THE INFLUENCE OF BODY MASS INDEX ON GLOBAL DNA METHYLATION LEVELS IN BLOOD LEUKOCYTES

by

Nora Zwingerman

A thesis submitted to the Department of Community Health and Epidemiology

In conformity with the requirements for

the degree of Masters of Science

Queen's University

Kingston, Ontario, Canada

March, 2012

Copyright © Nora Zwingerman, 2012

Abstract

Introduction: Body Mass Index (BMI) is a relative measure of whether an individual's weight is at a healthy level for their height. A higher BMI is associated with an increased risk of cancer and cardiovascular disease (CVD). However, the biologic mechanisms are not well understood. One proposed mechanism is through changes in global DNA methylation levels, particularly global DNA hypomethylation. Global DNA hypomethylation refers to lower levels of DNA methylation across the entire genome and hypermethylation refers to higher levels of DNA methylation across the entire genome. Changes in methylation levels can affect gene expression, genomic stability, and chromosomal structure. The methylation status of repetitive sequences in the DNA, such as LINE-1, is commonly used to represent a surrogate measure of global DNA methylation levels.

Objectives: 1. Quantify and describe LINE-1 DNA methylation in leukocytes in a large sample of healthy volunteers.

2. Examine the relationship between BMI and LINE-1 DNA methylation levels.

3. Assess if sex is an effect modifier of the relationship between BMI and LINE-1 DNA methylation levels.

Methods: A nested cross-sectional study was composed of 502 healthy volunteers between the ages of 20 and 50. Subjects completed a study questionnaire and provided blood samples for laboratory analyses. For each subject, DNA was isolated, underwent bisulfite conversion, and LINE-1 DNA methylation levels were measured by Polymerase Chain Reaction (PCR) High-Resolution Melting Curve analysis. For the main analysis, a multivariate linear

regression model was used to examine the relationship between BMI and LINE-1 DNA methylation levels, while controlling for confounders.

Results: LINE-1 DNA methylation was normally distributed with a mean of 84.52% and a standard deviation of 3.19%. BMI (normal, overweight, and obese categories) was not significantly associated with LINE-1 DNA methylation levels in the adjusted linear regression model (p=0.41) and the interaction term between BMI and sex was not significant (p=0.50).

Conclusions: LINE-1 DNA methylation was measured with a high degree of reliability in a sample of healthy volunteers. This research provided a description of LINE-1 DNA methylation levels in a large healthy population and showed that BMI was not associated with global DNA methylation.

Acknowledgements

This process was unlike any I have experienced. I could not have completed it without the assistance of many people and I am grateful to everyone who made my journey possible and wonderful.

I want to express my sincerest gratitude to Dr. Will King for his support throughout this process and for constantly broadening my understanding of epidemiology. Working with you was truly inspiring. I would like to thank Dr. Stephen Pang for all of his assistance and for welcoming me into his laboratory, which was a great atmosphere to be able to work in.

Thanks to the Department of Community Health and Epidemiology for allowing me to partake in this program and for providing a positive work environment. In the department, I feel honoured to have been surrounded by a research group who is supportive, motivating, and always fun to be around. I am very grateful to have had the opportunity to have learnt from all of you. A special thanks to Vikki Ho, Sarah Peacock, Melanie Walker, Celine Simmers, and Dr. Janet Ashbury. I appreciate all the time I was able to spend with Dr. M. Yat Tse, it was a wonderful experience to have your expertise to guide and help me in the lab.

Finally, thanks to my family for consistently being there for me whenever I needed them. Specifically, my parents for their moral support and willingness to listen to my ranting even when they do not understand what I am talking about and my sister for being a best friend to me and her comic relief.

Table of Contents

List of Figures

List of Tables

List of Abbreviation

BMI	Body Mass Index
BSC	Bisulfite Converted
Cp	Crossing Point
CpG	Cytosine-phosphate-Guanine
CVD	Cardiovascular Disease
DEXA	Dual-Energy X-ray Absorptiometry
DNMT	DNA Methyltransferase
DNA	Deoxyribonucleic Acid
Hcy	Homocysteine
HDL	High Density Lipoprotein
HIF	Hypoxia-Inducible Factor
HPLC	High Performance Liquid Chromatography
HRM	High-Resolution Melt
IGF	Insulin-like Growth Factor
IL	Interleukin
IPAQ	International Physical Activity Questionnaire
LDL	Low Density Lipoprotein
LINE	Long Interspersed Nuclear Element
MET	Metabolic Equivalent
MS	Mass Spectrometry
MRI	Magnetic Resonance Imaging
PCR	Polymerase Chain Reaction
REB	Research Ethics Board
SAH	S-Adenosyl-Homocysteine
SAM	S-Adenosyl-Methionine
Sat2	Human Satellite 2 Repeat
SINE	Short Interspersed Nuclear Element
VEGF	Vascular Endothelial Growth Factor
WHO	World Health Organization

Chapter 1

Introduction

1.1 General Introduction

Body Mass Index (BMI) is a relative measure of whether an individual's weight is at a healthy level for their height. BMI is the most widely used measure to categorize individuals as underweight, normal, overweight, or obese. The World Health Organization (WHO) has recommended widely accepted guidelines for categorizing individuals based on their BMI (1,2). In this study, the BMI variable was used to classify individuals as normal, overweight, or obese. This classification has been used in many studies to evaluate the increased risk for disease, including cancer and cardiovascular disease (CVD), for individuals who are overweight or obese.

In a large study population, BMI is used as a practical measure of body fat because BMI is an easily obtainable measure at minor costs. In adult populations the correlations between BMI and densitometry measurements range from moderate to strong (3).

The prevalence of overweight and obesity in developed countries has been increasing over the last couple of decades and is continuing to increase. In Canada, it is currently estimated that over half of adults are overweight or obese (4). Those who are overweight or obese are at increased risk for many adverse health outcomes, including cancer and CVD. However, the biologic pathway between overweight or obese and disease outcomes is not well understood.

The aim of this research was to gain insight into a potential intermediate endpoint, DNA methylation, through which obesity may confer an increased risk of adverse health outcomes, including cancer or CVD.

