Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Michael Wong Date

Atherosclerosis, Carotid Intima-Media Thickness, and DNA Methylation

By

Michael Wong

MPH

Epidemiology

_____

[Faculty Thesis Advisor's signature]

Yan Sun

Faculty Thesis Advisor

Atherosclerosis, Carotid Intima-Media Thickness, and DNA Methylation

by

Michael Wong

Bachelor of Arts

Johns Hopkins University

2013

Faculty Thesis Advisor: **Yan Sun**, PhD

An abstract of

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Public Health

in Epidemiology

2016

Abstract

Atherosclerosis, Carotid Intima-Media Thickness, and DNA Methylation

by Michael Wong

The rise of chronic disease has made the study of cardiovascular disease and atherosclerosis increasingly important. Cardiovascular disease is incredibly prevalent in the United States (over 1 in 3 adults) and is the single largest contributor to mortality nationally and globally. While traditional cardiovascular disease risk factors, such as smoking and high blood pressure, have been identified and studied for decades, advances in technology have helped identify risk factors such as carotid intima-media thickness (CIMT), measurable through ultrasonography, and DNA methylation, quantifiable through next generation sequencing (NGS). Using CIMT as a measure of the development of atherosclerosis, 13 CpG sites were identified to have an association between DNA methylation and the IMT variables.

Atherosclerosis, Carotid Intima-Media Thickness, and DNA Methylation

by

Michael Wong

Bachelor of Arts

Johns Hopkins University

2013

Faculty Thesis Advisor: **Yan Sun**, PhD

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Public Health

in Epidemiology

2016

Table of Contents:

Background:

Epidemiology:

Cardiovascular disease (CVD) is the leading cause of disability and premature mortality worldwide. (1,2) In the United States today, approximately 85 million American adults over the age of 20 (over one in three) are living with CVD, with 43.7 million of those over the age of 60. (3) CVD is a chronic disease that can lead to a fatal outcome such as heart attack and stroke. While the rate of mortality from this disease has decreased by 30% from 2001 to 2011 due to improving medical capabilities, it remains the leading cause of illness and death in the United States today, accounting for 31% of all mortality. Over 2000 Americans die of cardiovascular disease or related complications everyday, or about 230 per 100,000. (3,4,5) By 2030, the American Heart Association projects that 44% of the United States population will have CVD. (3) One study suggests that eliminating all CVD would increase average life expectancy of Americans by seven years. (6)

On a global scale, the World Health Organization (WHO) estimates 20 million deaths attributable to cardiovascular disease in 2015, which would account for 30% of the world's total deaths. (7) As the world continues to shift away from infectious disease towards chronic disease, the absolute number and proportions of CVD death are projected to continue to increase, and CVD will continue to be the United States' and the world's greatest contributor to mortality.

Pathogenesis:

Atherosclerosis is often a precursor for CVD. It is a disease that starts in childhood and can develop over decades. (8) It is a condition in which the walls of the arteries thicken and harden due to inflammatory response from white blood cells. The fatty streak is a visible lesion found in the innermost layer of the artery, often found in children. (9) It consists of lipids, macrophages, smooth muscle cells, and cellular waste products. A plaque builds as white blood cells respond to the lesion, which becomes coated with a tough, fibrous cap made up of connective tissue and smooth muscle, making it difficult to treat. As atherosclerosis advances, the blood flow in the artery can be partially blocked, and a cardiovascular disease can manifest in the form of a heart attack or stroke. (10) Atherosclerosis can also lead to carotid artery disease, where the walls of the arteries that carry oxygen-rich blood to the brain narrow, or coronary artery disease, where the arteries going to the heart are blocked. (11)

Risk Factors:

While the exact etiology of atherosclerosis is not perfectly understood, certain risk factors for atherosclerosis and carotid artery disease can accelerate the complications of the disease. Major risk factors for atherosclerosis include smoking, high blood pressure, high blood cholesterol, diabetes, age, physical activity, obesity, and genetics. Due to the relationships between these risk factors

and CVD, studying these traditional, yet important contributors is critical in order to understand and treat atherosclerosis. One study estimates that not smoking, maintaining healthy weight through diet and regular exercise, not having diabetes, and maintaining reasonable blood pressure and cholesterol levels can prevent 80% of cardiovascular disease. (3) Meanwhile, data from NHANES II shows that mortality from cardiovascular disease was reduced by about 60% for non-smokers without hypertension and high cholesterol. (12)

High carotid intima-media thickness (IMT) is commonly used as a quantitative measurement to assess atherosclerosis and overall cardiovascular health. (13, 14, 15, 16) IMT has been shown to be a low-cost, non-invasive, and reproducible approach to predicting cardiovascular outcomes independently from the more traditional risk factors, like high blood pressure. (17) Progression of atherosclerosis is often measured through change in IMT over time, which is popularly used as the outcome of interest in many risk factor intervention and observational studies. (16) IMT represents the thickness of the two innermost layers of an artery wall. It is measured through non-invasive B-mode ultrasound. (18) IMT can be measured from either the carotid artery wall nearer to or farther from the ultrasound transducer, known as near wall and far wall measurements respectively. (19) Research suggests that far wall measurements are more accurate than near wall measurements in predicting CVD outcomes, while far wall measurements combined with near wall measurements have also produced

reliable results. (20,21) Measurement values larger than 1mm suggest significant plaque buildup in the artery, and even values greater than 0.80mm for common carotid artery IMT have shown a two-fold increase in CVD outcomes. (13,22)

Epigenetics:

Epigenetic processes are important regulators of gene expression. The most widely studied mechanism in the field of epigenetic epidemiology is DNA methylation (DNAm), a form of epigenetic modification of cytosine that can lead to modification of gene expression in regulatory regions. (23) DNAm has been linked to many chronic diseases including atherosclerosis, as well as aging. (1) DNA methylation levels at a specific locus naturally change over time, and as monozygotic twins grow older, their epigenetic profiles show greater variability, likely through slight changes over years of successive cell divisions. (24) This is known as “epigenetic drift,” which is associated with aging. (25)

The most common type of methylation involves CpG sites, where a methyl group is reversibly added to the 5’ carbon of cytosine to potentially alter expression. (26) In humans, healthy somatic cells exhibit methylation on about 1% of cytosines, although variation from that number can occur in different kinds of tissues. (27) Approximately 70-80% of cytosines in CpG sites are methylated. (28) These stretches of methylation are important in silencing in noncoding DNA such as introns. (1) There are also regions with high unmethylated CpG counts,

known as CpG islands, which are found in gene promoters. (24) Alterations by methylation found in CpG islands are of particular interest in atherosclerosis research for their role in gene expression. It has been proposed that these epigenetic changes can modify DNA methylation patterns in arterial smooth muscle cell walls. Hypermethylation of CpG sites usually results in transcriptional silencing, with some studies suggesting increased DNAm in CVD patients. (29) Other studies propose a relationship between inflammation and hypermethylation in blood lymphocytes. (30) Research also suggests that hypomethylation of CpG sites may occur in atherosclerotic lesions and plaques. (8)

The study of epigenetic factors in the field of atherosclerosis research is rather new, but is becoming increasingly important. A person's DNA methylation profile presents a history of that person's environmental exposures that can be useful in studying disease risk. Thus, examining the relationship between DNA methylation, atherosclerosis, and cardiovascular disease risk factors could be crucial in understanding the etiology of atherosclerosis.

