller, N. et al. Class of antiretroviral drugs and the risk of myocardial infarction. *The New England journal of medicine* **356**, 1723-35 (2007).
71. Kingsley, L. a et al. Subclinical coronary atherosclerosis, HIV infection and antiretroviral therapy: Multicenter AIDS Cohort Study. *AIDS (London, England)* **22**, 1589-99 (2008).
72. Seaberg, E.C. et al. Association Between Human Immunodeficiency Virus Infection and Stiffness of the Common Carotid Artery. *Stroke; a journal of cerebral circulation* (2010).doi:10.1161/STROKEAHA.110.583856
73. Mangili, A. et al. Risk of cardiovascular disease in a cohort of HIV-infected adults: a study using carotid intima-media thickness and coronary artery calcium score. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **43**, 1482-9 (2006).
74. Tishkoff, S.A. & Kidd, K.K. Implications of biogeography of human populations for "race" and medicine. *Nature genetics* **36**, S21-7 (2004).
75. Pritchard, J.K. & Rosenberg, N. a Use of unlinked genetic markers to detect population stratification in association studies. *American journal of human genetics* **65**, 220-8 (1999).
76. Divers, J. et al. Comparing self-reported ethnicity to genetic background measures in the context of the Multi-Ethnic Study of Atherosclerosis (MESA). *BMC genetics* **12**, 28 (2011).
77. Arnedo, M. et al. Contribution of 20 single nucleotide polymorphisms of 13 genes to dyslipidemia associated with antiretroviral therapy. *Pharmacogenetics and genomics* **17**, 755-64 (2007).
78. Aulchenko, Y.S. et al. Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. *Nature genetics* **41**, 47-55 (2009).
79. Bauerfeind, A. et al. Concordant association of lipid gene variation with a combined HDL/LDL-cholesterol phenotype in two European populations. *Human heredity* **61**, 123-31 (2006).
80. Bersano, a, Ballabio, E., Bresolin, N. & Candelise, L. Genetic polymorphisms for the study of multifactorial stroke. *Human mutation* **29**, 776-95 (2008).
81. Boekholdt, S.M. et al. Cholesteryl ester transfer protein TaqIB variant, high-density lipoprotein cholesterol levels, cardiovascular risk, and efficacy of pravastatin treatment: individual patient meta-analysis of 13,677 subjects. *Circulation* **111**, 278-87 (2005).
82. Engler, M.B. et al. Genetic variation in phospholipid transfer protein modulates lipoprotein profiles in hyperalphalipoproteinemia. *Metabolism: clinical and experimental* **57**, 1719-24 (2008).
83. Helgadottir, A. et al. The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm. *Nature genetics* **40**, 217-24 (2008).
84. Helgadottir, A. et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science (New York, N.Y.)* **316**, 1491-3 (2007).
85. Kathiresan, S. et al. Polymorphisms associated with cholesterol and risk of cardiovascular events. *The New England journal of medicine* **358**, 1240-9 (2008).
86. Kathiresan, S. et al. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nature genetics* **40**, 189-97 (2008).
87. Kathiresan, S. et al. Common variants at 30 loci contribute to polygenic dyslipidemia. *Nature genetics* **41**, 56-65 (2009).
88. Keating, B.J. et al. Concept, design and implementation of a cardiovascular gene-centric 50 k SNP array for large-scale genomic association studies. *PloS one* **3**, e3583 (2008).
89. Knoblauch, H. et al. Common haplotypes in five genes influence genetic variance of LDL and HDL cholesterol in the general population. *Human molecular genetics* **11**, 1477-85 (2002).
90. Knoblauch, H. et al. Haplotypes and SNPs in 13 lipid-relevant genes explain most of the genetic variance in high-density lipoprotein and low-density lipoprotein cholesterol. *Human molecular genetics* **13**, 993-1004 (2004).
91. Morabia, A. et al. Association of extreme blood lipid profile phenotypic variation with 11 reverse cholesterol transport genes and 10 non-genetic cardiovascular disease risk factors. *Human molecular genetics* **12**, 2733-43 (2003).
92. Morabia, A. et al. APPENDIX : Tables A – E and Figures A , B Association of Extreme Blood Lipid Profile Phenotypic Variation with 11 Reverse Cholesterol Transport Genes and 10 Nongenetic Cardiovascular Disease Risk Factors Website : *New York*
93. Sabatti, C. et al. Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. *Nature genetics* **41**, 35-46 (2009).
94. Wang, S.S. et al. Identification of pathways for atherosclerosis in mice: integration of quantitative trait locus analysis and global gene expression data. *Circulation research* **101**, e11-30 (2007).

95. Willer, C.J. et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nature genetics* **40**, 161-9 (2008).
96. Wootton, P.T.E. et al. Tagging SNP haplotype analysis of the secretory PLA2-V gene, PLA2G5, shows strong association with LDL and oxLDL levels, suggesting functional distinction from sPLA2-IIA: results from the UDACS study. *Human molecular genetics* **16**, 1437-44 (2007).
97. Manolio, T.A. Cohort studies and the genetics of complex disease. *Nature genetics* **41**, 5-6 (2009).
98. Boekholdt, S.M. & Thompson, J.F. Natural genetic variation as a tool in understanding the role of CETP in lipid levels and disease. *Journal Of Lipid Research* **44**, (2003).
99. Sabatti, C. et al. Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. *Nature Genetics* **41**, 35-46 (2009).
100. Tsai, C.-T. et al. Interaction of gender, hypertension, and the angiotensinogen gene haplotypes on the risk of coronary artery disease in a large angiographic cohort. *Atherosclerosis* **203**, 249-56 (2009).