Epigenetic changes are defined as modifications to DNA that leave the DNA sequence intact. DNA methylation (the most common epigenetic change) refers to the addition of a methyl group to the 5-carbon position of a cytosine to form a 5-methyl cytosine in a CpG (Cytosine-phosphate-Guanine) dinucleotide sequence in DNA. DNA methylation plays a vital role in cells by controlling cell function, regulating gene expression, and impacting upon DNA stability in the structure of the chromatin (5). The One-Carbon Metabolism Cycle is the intracellular process that is responsible for methylating biochemical compounds, including DNA. This cycle is primarily comprised of S-Adenosyl-Methionine (SAM), S-Adenosyl-Homocysteine (SAH), and homocysteine. The One-Carbon Metabolism Cycle is affected by a variety of dietary factors including methionine, Vitamin B12 and B6, zinc, and folate. Abnormal levels of cofactors involved in the One-Carbon Metabolism Cycle have been associated with several adverse health outcomes (6-9). Two distinct types of DNA methylation are considered in relation to adverse health outcomes: global and gene-specific. Global hypomethylation, which refers to an overall decrease in the number of methylated cytosines across the entire genome, has been suggested to be an early event in carcinogenesis and is the focus of this thesis. Global DNA hypomethylation has been shown to adversely affect the regulation of DNA by changing genomic stability, increasing mutation rates, and modifying gene expression (10). Paradoxically, global hypomethylation is often accompanied by gene-specific increases in methylation levels (11).

DNA methylation is an emerging field of research. Reliable methods for measuring DNA methylation have recently been developed that can be applied to a large study population. The method used for this thesis measured the methylation level of the Long Interspersed Nuclear Element 1 (LINE-1) sequence, which is the most abundant long repetitive sequence in human DNA, present over 500,000 times. The methylation of LINE-1 DNA repeats has been shown to be correlated with global DNA methylation levels (12). LINE-1 DNA methylation levels represent a surrogate measure for global DNA methylation.

The use of DNA methylation levels in epidemiologic research, as both a health outcome and as a predictor of disease, is emerging. However, very few studies have examined the determinants of global DNA methylation changes, such as BMI. This research focused on the relationship between BMI and global DNA methylation. Also, factors affecting the One-Carbon Metabolism Cycle, and therefore indirectly potential determinants of global DNA methylation, were considered as potential confounders. As this thesis is nested within a larger study funded by the Canadian Institutes of Health Research (CIHR), which focused on identifying the determinants of the One-Carbon Metabolism Cycle including: age, gender, alcohol, smoking, and dietary factors, information regarding potential confounders of interest was available for this thesis.

A potential biologic pathway between increased BMI and DNA methylation changes is through inflammation. Adipose tissue is an endocrine organ that secretes adipokines. Adipokines are biologically active molecules that regulate lipid levels, immune function, blood pressure, insulin sensitivity, and vascular growth factors (13,14). In overweight or obese individuals, higher levels of adipokines result in systemic low-grade inflammation (3,13). Studies have identified that inflammation may affect global DNA methylation levels

(15-17). An alternate plausible biologic mechanism through which BMI could influence DNA methylation levels is through oxidative stress (17).

1.2 Rationale

Elevated BMI has been consistently suggested as a risk factor for several cancer sites and CVD. However, these relationships are not well understood with respect to their underlying biologic mechanisms. There is strong evidence to suggest that increased BMI and alterations in DNA methylation levels are both individually associated with adverse health events. However, it is unclear whether increased BMI adversely affects DNA methylation levels leading to detrimental health events.

The use of global DNA methylation as a potential intermediate endpoint on a causal pathway for disease outcomes is beneficial to epidemiologic studies because it permits the use of a healthy study population, the examination of a relationship with a shorter latency and more frequent events, and a potentially stronger relationship. The goal of this study was to investigate DNA methylation as a potential intermediate endpoint linking elevated BMI to increased cancer and CVD risk. This study contributes to the existing literature because few studies have investigated the relationship between BMI and global DNA methylation levels measured in white blood cells in a large healthy population while controlling for potential confounders.

In addition, it is difficult to disentangle the independent role of increased BMI from other related risk factors, such as physical activity, which are strongly correlated with weight. This research will attempt to isolate the relationship between BMI and global DNA methylation levels.

1.3 Objective

The objectives of this thesis were:

1) To quantify and describe LINE-1 DNA methylation levels in leukocytes in a large sample of healthy volunteers.

2) To examine the relationship between Body Mass Index and leukocyte LINE-1 DNA methylation levels in the study sample.

3) To investigate whether sex is an effect modifier of the relationship between BMI and LINE-1 DNA methylation levels.

1.4 Context of Research

This study was conducted within a larger cross-sectional health research program aimed at understanding environmental and lifestyle influences on colon cancer risk through the use of intermediate events. The primary objective of the larger study was to examine the influence of water disinfection by-products on the One-Carbon Metabolism Cycle. The exposure of interest was selected by the student investigator due to an interest in understanding how elevated BMI was associated with a variety of disease outcomes. The larger study collected information regarding lifestyle variables and laboratory measures. Nora Zwingerman contributed to the existing study database by conducting the laboratory analysis on stored whole blood samples to determine the LINE-1 DNA methylation levels for over 500 subjects using a High-Resolution-Melt (HRM) method. This approach was facilitated through collaboration with Dr. S. C. Pang and Dr. M. Yat Tse from the Department of Biomedical and Molecular Sciences at Queen's University.

1.5 Outline of Thesis

This thesis is organized into five chapters. Chapter 2 is a review of the published literature in this area of research and outlines the potential relevance of this research to both cancer and CVD. Chapter 3 contains information regarding the study design and methods, including a description of the study population, data collection and validity, and strategies for data analysis. The results, the main analysis, findings and sensitivity analyses are presented in Chapter 4. Chapter 5, the discussion, focuses on methodological considerations and also describes the generalizability, contributions, and future directions for further research.