Twin Studies of Epigenetics:

Twin studies have been well documented as being valuable in genetic epidemiology studies. Being able to compare monozygotic and dizygotic twins can help estimate the heritability of traits (i.e., the variance in traits due to genetic

factors). (31) Twins are reliable for studying many traits because of the ability to ignore age effects and separate genetic and environmental factors. (24)

Monozygotic twin studies are especially helpful in epigenetic research and understanding epigenetic profiles. Differences in phenotype between co-twins imply that environmentally related epigenetic factors, such as diet, smoking, or stress may be involved. (32) Monozygotic twins are matched for age, genotype, maternal environment, and sex. In recent research, for some diseases such as autism spectrum disorder, psychosis, and Type 1 diabetes, there is often discordance between monozygotic twins, suggesting that differences in environment may play a role. (24,31) Research supports the idea that the effect of the environment on the human genome can be mediated through epigenetics. Using twin studies in epigenetics is advantageous because it helps to compare the epigenetic variance at a single locus due to genetic variation and environmental factors between twins. (33)

Methods:

Subjects:

The DNA methylation data comes from the Emory Twin Study, which examined behavioral, biological, and psychological risk factors for cardiovascular disease among twins. In these studies, male monozygotic and dizygotic twins born between 1946 and 1956 were randomly selected from the Vietnam Era Twin

(VET) Registry, which is composed of 7369 twin pairs and is one of the largest twin registries in the United States. (34) The twins were examined together at the Emory University General Clinical Research Center between March 2002 and March 2006. (35) The study protocol was approved by the Emory University Institutional Review Board, and all subjects submitted informed consent. (36)

Data Methods:

142 monozygotic twins were epityped using Illumina Infinium 450K methylation BeadChip array to examine over 480,000 CpG sites. This array uses 500 nanograms of genomic DNA per sample from peripheral blood leukocytes and can fit 12 samples per array. It examines methylation levels of 485,577 CpG sites, covering 99% of RefSeq genes and 96% of CpG islands. Methylation level is measured with a probe that uses red and green fluorescent labeling to determine levels of cytosine or thymine resulting from bisulfate transformation of unmethylated cytosine. Methylation level is based on the signal intensities of the fluorescent labels. Computed beta values, ranging from 0 to 1, are ratios of fluorescence intensity of methylated probes to the sum of intensities of both methylated and unmethylated probes. Quantile-normalized beta values were performed to adjust for potential technical shift between methylation signals across differing probe categories. (37) From these beta values, no control probe values were greater than 4 standard deviations from their mean values. Methylation sites were excluded if they overlapped with known single nucleotide

polymorphisms (SNPs) or if they could not be uniquely matched to a genomic location on the reference genome. (38) After these quality control measures, the final analysis included 409,967 autosomal CpG sites and 140 MZ twins.

Twins were given identical diets, fasted overnight, and were not allowed to smoke the night before testing. (36) Medical history was taken at time of examination, and all measurements were performed in the morning. Body mass index (BMI) was measured by dividing weight (kg) by the square of height (meters), with obesity being a BMI greater than 30. Low- (LDL) and high-density lipoprotein (HDL) cholesterol levels were measured using plasma samples. Blood pressure was taken in two measurements using a mercury sphygmomanometer on the right arm. Smoking status was obtained through questionnaire and was divided into three groups: never smokers, current smokers, and past smokers, who reported to have smoked more than 100 cigarettes. Physical activity was measured with the Baecke questionnaire, which uses 16 questions to determine the level of activity. Intima-media thickness (IMT) was measured using B-mode ultrasound. The ultrasound transducer is moved around the circumference of the neck to scan multiple carotid artery segments. (39)

Research suggests that DNA methylation profile is cell type specific. The shift of cell type proportions in peripheral blood can be a major confounder for blood cell-based epigenetic association studies. (40) The white blood cell types - B

lymphocytes, granulocytes, monocytes, natural killer (NK) cells, and T cells (95% of which are CD^4+ T cells and CD^8+ T cells) - should be controlled for, because the modified proportions of leukocyte subtypes can shift the cell type specific DNA methylation profiles, leading to differences in methylation measures between cases and controls. (41) Houseman et al have developed a regression model that estimates cell compositional differences using sorted leukocyte references.

All statistical analyses were performed using R version 3.2.2 (https://www.r-project.org/). The R package *nlme* was used to implement a linear mixed effect model for batch and twin effects.

Two different linear mixed effects models were run to test the relationship between intima-media thickness and DNA methylation. DNA methylation was measured using the beta values calculated using fluorescence intensity.

1. DNA methylation ~ (β_1)IMT_max_FW + (β_2)current smoking status + (β_3)age + (β_4)BMI + (β_5)HDL cholesterol + (β_6)LDL cholesterol + (β_7)physical activity + (β_8)cell type proportions + ε
2. DNA methylation ~ (β_1)IMT_max_tot + (β_2)current smoking status + (β_3)age + (β_4)BMI + (β_5)HDL cholesterol + (β_6)LDL cholesterol + (β_7)physical activity + (β_8)cell type proportions + ε

3. Results:

For the 140 monozygotic twins, each twin in a pair was randomly assigned to one of two groups to compare the distribution of variables between co-twins. **Table 1** shows the descriptive statistics of the study group and related risk factors.

To determine the effect of IMT on DNA methylation, we used a linear mixed effect (LME) model. The variable IMT_max_FW represents the intima-media thickness measured from the far wall, while the variable IMT_max_tot represents the intima-media thickness from the near wall. Twin effects were considered random effects. Never smokers and past smokers were combined as non-current smokers. Both HDL and LDL cholesterol levels were included.

The results of the top epigenetics associations (p-value $< 2 \times 10^{-5}$) from two linear mixed effect models are shown in **Table 2**. DNA methylation and IMT measured from the far wall were most strongly associated at CpG sites cg04085789 and cg06945936. DNA methylation and IMT measured from the near wall were most strongly associated at CpG sites cg04085789, cg00157962, cg00187229, cg02005758, cg03534505, cg05257202, cg09827751, cg10799327, cg11551560, cg26217402, and cg26532621. Only cg04085789 was strongly associated with both IMT traits. The analyses produced 13 CpG sites that suggest associations between methylation and intima-media thickness. While the two measurements of IMT are similar, the far wall results showed associations

with low p-values. All effect sizes were positive values, suggesting an increase in thickness of the arterial walls is associated with DNA hypermethylation.