101. Sandhu, M.S. et al. LDL-cholesterol concentrations: a genome-wide association study. *Lancet* **371**, 483-91 (2008).
102. McPherson, R. et al. A common allele on chromosome 9 associated with coronary heart disease. *Science (New York, N.Y.)* **316**, 1488-91 (2007).
103. O'Donnell, C.J. et al. Genome-wide association study for subclinical atherosclerosis in major arterial territories in the NHLBI's Framingham Heart Study. *BMC medical genetics* **8 Suppl 1**, S4 (2007).
104. Saxena, R. et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science (New York, N.Y.)* **316**, 1331-6 (2007).
105. Engler, M.B. et al. Genetic variation in phospholipid transfer protein modulates lipoprotein profiles in hyperalphalipoproteinemia. *Metabolism: clinical and experimental* **57**, 1719-24 (2008).
106. Love-Gregory, L. et al. Variants in the CD36 gene associate with the metabolic syndrome and high-density lipoprotein cholesterol. *Human molecular genetics* **17**, 1695-704 (2008).
107. Levy, D. et al. Framingham Heart Study 100K Project: genome-wide associations for blood pressure and arterial stiffness. *BMC medical genetics* **8 Suppl 1**, S3 (2007).
108. Samani, N.J. et al. Genomewide association analysis of coronary artery disease. *The New England journal of medicine* **357**, 443-53 (2007).
109. Wellcome, T., Case, T. & Consortium, C. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661-78 (2007).
110. Guidelines for the use of antiretroviral agents in HIV-infected adults and adolescents. Department of Health and Human Services and Henry J. Kaiser Family Foundation. *MMWR. Recommendations and reports : Morbidity and mortality weekly report. Recommendations and reports / Centers for Disease Control* **47**, 43-82 (1998).
111. Singer, V.L., Jones, L.J., Yue, S.T. & Haugland, R.P. Characterization of PicoGreen reagent and development of a fluorescence-based solution assay for double-stranded DNA quantitation. *Analytical biochemistry* **249**, 228-38 (1997).
112. Chen, X., Levine, L. & Kwok, P.Y. Fluorescence polarization in homogeneous nucleic acid analysis. *Genome research* **9**, 492-8 (1999).
113. Kwok, P.-Y. SNP genotyping with fluorescence polarization detection. *Human mutation* **19**, 315-23 (2002).
114. Barrett, J.C., Fry, B., Maller, J. & Daly, M.J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics (Oxford, England)* **21**, 263-5 (2005).
115. de Bakker, P.I.W. et al. Efficiency and power in genetic association studies. *Nature genetics* **37**, 1217-23 (2005).
116. Weeks, J.P. Plink: An R Package for Linking Mixed-Format Tests Using IRT-Based Methods. (2010).at <http://www.jstatsoft.org/v35/i12/>
117. Team, R.D.C. R: A language and environment for statistical computing. (2008).at <http://www.r-project.org>
118. Kuo, C.-L. & Feingold, E. What's the best statistic for a simple test of genetic association in a case-control study? *Genetic epidemiology* **34**, 246-53 (2010).
119. Clayton, D.G. et al. Population structure, differential bias and genomic control in a large-scale, case-control association study. *Nature genetics* **37**, 1243-6 (2005).
120. Kosoy, R. et al. Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America. *Human mutation* **30**, 69-78 (2009).
121. Rasmussen-Torvik, L.J. et al. Impact of repeated measures and sample selection on genome-wide association studies of fasting glucose. *Genetic epidemiology* **34**, 665-73 (2010).

122. Pendergrass, S., Dudek, S.M., Roden, D.M., Crawford, D.C. & Ritchie, M.D. Visual integration of results from a large DNA biobank (biovu) using synthesis-view. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing* 265-75 (2011).at <http://www.ncbi.nlm.nih.gov/pubmed/21121054>
123. Jansen, H. Hepatic lipase: a pro- or anti-atherogenic protein? *The Journal of Lipid Research* **43**, 1352-1362 (2002).
124. Perret, B. et al. Hepatic lipase: structure/function relationship, synthesis, and regulation. *Journal of lipid research* **43**, 1163-9 (2002).
125. Zambon, A., Deeb, S.S., Pauletto, P., Crepaldi, G. & Brunzell, J.D. Hepatic lipase: a marker for cardiovascular disease risk and response to therapy. *Current opinion in lipidology* **14**, 179-89 (2003).
126. Thompson, J.F., Wood, L.S., Pickering, E.H., Dechairo, B. & Hyde, C.L. High-density genotyping and functional SNP localization in the CETP gene. *Journal of lipid research* **48**, 434-43 (2007).
127. Weber, O., Bischoff, H., Schmeck, C. & Böttcher, M.-F. Cholesteryl ester transfer protein and its inhibition. *Cellular and molecular life sciences : CMLS* 3139-3149 (2010).doi:10.1007/s00018-010-0418-3
128. de Grooth, G.J. et al. A review of CETP and its relation to atherosclerosis. *Journal of lipid research* **45**, 1967-74 (2004).
129. Tai, E.S. & Ordovas, J.M. Clinical significance of apolipoprotein A5. *Current opinion in lipidology* **19**, 349-54 (2008).
130. Hubáýek, J.A. et al. Apolipoprotein A5 in Health and Disease Description and function of apolipoprotein. *Clinical Chemistry* **58**, (2009).