Chapter 2

Background and Literature Review

2.1 Body Mass Index

Body Mass Index (BMI) is the most widely used measure to represent an individual's body composition. This section will present relevant information regarding BMI and highlight the importance of expanding the current knowledge of the biological processes linking BMI to adverse health outcomes. The implications of preventing, understanding and managing obesity, along with health consequences directly associated with obesity, are of utmost value due to the increasing prevalence of overweight and obesity in the Canadian population.

2.1.1 Definition

Body Mass Index (BMI), or Quetelet's Index, is a relative measure of whether an individual's weight is at a healthy level for their height. To calculate an individual's BMI, weight in kilograms is divided by height in metres squared (18,19). BMI is the most widely used measure of body composition because it is obtainable at minor costs and permits a standard measure within and between studies (3). The World Health Organization (WHO) has recommended guidelines for categorizing individuals based on their BMI that are widely accepted (1,2,20). BMI is typically divided into categories for underweight ($< 18.5 kg/m^2$), normal ($18.5 - < 25.0 kg/m^2$), overweight ($25.0 - < 30.0 kg/m^2$), and obese ($\geq 30.0 kg/m^2$) (1,2,20). These BMI categories were largely created on the basis of observational epidemiological studies of BMI and associated overall mortality (2).

2.1.2 Descriptive Epidemiology

Over the past 25 years, the BMI of Canadian adults has been increasing (21). In Canada, it is currently estimated that over half of adults are overweight or obese (4). Between 2007 and 2009, the Canadian Health Measures Survey measured BMI for a large sample of Canadians and reported that for 20 to 39 year olds, 37% of males and 23% of females were overweight, and 19% of males and 21% of females were obese (22). Furthermore, the prevalence of overweight and obesity increased in the higher age categories (22). Comparing BMI values from the 2007 - 2009 Canadian Health Measures Survey with data from 1986 - 1992 showed an increase in average BMI, with a shift in the distribution towards higher BMI values (21). The increase in BMI over time and with increasing age corresponded with an increase in waist circumference and skin-folds thickness measures (22). These results indicate that not only are BMI values increasing, but there is also an increase in the amount of abdominal and subcutaneous adipose tissue.

2.1.3 The Relationship between BMI and Body Composition

There are many methods to measure or estimate total body fat. An ideal measure would be able to differentiate between all of the components of an individual's body mass and subtypes of tissues. However, these measurements are costly and require expensive equipment operated by trained technicians. Different measures available include: underwater weighing (hydrometry), bioelectrical impedance, Dual-Energy X-ray Absorptiometry (DEXA), Magnetic Resonance Imaging (MRI), Computed Tomography (CT), and isotope dilution (1). These methods are impractical for large-scale studies and therefore, anthropometric measures such as BMI, waist circumference, and waist-to-hip ratio are more appropriate. In the literature, the correlations between BMI and densitometry (i.e., DEXA)

have ranged from moderate (0.6) to strong (0.9) in adult populations (3,23). The assessment of body fat percentage from BMI was most accurate when taking into account age and sex, meaning that the reliability of BMI as a measure of body fat differs by age group and sex (23,24). Additionally, the traditional BMI calculation may not be appropriate to use in children, adolescents, or the elderly. In the elderly adipose tissue tends to be redistributed and BMI becomes a less valid measure (25,26).

Generally, the healthy or normal category for BMI (18.5 - < 25.0kg/m^2) equates to 15 - 20 percent body fat in adult males and 25 - 30 percent body fat in adult females (27). Sex should be considered when assessing body fat because the healthy range of percentage of body fat is considerably different for males and females and the distribution of body fat in these two groups tends to vary. Males typically present as android or 'apple shape' with a larger amount of abdominal fat, while females are described as gynoid or 'pear shape' with a greater amount of subcutaneous fat located in their thighs and hips (27). A recent study by Romero-Corral et al. evaluated the accuracy of BMI for determining body fat percentage using bioelectrical impedance as the gold standard. This study found that the BMI cut-off for obesity at 30 kg/m^2 had a low sensitivity (43%) and a very high specificity (96%) for detecting excess body fat (26). When the study population was stratified by age, BMI had a higher sensitivity and specificity for individuals less than 60 years of age compared to individuals 60 years of age or older (26).

Even though BMI does not directly measure the quantity or characterize the amount of adipose tissue, clinical predictors of body fat are associated with BMI. These factors include blood triglycerides, cholesterol, blood pressure, estrogens, glucose levels, and insulin levels (3,28).

Often BMI is calculated based on self-reported values of height and weight. Validation studies have shown that self-reported and measured height and weight are highly correlated (Spearman r>0.9). In addition, self-reported BMI values had a high specificity and sensitivity when compared to measured BMI values. However, weight tended to be systematically underestimated in self-reported data (29-31). Weight is more accurately reported for those less than 60 years of age (29).

2.1.4 Biological Implications

Adipose tissue consists of adipocytes, adipose cells, which store fat. Adipose tissue expands when carbohydrates are converted into triglycerides, the most prominent storage fat stored in adipocytes (3). There are two main types of adipose tissue. Subcutaneous adipose tissue is primarily located between the skin and muscle layers. Visceral adipose tissue is found in large body cavities, predominantly in the abdominal cavity. Adipose tissue is recognized to be a complex organ with endocrine functions (32), with visceral adipocytes having greater metabolic activity than subcutaneous adipocytes. Thus, the quantity of intra-abdominal visceral adipose tissue has a stronger relationship to adverse health outcomes than subcutaneous adipose tissue due to the detrimental effects of higher metabolic activity levels (13).