Figure 1 shows the summary plots (a. Q-Q plot, b. Manhattan plot, c: regional plot) for intima-media thickness measured from far wall, while **Figure 2** shows the summary plots for intima-media thickness measured from near wall. The Q-Q plot **(a)** shows the association between DNA methylation and the IMT variables. The inflation factor (λ) value represents the ratio between the medians of the chi square distributions of the expected and observed values. (42) A ratio of 1 indicates the data follows a normal chi-square distribution, while a ratio greater than 1 indicates inflation of the null distribution. Both the far wall plot and near wall plot show moderately deflated values (λ=0.929 and λ=0.907, respectively). The higher IF value in the far wall plot indicates that the observed pattern was closer to the expected null distribution compared to that of the near wall plot. The Manhattan plot **(b)** maps out the p-values of all tested CpG sites in the genome. More significant p-values are smaller, and as a result, would have larger –log(p-values). Using 1.5×10^{-6} as the baseline for significance, only one site was found to be truly significant, methylation of cg06945936 on chromosome 19 and the far wall measurement. The highest purple dot on chromosome 22 (cg04085789) appears in both plots, showing a strong association in both measurements. After the strongest associations are identified, the CpG regional plots are used to depict the location of that CpG site, its position on the chromosome, and its

proximity to any genes. A regional plot **(c)** is created to look at the strength of an association in the context of surrounding genes and CpG sites. These plots may display multiple closely related sites that show varying degrees of association. (43) Figure 1c shows the regional plot for cg06945936 (KIRREL2), cg04085789 (SLC25A1), and cg10740660 (VARS). Figure 2c shows the three strongest (lowest p-values), which were identified as cg05257202 (UBAC2), cg11551560 (TLE3), and cg04085789 (SLC25A1).

Discussion:

CpG Locations

Cg10751811 (p-value: 1.68 x 10^{-5}) is located on chromosome 1 at the Cytochrome P450 4X1 (CYP4X1) gene, which is expressed in trachea and aorta. (44) This gene encodes the CYP4X1 protein, a cytochrome P450 protein, which is involved in drug breakdown and lipogenesis (lipid formation). Also, increased expression of CYP4X1 has been linked to intake of flavonoids, which has been shown to reduce inflammation to reduce CVD. (45,46) Another site, cg09827751, is found at the Epsin-3 (EPN3) gene, which has been shown to increase expression with increased levels of ethinyl estradiol, which in turn has been shown to protect artery walls against atherosclerotic lesions from high cholesterol diets. (47,48) Lastly, the site cg26532621 is found at the PROP paired-like homeobox (PROP1) gene. Hydrogen peroxide has been show to create oxidative stress leading to artery wall cell death and inflammation. (49) PROP1 gene

mutants have shown a resistance to hydrogen peroxide, and thus could resist hydrogen peroxide exposures. (50)

Causality in Epigenetic Studies:

Statistical significance does not always imply biological significance, so it is useful to look at the strength of the association as well as the p-values. (41) In our study, the top associations were not found to be statistically significant after correction for multiple testing, but may have value biologically. While many epigenetic studies have found significant associations between DNA methylation and gene expression differences, the correlations can be somewhat small and correlation does not always imply a causal relationship. Epigenetic patterns can be fluid and subject to modification, so while epigenetic variation and a phenotype may be associated, it may be a challenge to establishing causality. (51) For example, recent research has identified epigenetic biomarkers in bladder and ovarian cancers. However, a causal relationship between these biomarkers and phenotype can be very complex. Studies must address confounding, reverse causation, and measurement error. While genetic variables are not subject to classical epidemiologic confounders, there may be confounding through linkage disequilibrium, which can be addressed through population stratification. Reverse causation can arise in epigenetics, as a trait or disease can influence health behaviors, thus changing the epigenome. Hypothetically, cancer can increase expression of the gene SREBF2, which changes the epigenome by modifying

lysine residues. (52,53) Hypothetically, Studying both genomic and epigenomic factors is a growing, yet critical approach to understanding disease risk prediction models.

Relton et al proposed a quantitative approach called the Mendelian randomization strategy, in which the impact of exposures or risk factors on the epigenetic profile is identified, after which the causality between the epigenetic profile and the disease outcome is assessed. (54) This process requires a genetic proxy (e.g., a SNP) for the exposure or risk factor, such as alcohol consumption or smoking. The SNP is used as a proxy for DNA methylation levels to assess the causality with the disease outcome. This approach has been applied in CVD research to establish the causal relationship between lipid levels and alcohol intake on disease risk. (55,56)

Strengths and Weaknesses:

One strength of this study involves how the data was collected. Data for the twins was collected together at the same time of day. Also, all DNA methylation samples were taken from the same source, blood, to help minimize variability. Several studies have tested and validated the effectiveness of Illumina Infinium BeadChip 450k methylation array. (40, 57, 58, 59, 60) The ability to examine 485,577 CpG sites at a time is definitely advantageous.

One limitation of this study is that the study population was small in scope and does not represent the general population. Studying only male twins born within a certain decade (1946-1956) that served in the Vietnam War is a very specific cohort. The smaller sample size may limit the external validity of this study, and may also increase inaccuracies in the results. Also, the per sample cost of epigenome-wide profiling remains high. A substantial increase of the sample size would require additional funding support. In addition, the linear mixed effect model did not take into account twin effects to study the effect of environmental differences between twins, as each twin was randomized into each group.

Future Directions:

The next step in this study would be to develop a twin-specific model that would address the environmental differences within twin pairs. Carlin proposes a paired-difference model of the basic format: $E(Y) = \beta_0 + \beta_w(x - \bar{x}) + \beta_B(\bar{x})$ that uses within- and between-pair effects, where the expected value changes with every unit increase in the difference between the observed x and average, $\bar{x}$, or for every unit increase in $\bar{x}$. (61) Adding the within-pair effect to the model prevents treating twin pairs as individuals. Future research could also tackle some of the weaknesses above, such as small sample size. Also, after identifying methylated CpG sites significantly associated with intima-media thickness, future studies could further analyze the DNA methylation at these sites and possibly in CpG islands surrounding them. Identifying CYP4X1 as a gene of interest could allow

future research to study how DNA methylation affects its ability to regulate cholesterol and steroid formation. EPN3 and PROP1 could also be further examined to study how DNA methylation and inflammatory response are related. In addition, the Emory Twin Study recorded mental health and stress data that could also be studied to find a potential relationship with DNA methylation.