Adipocytes are biologically active cells that release fatty acids, cholesterol, steroid hormones, prostaglandins, and active proteins (13). Adipokines is a term used to describe the active proteins secreted from adipocytes. These proteins play a vital role in regulating lipid levels, immune function, blood pressure, insulin sensitivity, vascular growth factors, energy balance, and inflammation (13,14). Over 50 adipokines have been identified, including: leptin, adiponectin, C-reactive protein, tumor necrosis factor alpha (TNF-α), interleukin-6

(IL-6), acylation stimulating protein (ASP), insulin-like growth factor-1 (IGF-1), and vascular growth factors (VEGF) (13,33). These factors can act locally and at distant sites in the body. The secretion of these proteins is affected by adipocyte size. For example, as an adipocyte increases in size (i.e. hypertrophies) the cell generally secretes a greater amount of adipokines. An exception to this trend is the adipokine adiponectin, an anti-inflammatory cytokine, which is released at lower levels in larger cells. The adipokine C-reactive protein is a proinflammatory signal that is positively correlated to increases in adipose tissue (3,13). Obesity is recognized as a chronic state of generalized low-grade inflammation due to higher levels of circulating adipokines (14) that is hypothesized to account for increased incidence of gastrointestinal diseases (i.e., Crohn's disease, colon cancer, and fatty liver) in obese individuals (13).

There is an interaction between adipocyte growth, inflammation and other biologic processes. For example, as adipose tissue enlarges it may not have an adequate oxygen supply, referred to as hypoxia, which in turn increases the demand for blood supply causing the promotion of angiogenesis potentially mediated by hypoxia-inducible factor-1 (HIF-1) (13). HIF-1 is also a characteristic of neoplasms because as cells multiply and grow they become hypoxic and require nutrients in order to expand. In order to facilitate nutrient supply angiogenesis is promoted (34). HIF-1 induces angiogenic factors, such as vascular endothelial growth factor (VEGF), which increase the production of endothelial cells (14).

As adipocytes grow, the cellular structure adapts to increasing demand and requires more glucose to produce adipokines. The increased amount of glucose results in a greater amount of reactive oxygen species being generated from the mitochondria and causes oxidative stress (13).

In males and post-menopausal women, adipose tissue is the primary site for the production of sex steroids such as estrogen and estradiol. These hormones play an important role in cellular function by regulating differentiation, proliferation, and apoptosis (3). IGF-1 is a peptide hormone which functions in a similar capacity to insulin. *In vitro* IGF-1 has been demonstrated to promote cell proliferation and inhibit apoptosis. In animal models these mechanisms result in an increase in carcinogenesis (3).

Overall, there are many cellular level changes and pathways that are activated as a result of adipose tissue. Several candidate mechanisms may explain the association between adipose tissue and adverse health events. These different candidate mechanisms have the potential to interact with one another. Currently, the biological mechanisms linking elevated BMI to adverse health outcomes, such as cancer and CVD, are not well understood and they could potentially act through processes associated with inflammation, lipids, hormones, and oxidative stress (13,35,36).

2.1.5 Determinants of BMI

An elevated BMI is a result of the combination of multiple factors including: genetic, metabolic, behavioral, and environmental factors. Some of the behavioral risk factors for obesity include: diet, physical activity, alcohol consumption, and smoking. Diet has consistently been shown to be a determinant of body size. There is a positive relationship between increased BMI and the intake of fat and calories (37). More specifically, energy-dense food, sugary drinks, and fast food consumption are associated with an increased risk for weight gain (27).

Not only is the quantity and type of food consumed related to BMI but also the amount of energy expended during daily activities and exercise are determinants of BMI.

Greater amounts of physical activity are associated with a lower risk for weight gain and obesity (27,37-39). Furthermore, a sedentary or inactive lifestyle has been associated with an increased risk of weight gain and obesity (27).

Current smoking status has been reported to have an inverse association with BMI (40,41); however, Canadian men and women who are former smokers appear to be more likely to be obese (21). Inconsistent results have been reported for alcohol consumption. Overall, more studies report a positive relationship between high alcohol consumption and elevated BMI (42-44).

Non-modifiable risk factors for an overweight and obese BMI are age, sex, and ethnicity. Increased BMI becomes more prevalent with increasing age for males and females (21,41). Males tend to have a greater proportion of intra-abdominal adipose tissue compared to women prior to menopause. However, women after menopause have a similar distribution of adipose to males (45). Ethnicity is related to body size and distribution of fat. Similar BMI values in different ethnic population may represent different body compositions (distribution of adiposity) (46).

Socioeconomic status has been associated with body size. Lower education level and lower income levels have typically been associated with lower BMI values (21,41,47-49).

2.1.6 Adverse Outcomes associated with BMI

Elevated BMI is associated with an increased relative risk between 1.0 and 2.0 for impaired fertility and hormonal abnormalities (27). Stronger relationships with relative risks between 2 and 3 in obese individuals have been reported for coronary heart disease, hypertension, and hyperuricemia (27). The strongest relationships for adverse health events, with relative risks estimates greater than 3 have been reported for type 2 diabetes, gallbladder

disease, dyslipidemia, insulin resistance, and sleep apnea (27). In addition, an elevated BMI is associated with increased risk for several cancer sites. Moderate relative risks (1.5 to 2.0) for obese BMI compared to normal BMI are reported for colorectal cancer and postmenopausal breast cancer. Moderate to strong relative risks (2.0 to 3.0) are reported for endometrial, kidney, and esophageal cancer. Obesity is also likely a risk factor for other cancer sites including the pancreas, liver, gall bladder, and the stomach (35). Overall, obesity significantly decreases life expectancy by approximately 7 years compared to a healthy weight (27).

According to the WHO and Health Canada, normal BMI is associated with the lowest risk for adverse health events. The relationship between BMI and health outcomes tends to show a J-shaped curve due to the increased risk for those who are underweight (50). Underweight and overweight BMI categories confer an increased risk for adverse health events compared to the normal category. Among those with higher than normal BMI there is a pattern of increasing risk for adverse health events with increasing BMI (1,51).

2.1.6.1 Cancer

In 2007, the World Cancer Research Fund conducted a review of the existing evidence and reported an association between increased body fat usually measured by BMI and an increased risk for many cancers such as esophageal, pancreatic, colorectal, breast (post-menopause), endometrial, and renal cancers (27,52). In addition, a large meta-analysis of prospective cohort studies found a significant relationship between elevated BMI and an increased risk for all cancers listed above, as well as thyroid and gallbladder cancers, leukemia, and melanoma (35). Some of these risk estimates varied by sex and ethnic origin. For example, there was a much stronger relationship between high BMI and colon cancer for

males compared to females (35). Not only are those with an increased BMI at a higher risk for cancer but they also have an increased risk of death from the cancer compared to those who are at a healthy weight at the time of cancer diagnosis (50). Current estimates suggest that 15 to 20 percent of all cancer deaths in North America can be attributed to overweight and obesity (50). Due to the rising prevalence of overweight and obesity, the amount of disease that can be attributed to increased body size (i.e. attributable risk) is expected to continue to rise.

2.1.6.2 Cardiovascular Disease (CVD)

The American Heart Association describes obesity as a major risk factor for coronary heart disease (53). Research has identified increased BMI as an independent risk factor for CVD and suggested that the relationship between obesity and CVD is mediated through diabetes, hypertension, and dislipidaemia (54,55). Using slightly different BMI cut-off values compared to the WHO classification, Rimm et al. found that the relative risk for overweight men under 65 years of age was 1.72 and reported relative risks of 2.61 and 3.44 for the two obese categories (obese, very obese) respectively (56). Additionally, obesity, particularly intra-abdominal adiposity, has been shown to be associated with an increase in coronary artery disease mortality (57,58).

2.2 DNA Methylation

Epigenetics is an emerging field of research and may play a vital role in the understanding of biologic pathways linking modifiable or lifestyle risk factors with disease states. Epigenetic changes are defined as reversible modifications to DNA that leave the DNA sequence intact. DNA methylation is the most common epigenetic change. Epigenetics

plays an integral component to normal cell functioning. Specifically, DNA methylation regulates gene expression.

2.2.1 Definition and Description

DNA methylation is a postreplicative modification of DNA that occurs when a methyl group (CH_3) is added to the 5-carbon position on a cytosine nucleotide that is located in a Cytosine-phosphate-Guanine (CpG) dinucleotide pair (59,60). Figure 2-1 depicts the nucleotide cytosine in an unmethylated and methylated form.

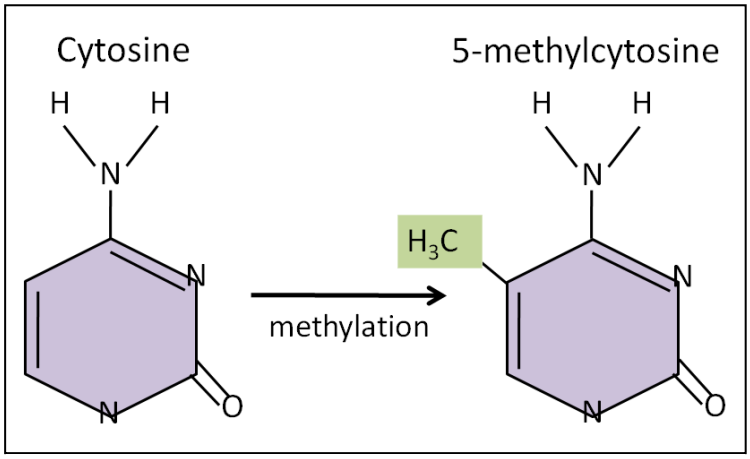

Figure 2-1. Methylation of a Cytosine Nucleotide. Chemical structure of a cytosine nucleotide in an unmethylated and a methylated form.

DNA methylation patterns are conserved during replication. The CpG dinucleotide pairing is unique because it has the same sequence, in the opposite orientation, on the complementary strand of DNA. This is depicted in Figure 2-2. This permits an enzyme called maintenance DNA methyltransferase (DNMT1) to maintain DNA methylation during replication (61).

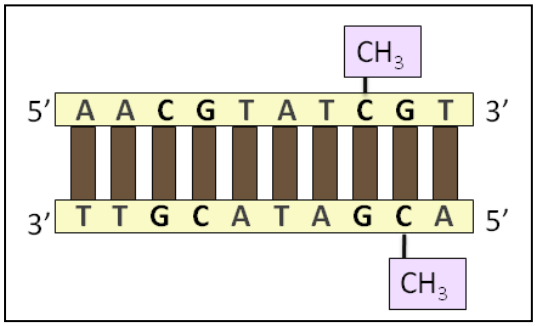

Figure 2-2. Complementary Pairing of CpG Dinucleotides. A segment of DNA with DNA methylation at complementary CpG sites.

2.2.1.1 One-Carbon Metabolism Cycle

The One-Carbon Metabolism cycle (depicted in Figure 2-3) is the process responsible for attaching a methyl group to a CpG dinucleotide. Methionine, mainly from the diet, is converted into S-Adenosyl-Methionine (SAM), the methyl donor. DNA methyltranferases (DNMTs) are the enzymes responsible for removing a methyl group from SAM and attaching it onto a cytosine nucleotide, which converts SAM to S-Adenosyl-Homocysteine (SAH). Then SAH is converted to homocysteine (Hcy), which can be recycled back to methionine by the enzyme methionine-synthase. The normal steady state of this cycle relies on the removal or conversion of homocysteine because biochemically the reaction favors the production of SAH from homocysteine. SAH binds to DNMTs with a higher affinity than SAM resulting in the inhibition of DNA methylation by SAH (62). Yi et al. demonstrated that increased homocysteine levels were positively related to SAH levels, that there were strong correlations between SAH levels in plasma and in lymphocytes, and that SAH levels were negatively

correlated with DNA methylation changes (62-64). Abnormal levels of homocysteine, SAH, or SAM indicate an imbalance in the pathway (65,66).