References:

1. Dong C, Yoon W, Goldschmidt-Clermont PJ. DNA Methylation and Atherosclerosis. Journal of Nutrition 2002; 132:2406S-2409S.
2. Rodriguez-Saldana J, Rodriguez-Flores M, Cantu-Brito C, Aguirre-Garcia J. A Pathological Study of the Epidemiology of Atherosclerosis in Mexico City. Cardiolog Research and Practice 2014; 1-8.
3. Mozaffarian D, et al. Heart Disease and Stroke Statistics – 2015 Update. Circulation 2015;131:e1-e194.
4. Go AS, Mozaffarian D, Roger VL, Benjamin EJ, Berry JD, Blaha MJ, Dai S, Ford ES, Fox CS, Franco S, Fullerton HJ, Gillespie C, Hailpern S, Heit JA, Howard VJ, Huffman MD, Judd SE, Kissela BM, Kittner SJ, Lackland DT, Lichtman JH, Lisabeth LD, Mackey RH, Magid DJ, Marcus GM, Marelli A, Matchar DB, McGuire DK, Mohler ER, Moy CS, Mussolino ME, Neumar RW, Nichol G, Pandey DK, Paynter NP, Reeves MJ, Sorlie PD, Stein J, Towfighi A, Turan T, Virani S, Wong ND, Woo D, Turner M. Heart Disease and Stroke Statistics – 2014 Update: A Report from the American Heart Association. Circulation 2013, 129:e-28-e292.

5. Mahmood SS, Levy D, Vasan RS, Wang TJ. The Framingham Heart Study and the epidemiology of cardiovascular disease: a historic perspective. Lancet 2014; 383:999-1008.
6. Danaei G, Ding EL, Mozaffarian D, Taylor B, Rehm J, Murray CJL, Ezzati M. The preventable causes of death in the United States: Comparative risk assessment of dietary, lifestyle, and metabolic risk factors. PLoS Medicine 2009;6(4):e1000058.
7. Fuster V, Kelly B. Institute of Medicine. Promoting Cardiovascular Health in the Developing World: A Critical Challenge to Achieve Global Health 2010.
8. Turunen MP, Aavik E, Yla-Herttuala S. Epigenetics and atherosclerosis. Biochimica et Biophysica Acta 2009; 1790:886-891.
9. Ross R. The pathogenesis of Atherosclerosis – An Update. New England Journal of Medicine 1986; 314(8):488-500.
10. Bis JC, White CC, Franceschini N, Brody J, Zhang X, Muzny D, Santibanez J, Gibbs R, Liu X, Lin H, Boerwinkle E, Psaty B, North K, Cupples LA, O'Donnell CJ. Sequencing of 2 Subclinical Atherosclerosis Candidate Regions in 3669 Individuals: Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium Targeted Sequencing Study. Circ Cardiovasc Genet 2014; 7:359-364.
11. Jashari F, Ibrahimi P, Nicoll R, Bajraktari G, Wester P, Henein M. Coronary and carotid atherosclerosis: Similarities and differences. Atherosclerosis 2013; 227:193-200.

12. Mensah GA, Brown DW, Croft JB, Greenlund KJ. Major Coronary Risk Factors and Death from Coronary Heart Disease. American Journal of Preventive Medicine 2005; 29(5):68-74.
13. Lorenz MW, von Kegler S, Steinmetz H, Markus HS, Sitzer M. Carotid Intima-Media Thickening Indicates a Higher Vascular Risk Across a Wide Age Range: Prospective Data from the Carotid Atherosclerosis Progression Study (CAPS). Stroke 2006; 37:87-92.
14. Peters SAE, Bots ML. Carotid Intima-Media Thickness Studies: Study Design and Data Analysis. Journal of Stroke 2013; 15(1):38-48.
15. Stein JH, Korcarz CE, Hurst RT, Lonn E, Kendall CB, Mohler ER, Najjar SS, Rembold CM, Post WS. Use of Carotid Ultrasound to Identify Subclinical Vascular Disease and Evaluate Cardiovascular Disease Risk: A Consensus Statement from the American Society of Echocardiography Carotid Intima-Media Thickness Task Force Endorsed by the Society for Vascular Medicine. Journal of the American Society of Echocardiography 2008; 21(2):93-111.
16. Bots ML, Baldassarre D, Simon A, de Groot E, O'Leary DH, Riley W, Kastelein JJ, Grobbee DE. Carotid intima-media thickness and coronary atherosclerosis: weak or strong relations? European Heart Journal 2007; 28:398-406.
17. Onut R, Balanescu S, Constantinescu D, Calmac L, Marinescu M, Dorobantu M. Imaging Atherosclerosis by Carotid Intima-media Thickness in vivo: How to, Where and in Whom? Journal of Clinical Medicine 2012; 7(2):153-162.

18. Salonen JT, Salonen R. Ultrasound B-mode imaging in observational studies of atherosclerotic progression. Circulation 1993; 87:1156-65.
19. Polak JF, Szklo M, O'Leary D. Associations of Coronary Heart Disease with Common Carotid Artery Near and Far Wall Intima-Media Thickness: The Multi-Ethnic Study of Atherosclerosis. Journal of the American Society of Echocardiography 2015. 28(9):1114-1121.
20. Wendelhag I, Gustavsson T, Suurküla M, Berglund G, Wikstrand J. Ultrasound measurement of wall thickness in the carotid artery: fundamental principles and description of a computerized analyzing system. Clinical Physiology 1991. 11:565-577.
21. Polak JF, Pencina MJ, Pencina KM, O'Donnell CJ, Wolf PA, D'Agostino RB. Carotid-Wall Intima-Media Thickness and Cardiovascular Events. New England Journal of Medicine 2011. 365(3):213-221.
22. Casella IB, Presti C, Porta RMP, Sabbag CRD, Bosch MA, Yamazaki Y. A Practical Protocol to Measure Common Carotid Artery Intima-Media Thickness. Clinics 2008;63:515-520.
23. Hiltunen MO, Yla-Herttuala S. DNA Methylation, Smooth Muscle Cells, and Atherogenesis. Arterioscler Thromb Vasc Biol 2003; 23:1750-1753.
24. Bell JT, Spector TD. Approach to unraveling epigenetics. Trends in Genetics 2011; 27(3):116-125.

25. Gentilini D, Mari D, Castaldi D, Remondini D, Ogliari G, Ostan R, Bucci L, Sirchia S, Tabano S, Cavagnini F, Monti D, Franceschi C, Di Blasio AM, Vitale G. AGE 2013. 35:1961-1973.

26. Dick KJ, Nelson CP, Tsaprouni L, Sandling JK, Aissi D, Wahl S, Meduri E, Morange P, Gagnon F, Grallert H, Waldenberger M, Peters A, Erdmann J, Hengstenberg C, Cambien F, Goodall AH, Ouwehand WH, Schunkert H, Thompson JR, Spector TD, Gieger C, Tregouet D, Deloukas P, Samni N. DNA methylation and body-mass index: a genome-wide analysis. Lancet 2014;383:1990-1998.