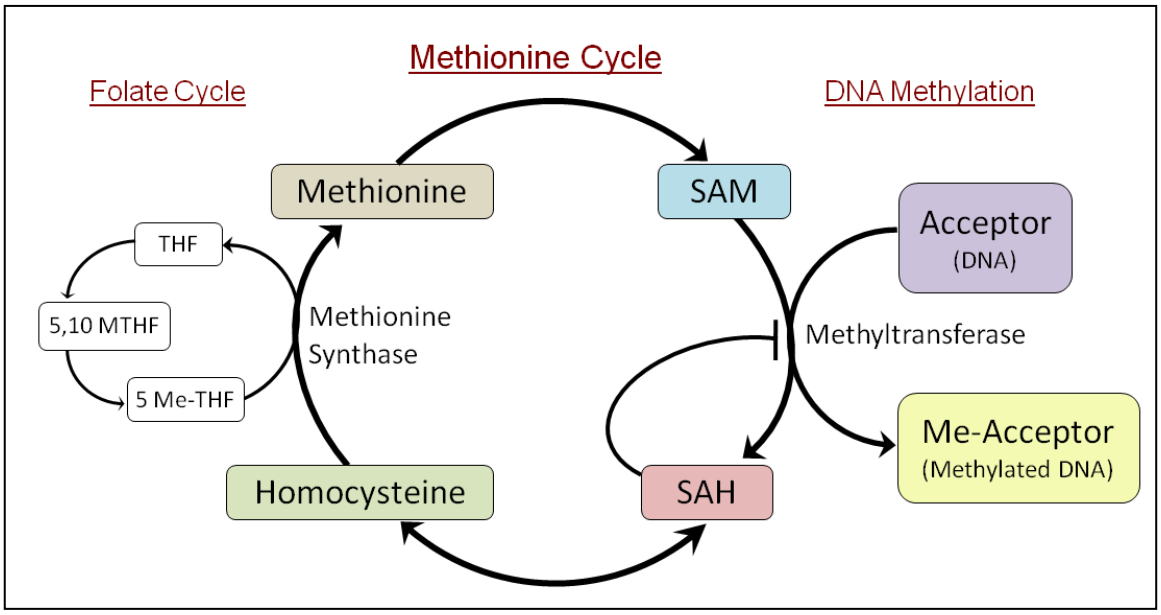

Figure 2-3. The One-Carbon Metabolism Cycle. The cycle is responsible for methylating compounds through the activity of methyltransferases that take a methyl group from SAM. Legend: S-Adenosyl-Methionine (SAM), S-Adenosyl-Homocysteine (SAM), Tetrahydrofolate (THF), Methylene-Tetrahydrofolate (MTHF), Methyl-Tetrahydrofolate (Me-THF).

The One-Carbon Metabolism Cycle is a complex cycle involving many additional cofactors. Some of these dietary factors include folate, Vitamin B12 and B6, choline, and methionine (67). Furthermore, genetic polymorphisms and expression levels of enzymes involved in the One-Carbon Metabolism Cycle may have an impact on the functionality of the cycle (63,67,68).

2.2.2 Prevalence of Methylation in the Human Genome

CpG dinucleotides are predominantly found in two locations in the human genome and have varying standard DNA methylation levels. Fifteen to twenty percent of CpG

dinucleotides are located in CpG islands that are located upstream, in the 5' region, of approximately half of human genes. These CpG islands normally have low levels of methylation (10,61,69), while the other eighty percent of CpG dinucleotides are located in repetitive sequences and are normally highly methylated (10,61).

2.2.2.1 Global DNA Methylation

Global DNA methylation refers to the methylation level across the entire genome (10). Globally, 70 to 90 percent of CpG sites are methylated in human DNA (10,70), although this varies based on tissue site. Methylation levels of some repetitive sequences, in particular LINE-1, have been shown experimentally to be highly correlated with genome-wide methylation levels; however, this is specific to certain regions of the LINE-1 sequence (12). Alu is a short repetitive sequence that is 350 base pairs in length. The Alu and Sat2 (human satellite 2 repeat) have also been shown to be correlated with global DNA methylation levels (12).

It is unrealistic to measure overall genome-wide methylation on a large number of samples because the methods are prohibitively time-consuming and expensive, as discussed in section 2.2.3.2. Thus, DNA methylation in specific repetitive elements is used as a proxy measure for genome-wide methylation. For this study, the methylation status of LINE-1 repetitive sequence was chosen to represent a surrogate measure of overall global methylation as suggested by the literature (12).

2.2.2.2 Gene-Specific DNA Methylation

Gene-specific DNA methylation refers to the methylation level of a particular region, usually CpG islands or promoter regions of a gene. It is currently estimated that fifty to sixty

percent of genes have CpG islands (70). Site-specific DNA methylation levels normally tend to be lower than global methylation levels and functionally control gene activity by suppressing gene transcription. Changes in site-specific DNA methylation are also of interest in disease etiology, in particular cancer (71,72), but are not the focus of this thesis.

2.2.3 Measurement of DNA Methylation

A variety of methods have been established to quantify DNA methylation levels. Each of these methods possesses strengths and weaknesses. This section will discuss briefly the most common methods and their applicability. Many novel approaches to measure DNA methylation are being researched. It is important that the validity of any novel method is investigated. Epidemiologic studies of DNA methylation require a measure that is not only valid, but also feasible to perform on a large number of subject samples and that uses a small amount of DNA.

2.2.3.1 Stability of DNA Methylation in Blood Samples

DNA methylation is relatively stable given that the methyl compound is covalently bonded to the cytosine nucleotide. This is a stable bond that is maintained during routine DNA extraction procedures (73). The stability of DNA methylation in a biologic sample is supported by the observation that there was no correlation between the amount of time a sample was stored in a freezer and methylation levels (16,74).

2.2.3.2 Methods to Measure Global DNA Methylation

The first methods outlined have the ability to directly measure overall global DNA methylation levels. High Performance Liquid Chromatography (HPLC) or Mass Spectrometry (MS) based methods rely on enzymatic hydrolysis of DNA, which is the

separation of all nucleotides to be analyzed (75,76). These methods require large amounts of DNA in order to measure the methylation levels. Methods, such as the methyl acceptor assay, which have been developed to quantify the amount of methylated cytosines, incorporate an enzyme into the assay which adds a methyl group to all unmethylated CpG sites (75).

Pyrosequencing and Polymerase Chain Reaction (PCR) based methods rely on the use of bisulfite conversion of the starting DNA and then PCR to amplify a segment of DNA. Bisulfite conversion replaces all the unmethylated cytosine nucleotides with uracil, while the methylated cytosines remain intact. Nucleotides pair in distinct combinations, cytosine pairs with guanine and adenine pairs with thymine. However, bisulfite conversion replaces unmethylated cytosines with uracil nucleotides, which then bind with adenine. Therefore the original unmethylated cytosines are replaced by thymines in subsequent replications. This provides DNA in a format where methylated and unmethylated cytosines can be distinguished from each other.