27. Ehrlich M, Gama-Sosa MA, Huang L, Midgett RM, Kuo KC, McCune RA, Gehrke C. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or cells. Nucleic Acids Research 1982; 10(8):2709-2721.

28. Jabbari K, Bernardi G. Cytosine methylation and CpG, TpG (CpA) and TpA frequencies. Gene 2004;333:143-149.

29. Sharma P, Kumar J, Garg G, Kumar A, Patowary A, Karthikeyan G, Ramakrishnan L, Brahmachari V, Sengupta S. Detection of altered global DNA methylation in coronary artery disease patients. DNA Cell Biol 2008; 27(7):357:365.

30. Stenvinkel P, Karimi M, Johansson S, Axelsson J, Suliman M, Lindholm B, Heimburger O, Barany P, Alvestrand A, Nordfords L, Qureshi AR, Ekstrom TJ, Schalling M. Impact of inflammation on epigenetic DNA methylation – a novel risk factor for cardiovascular disease? Journal of Internal Medicine 2007; 261:488-

499.

31. Wong CCY, Meaburn EL, Ronald A, Price TS, Jeffries AR, Schalkwyk LC, Plomin R, Mill J. Methylomic analysis of monozygotic twins discordant for autism spectrum disorder and related behavioural traits. Molecular Psychiatry 2014; 19:495-503.
32. Kaminsky ZA, Tang T, Wang S, Ptak C, Oh GHT, Wong AHC, Feldcamp LA, Virtanen C, Halfvarson J, Tysk C, McRae AF, Visscher PM, Montgomery GW, Gottesman II, Martin NG, Petronis A. DNA methylation profiles in monozygotic and dizygotic twins. Nature Genetics 2009;41(2):240-245.
33. Bell JT, Saffery R. The value of twins in epigenetic epidemiology. International Journal of Epidemiology 2012; 41:140-150.
34. Goldberg J, Curran B, Vitek ME, Henderson WG, Boyko EJ. The Vietnam Era Twin Registry. Twin Res 2002. 5(5):476-481.
35. Vaccarino V, Khan D, Votaw J, Faber T, Veledar E, Jones DP, Goldberg J, Raggi P, Quyyumi AA, Bremner JD. Inflammation is Related to Coronoary Flow Reserve Detected by Positron Emission Tomography in Asymptomatic Male Twins. Journal of the American College of Cardiology 2011; 57(11):1271-1279.
36. Rooks C, Veledar E, Goldberg J, Bremner JD, Vaccarino V. Early Trauma and Inflammation: Role of Familial Factors in a Study of Twins. Psychosom Med 2012; 74(2):146-152.
37. Sun YV, Goldberg J, Jones D, Vaccarino V. Epigenome-wide association with soluble cell adhesion molecules among monozygotic twins. Circulation. 2014.

38. Chen Y, Lemire M, Choufani S, Butcher D, Grafodatskaya D, Zanke BW, Gallinger S, Hudson TJ, Weksberg R. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics 2013; 8(2):203-209.
39. Tang R, Hennig M, Bond MG, Hollweck R, Mancia G, Zanchetti A. Quality control of B-mode ultrasonic measurement of carotid artery intima-media thickness: the European Lacidipine Study on Atherosclerosis. Journal of Hypertension 2005. 23:1047-1054.
40. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, Wiencke JK, Kelsey KT. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics 2012; 13:86.
41. Bock C. Analysing and interpreting DNA methylation data. Genetics 2012; 13:705-719.
42. Devlin B, Roeder K. Genomic Control for Association Studies. Biometrics 1999. 55:997-1004.
43. Prium RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, WIller CJ. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 2010. 26:2336-2337.
44. Savas U, Hsu MH, Griffin KJ, Bell DR, Johnson EF. Conditional regulation of the human CYP4X1 and CYP4Z1 genes. Archives of Biochemistry and Biophysics 2005;436:377-385.

45. Luceri C, Giovannelli L, Pitozzi V, Toti S, Castagnini C, Routaboul JM, Lepiniec L, Larrosa M, Dolaro P. Liver and colon DNA oxidative damage and gene expression profiles of rats fed Arabidopsis thaliana mutant seeds containing contrasted flavonoids. Food and Chemical Toxicology 2008; 46:1213-1220.
46. Watkins BA, Hannon K, Ferruzzi M, Li Y. Dietary PUFA and flavonoids as deterrents for environmental pollutants. J Nutr Biochem 2007;18(3):196-205.
47. Heneweer M, Houtman R, Poortman J, Groot M, Maliepaard C, Peijnenburg A. Estrogenic effects in the immature rat uterus after dietary exposure to ethinylestradiol and zearalenone using a systems biology approach. Toxicological Sciences 2007;99(10:303-314.
48. Zandberg P, Peters JLM, Demacker PNM, de Reeder EG, Smit MJ, Meuleman DG. Comparison of the antiatherosclerotic effect of tibolone with that of estradiol and ethinyl estradiol in cholesterol-fed, ovariectomized rabbits. Meopause 2001;8(2):96-105.
49. Wiseman DA, Wells SM, Hubbard M, Welker JE, Black SM. Alterations in zinc homeostasis underlie endothelial cell death induced by oxidative stress from acute exposure to hydrogen peroxide. Am J Physiol Lung Cell Mol Physiol 2005;292:L165-L177.
50. Harper JM, Salmon AB, Chang Y, Bonkowski M, Bartke A, Miller RA. Stress resistance and aging: Influence of genes and nutrition. Mech Ageing Dev 2006;127(8):687-694.

51. Relton CL, Smith GD. Epigenetic Epidemiology of Common Complex Disease: Prospects for Prediction, Prevention, and Treatment. PLoS Medicine 2010; 7(10):e1000356.
52. Ding EL, Hu FB. Cancer and cholesterol: understanding the V-shaped association in patients with diabetes. Canadian Medical Association Journal 2008. 179(5):403-404.
53. Meaney S. Epigenetic regulation of cholesterol homeostasis. Frontiers in Genetics 2014. 5(311):1-10.
54. Relton CL, Smith GD. Two-step epigenetic Mendelian randomization: a strategy for establishing the causal role of epigenetic processes in pathways to disease. International Journal of Epidemiology 2012;41:161-176.
55. Timpson N, Wade KH, Smith GD. Mendelian randomization studies: recent methodological developments. American College of Cardiology 2015.
56. Kamstrup PR, Tybjaerg-Hansen A, Steffensen R, Nordestgaard BG. Genetically Elevated Lipoprotein(a) and increased risk of myocardial infarction. JAMA 2009;301(22):2331-2339.
57. Sandoval J, Heyn H, Moran S, Serra-Musach J, Pujana MA, Bibikova M, Esteller M. Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. Epigenetics 2011;6(6):692-702.
58. Wang D, Yan L, Hu Q, Sucheston LE, Higgins MJ, Ambrosone CB, Johnson CS, Smiraglia DJ, Liu S. Bioinformatics 2012; 28(5):729-730.