PCR-based methods measure methylation on a segment of DNA, typically 100-200 base pairs long, that is flanked by primers. Therefore these methods are unable to directly measure overall global DNA methylation levels. However, PCR-based methods are ideal for quantifying methylation levels at multiple sites or in repetitive sequences. Pyrosequencing is an expensive method used to determine the nucleotide order in the segment. MethyLight uses quantitative real-time PCR with probes and primers specific to methylated and unmethylated DNA (77). Another method, which will be discussed in further depth in the methods section, is Methylation-sensitive High-Resolution Melting (HRM). This method takes advantage of the different number of bonds which results in a strength difference between cytosine-guanine nucleotide pairing that has 3 hydrogen bonds, and adenosine-thymine nucleotide

pairing that has 2 hydrogen bonds. Quantitative PCR (qPCR) is used to amplify a sequence. Then the melt profile is used to differentiate along the spectrum from unmethylated to completely methylated strands (78,79). Stanzer et al. validated methylation levels using HRM with Pyrosequencing and MethyLight; the correlation coefficients were 0.943 and 0.949, respectively (80). This illustrates the utility of more cost-effective methods, such as HRM, for measuring DNA methylation levels.

Each method has its own set of potential limitations. When utilizing a novel method it is important to investigate the reliability and the validity of the method. For each method, sources of potential error should be addressed. These may include: starting quantity of DNA, number of CpG sites, PCR bias, assay efficiency, and reliability. For example, if using a Mass Spectrometry (MS) method it would be important to validate the starting quantity of DNA accurately and to ensure that a complete hydrolysis had occurred. PCR based methods have to be completed on multiple experimental runs, therefore the reliability between runs should be assessed. Similarly, the primers are typically designed based on a methylated DNA strand which has the potential to result in PCR bias, which should be assessed.

2.2.3.3 Locations

Repetitive elements are sequences of DNA that repeat across the human genome and do not encode genes. Interspersed repetitive elements are distributed widely throughout the genome. There are two main categories of interspersed repetitive DNA: long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINEs). SINEs are less than 500 base pairs in length. The most abundant SINE is the repetitive sequence Alu, which has over 1,000,000 copies and accounts for approximately 11 percent of the human genome (81,82). LINEs are sequences with more than 500 base pairs in length. LINE-1 has over

500,000 copies in the human genome and comprises approximately 17 percent of the human genome (83-85).

2.2.3.4 Tissue Specificity

Global DNA methylation levels in humans are tissue-specific (12,84,86). Firstly, the baseline levels in tissues may be different because methylation patterns are involved in the differentiation process to form organs during embryogenesis. Additionally, some tissues have a transsulfuration pathway to assist in the removal of homocysteine from cells, while some other tissues do not (62). Lymphocytes do not have the transsulfuration pathway and therefore may be associated with higher homocysteine levels and an increased sensitivity to SAH levels.

Cell types within a specific tissue may also have different methylation patterns. Leukocytes, commonly known as white blood cells, are comprised of five different white blood cell types: neutrophils, eosinophils, basophils, lymphocytes, and monocytes. All of these cell types originate in the bone marrow and may have different methylation patterns (74).

2.2.3.5 DNA Methylation in Leukocytes as a Biomarker

Research on whether DNA methylation in leukocytes is a useful biomarker for different health outcomes is limited. For a biomarker of an intermediate endpoint to be meaningful in understanding disease causation, the biomarker has to be correlated with the endpoint of interest and exposures should be related to the marker. Although research is currently limited, there is evidence to support these criteria for DNA methylation measured in leukocytes. Epidemiologic studies have reported associations between global DNA

methylation in leukocytes and several different cancers including cancers of the colon, bladder, stomach, breast, and head and neck (87-92). Epidemiologic studies have reported relationships between demographic factors (e.g. age, gender, and race), environmental exposures (e.g. benzene and air pollution), and other risk factors (e.g. cigarette smoking, alcohol consumption, and diet) and global DNA methylation in leukocytes (93,94). Furthermore, it has been suggested that there may be a systemic change in methylation levels, which may predispose individuals to adverse outcomes (89,95)

2.2.4 Levels of DNA Methylation

To be useful as an endpoint in an epidemiologic study a biomarker must have meaningful variability in the study population. There is considerable variation in DNA methylation levels in humans. In repetitive sequences the variability between healthy individuals is between 5 to 25 percent in blood (96). Strong evidence supports the ability of environmental or lifestyle factors to modify methylation levels both globally and at specific gene sites. Early studies from monozygotic twins depict diverging methylation patterns with age and with dissimilar exposures (93,97). Terminology has been coined to refer to relative states of methylation levels.

2.2.4.1 Hypomethylation and Hypermethylation

When DNA methylation levels of test samples are greater than average levels in healthy specimens the region is considered to hypermethylated. Conversely, when DNA methylation levels of test samples are lower than average levels in healthy specimens the region is considered to be hypomethylated. These are both relative states on a continuum of methylation level (10).

2.2.4.2 Stability of DNA Methylation Levels in Individuals

For DNA methylation to be a useful measure in epidemiologic studies it is important for methylation levels to be fairly stable over short periods of time. Kok et al. examined fasting blood samples from ten individuals longitudinally over time intervals of one week and ten months. No differences between the methylation measurements were observed (98). This provides support that DNA methylation levels are relatively stable in an individual over at least ten months.

2.2.4.3 Potential Mechanisms leading to Hypomethylation

Research is expanding to elucidate the mechanisms involved in regulating and altering methylation status. The primary cause of hypomethylation is related to the disruption in the One-Carbon Metabolism Cycle, which has been shown to influence the activity of DNMTs. Also, the failure of maintenance DNA methyltransferase to maintain methyl levels during replication would result in lower methylation levels (99). It has been suggested that DNA may be actively demethylated, however, the mechanisms and enzymes involved are still unknown (100). Lastly, if a methylated cytosine undergoes a deamination reaction it becomes a thymine nucleotide (99) and therefore would not be recognized by DNA repair enzymes as abnormal and consequently would not be replaced.