59. Dedeurwaerder S, Defrance M, Calonne E, Denis H, Sotiriou C, Fuks F. Evaluation of the Infinium Methylation 450k technology. Epigenomics 2011; 3(6):771-784.

60. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le J, Delano D, Zhang L, Schroth G, Gunderson K, Fan J, Shen R. High density DNA methylation array with single CpG site resolution. Genomics 2011; 98:288-295.

61. Carlin JB, Gurrin LC, Sterne JAC, Morley R, Dwyer T. Regression models for twin studies: a critical review. International Journal of Epidemiology 2005. 34:1089-1099.

Table 1: Summary Statistics

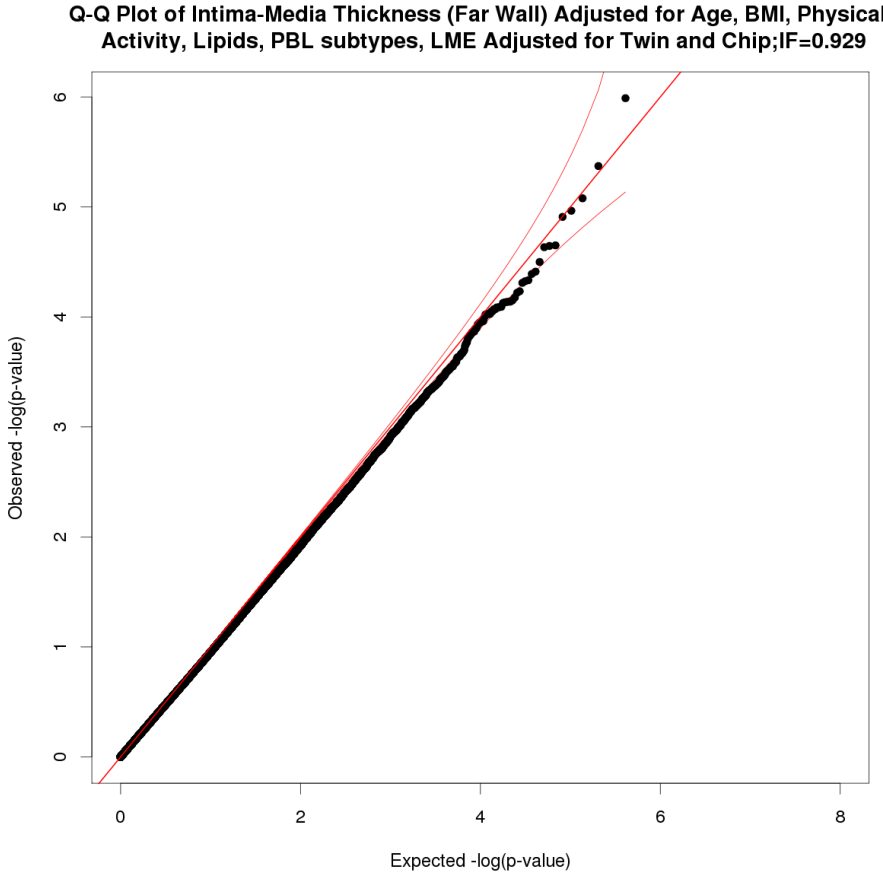
Variable		Group 1 mean	Group 2 mean
IMT_max_tot (mm)		0.7655	0.7686
IMT_max_FW (mm)		0.7659	0.7584
age (years)		55.63	55.69
body mass index (kilograms / meter)		28.38	29.87
current smokers (count)		26 (37%)	21 (30%)
physical activity (count)	low	7 (10%)	10 (14%)
	med	45 (64%)	43 (61%)
	high	18 (26%)	17 (24%)
HDL cholesterol (mg/dL)		38.95	37.54
LDL cholesterol (mg/dL)		117.97	124.20

Table 2: CpG Sites

*** Genes identified using UCSC Genome Browser. https://genome.ucsc.edu/**

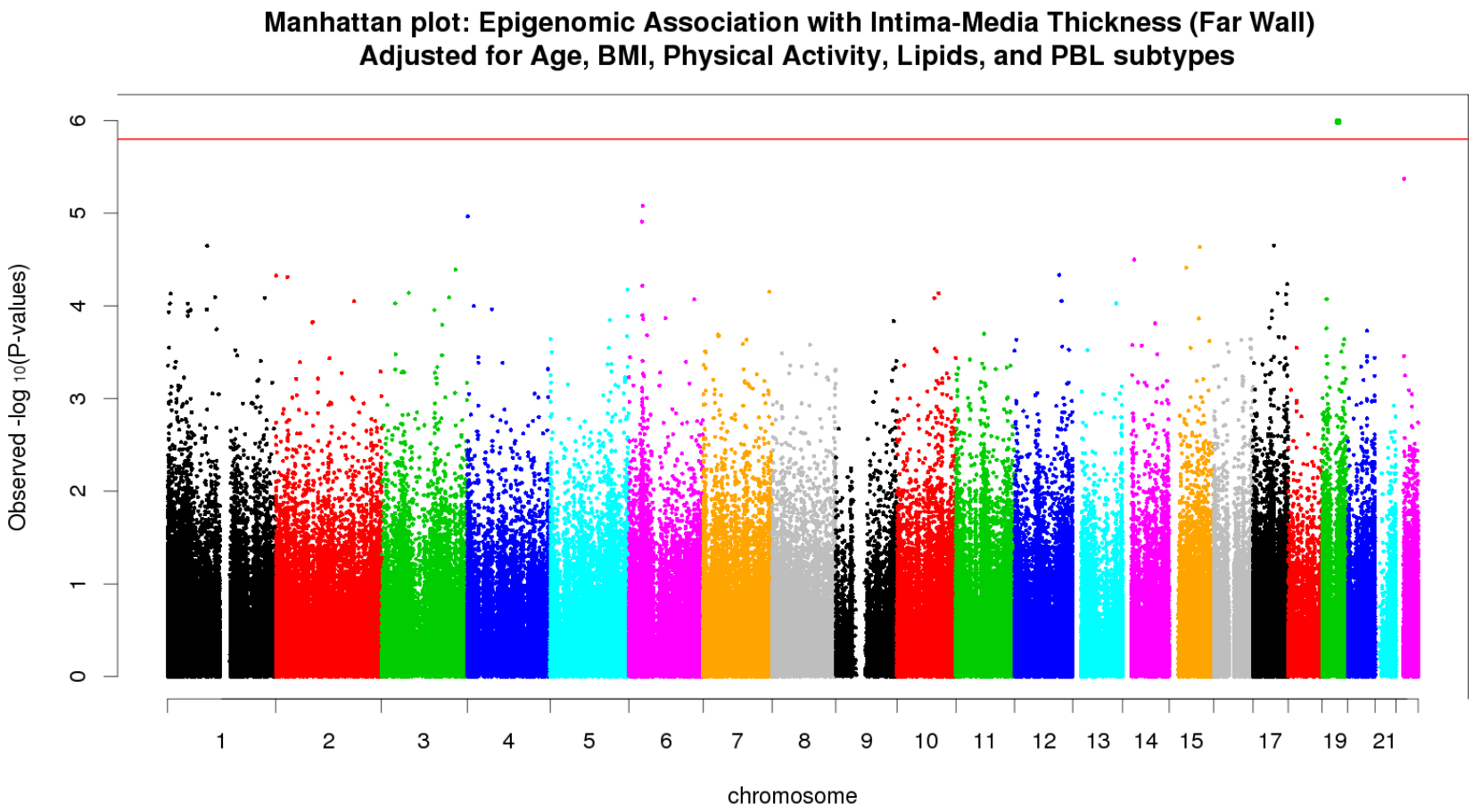
CpG site	Chromosome	RefSeq Gene*	IMT_max_FW E	IMT_max_FW p-value
cg04085789	22		0.1076	4.25×10^{-5}
cg06945936	19	KIRREL2	0.0378	1.03×10^{-6}
			IMT_max_tot E	**IMT_max_tot p-value**
cg04085789	22		0.1386	5.76×10^{-6}
cg00157962	1	FAM102B	0.1156	1.29×10^{-5}
cg00187229	4	CPLX1	0.1210	1.07×10^{-5}
cg02005758	3	RAB7A	0.0935	1.18×10^{-5}
cg03534504	5	STK10	0.0772	1.89×10^{-5}
cg05257202	13	UBAC2	0.1262	7.77×10^{-6}
cg09827751	17	EPN3	0.1035	1.50×10^{-5}
cg10751811	1	CYP4X1	0.2287	1.68×10^{-5}
cg10799327	3		0.0785	1.88×10^{-5}
cg11551560	15		0.1080	3.50×10^{-6}
cg26217402	14	C14orf43	0.0615	1.75×10^{-5}
cg26532621	5	PROP1	0.1065	1.39×10^{-5}

Figure 1: Q-Q Plot (a), Manhattan Plot (b), and CpG Regional Plot (c) for the Top Associations for Intima-Media Thickness (Far Wall)



(a)

(b)


Manhattan plot: Epigenomic Association with Intima-Media Thickness (Far Wall)
Adjusted for Age, BMI, Physical Activity, Lipids, and PBL subtypes
Observed -log 10(P-values)
0
1
2
3
4
5
6
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
17
19
21
chromosome

(c) CpG Regional Plot for cg10740660, cg06945936, cg04085789

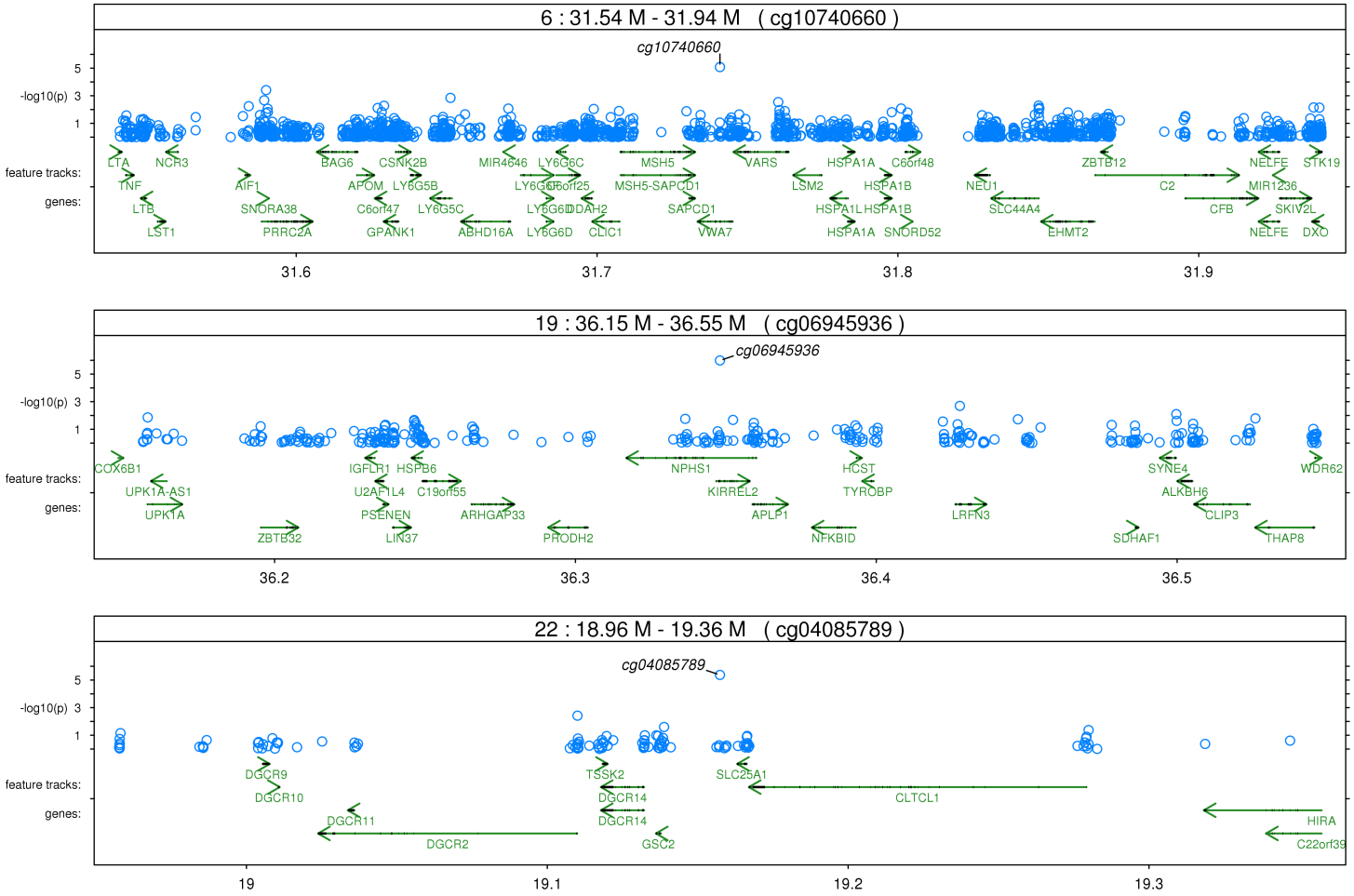
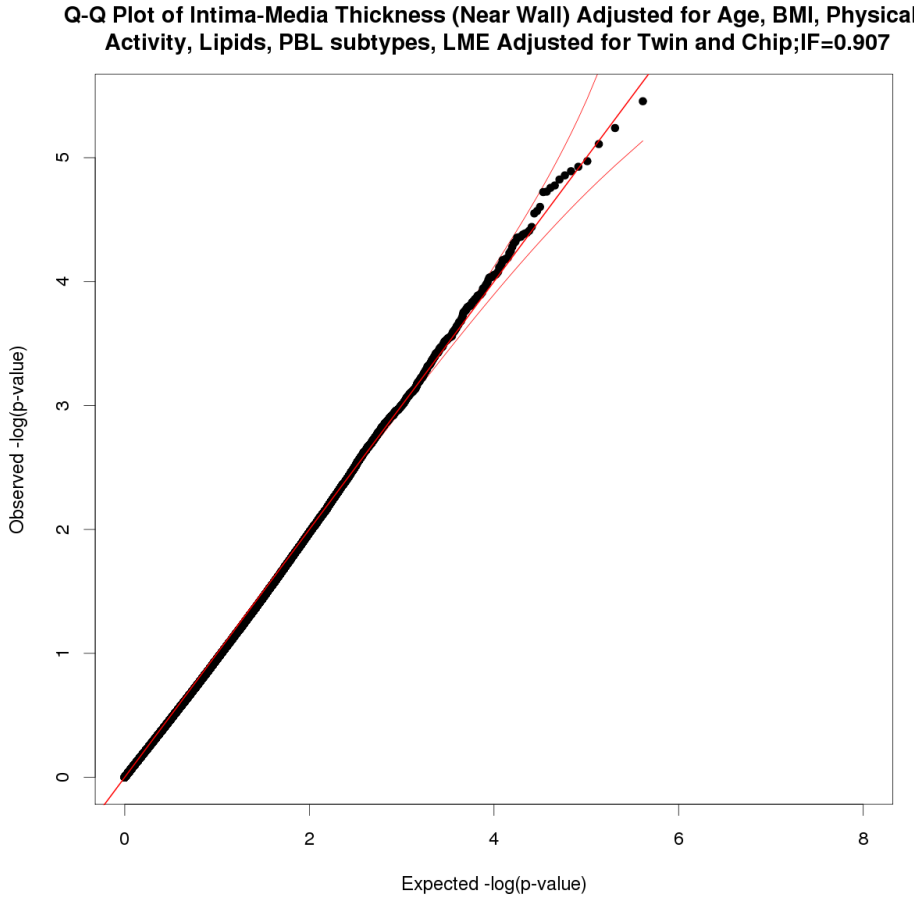
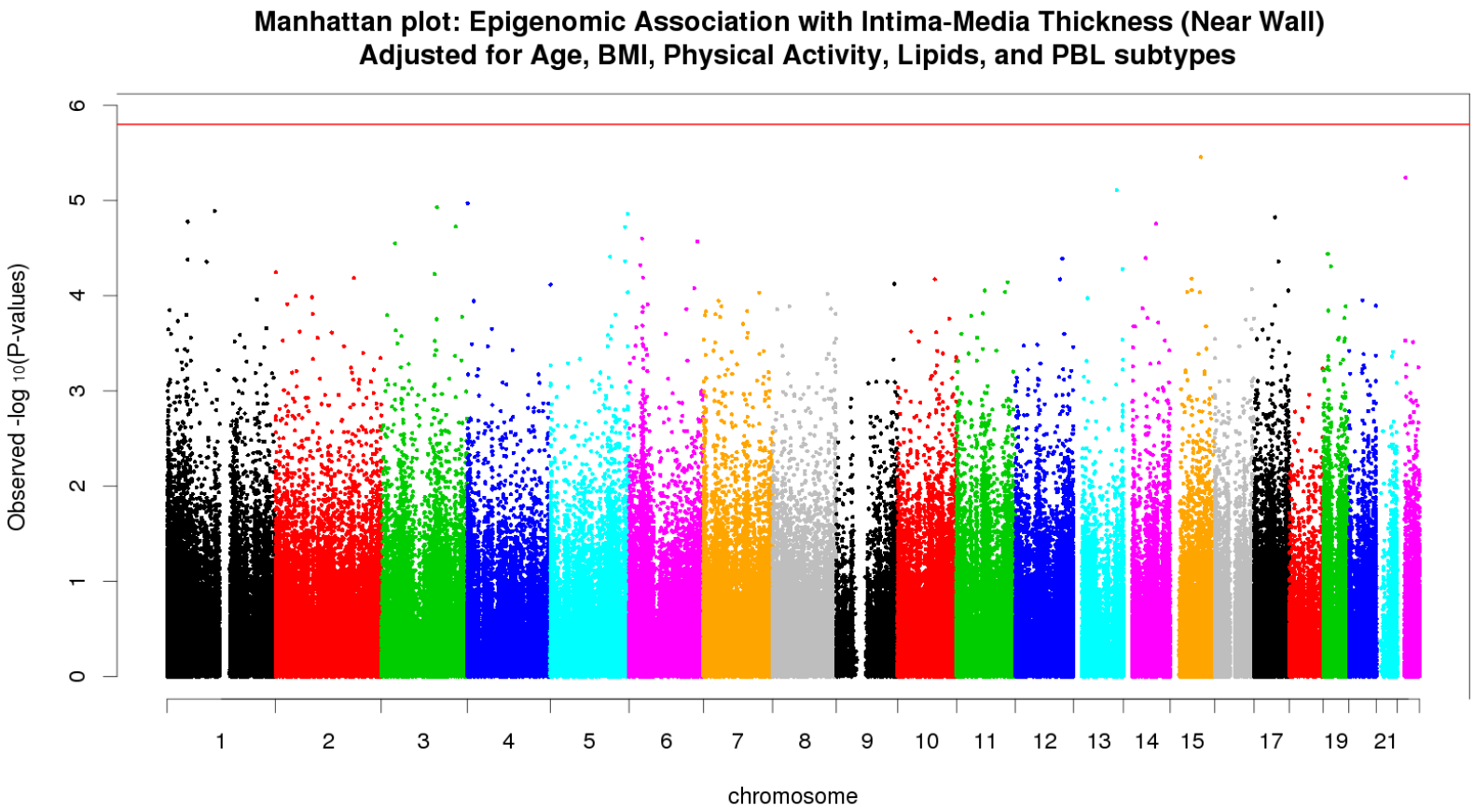


Figure 2: Q-Q Plot (a), Manhattan Plot (b), and CpG Regional Plot (c) for the Top Associations for Intima-Media Thickness (Near Wall)



(a)

(b)


Manhattan plot: Epigenomic Association with Intima-Media Thickness (Near Wall)
Adjusted for Age, BMI, Physical Activity, Lipids, and PBL subtypes
Observed -log 10(P-values)
0
1
2
3
4
5
6
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
17
19
21
chromosome

(c) CpG Regional Plot for cg05257202, cg11551560, and cg0405789.