2.2.5 Biologic Function of DNA Methylation in Cells

DNA methylation is critical during embryogenesis. In mouse models, knock-out of the DNMTs is embryonic lethal, indicating that the establishment and maintenance of methylation patterns are crucial for development (61). DNA methylation plays a vital role in the inactivation of an X-chromosome in females and gene imprinting (101). A number of congenital abnormalities and malignancies are associated with abnormal methylation patterns

or mutations in the DNMTs (102). This further illustrates the importance of methylation levels in healthy development.

The activity level of a gene is regulated by its levels of transcription from the genome. DNA methylation in a promoter region of a gene suppresses its transcription. Consequently, both hypomethylation and hypermethylation at specific sites (genes) and global hypomethylation can change the activity level of genes. Generally, hypomethylation would allow for greater activity while hypermethylation would down-regulate genes (93,102).

From cell line studies it has been shown that global DNA methylation functions to assist in genomic stability. Stabilizing the genome is done by suppressing recombination and preventing mutations. It has been suggested that hypomethylated DNA causes an increased frequency of mutational events (103).

Finally, DNA methylation is critical in coordinating the structure of chromatin. Euchromatin refers to an expanded or open chromatin structure that permits access of transcription factors to genes. Euchromatin is characterized as DNA being loosely bound to histones and DNA being hypomethylated. Conversely, when DNA is condensed or in a heterochromatin form, it is tightly woven around histones and the DNA is highly methylated. There are proteins called Methyl-Binding Domain proteins which specifically bind to methylated cytosine and act to suppress transcription (104-106). Thus, not only does methylation suppress transcription, but there is an interaction between methylation levels and chromatin structure that functions to further control gene expression and genomic stability.

2.2.6 Determinants of Global DNA Methylation

Information regarding the determinants of global DNA methylation levels in leukocytes is limited. This section will first outline the evidence for the determinants of leukocyte DNA methylation levels. Then the known and potential determinants of DNA methylation levels based on factors affecting the One-Carbon Metabolism Cycle will be discussed.

Few studies in the literature have examined the lifestyle determinants of leukocyte global DNA methylation levels. A number of non-modifiable factors (e.g. age and sex) have been shown to influence methylation levels. Although inconsistent findings have been reported, there is evidence to support that with increasing age, genome-wide methylation levels decrease (93,107,108). The effect of age on methylation levels tends to be more prominent in the Alu sequence rather than the LINE-1 repetitive sequence (107). Further support for an inverse relationship between age and methylation levels comes from two studies which measured methylation levels in blood samples collected at two separate times (107,109). One study found that one-third of their study population showed methylation level changes greater than 10 percent between samples collected 11 years a part (109).

Methylation levels in repetitive elements vary by sex. The majority of studies have found that males have higher overall global and LINE-1 DNA methylation levels compared to females (16,90,94,96,110). However, the type of repetitive sequence appears to influence the relationship between sex and DNA methylation levels. Inconsistent findings exist for the relationship between sex and the Alu repetitive elements (111), with studies reporting that females had higher Alu methylation levels compared to males (94,112). Lastly, studies have

reported inconsistent findings for the relationship between ethnicity and DNA methylation level (93,113).

Modifiable factors in relation to methylation levels are currently being researched. Smoking has been investigated in relation to methylation levels in several studies (16,93,94,112,114). Overall, the evidence does not support an association between smoking and methylation levels, either globally or in specific repetitive sequences. Mixed results exist for the relationship between alcohol and methylation levels. Lower Alu methylation levels (but not LINE-1) have been associated with increased alcohol consumption (94,110,112). Conversely, a case-control study that compared methylation levels in individuals diagnosed with alcoholism with controls found higher (8 to 10 percent) global methylation levels in those with alcoholism (115).

Associations have also been reported between DNA methylation levels and environmental exposures to chemicals and metals. Chemical exposure to benzene, air particulates (i.e., sulfate and black carbon), and medications (i.e., cytidine analogs) have been associated with decreased DNA methylation levels (93,116). Exposure to certain metals has been identified to affect methylation levels. Nickel and arsenic have been associated with hypermethylation and lead has been associated with hypomethylation (93,116).

It is important to note that one of the major studies that examined potential determinants of DNA methylation had a low amount of variability in methylation levels. The methods used to measure methylation levels and the different genome regions or sequences measured may result in different associations being observed. Overall, the determinants of DNA methylation levels are still largely unknown and may differ by location and tissue.

Further research is required to adequately address epidemiologic risk factors for global and repetitive sequence DNA methylation levels.

The One-Carbon Metabolism Cycle is the molecular cycle that regulates methylation capacity. It has been suggested that increased SAM levels may protect DNA from demethylation (117). However, in healthy individuals, it has been found that homocysteine has been inversely correlated with methylation levels (62). Dietary factors, especially those involved in the One-carbon Metabolism Cycle, such as folate, choline, and vitamins B12 and B6, may influence methylation levels. Most observational studies have not seen a relationship between dietary folate and leukocyte DNA methylation levels (93,98). However, in a randomized controlled trial participants randomized to folate supplementation had an increase in DNA methylation levels in leukocytes compared to the placebo group (118). This suggests that dietary factors may play a critical role in maintaining normal methylation levels.

Furthermore, other dietary factors that affect the One-Carbon Metabolism Cycle may influence DNA methylation levels. These additional factors include zinc, selenium, and methionine (10,62,119). Additionally, polymorphisms in the enzymes involved in the One-Carbon Metabolism Cycle have been associated with DNA hypomethylation in a few studies (67,68) but not others (93,98,120). *In vitro*, polyphenols and soya compounds were shown to inhibit DNMT activity in a dose-dependent manner, however the amount of these compounds required was greater than the nutritionally consumed dose (7).

2.2.7 Outcomes associated with Global DNA Methylation Levels

Global DNA methylation is recognized as an important factor in an array of adverse health outcomes. A potential mechanism for the increased risk of certain diseases with age

may be related to decreasing lifetime methylation levels. Specifically, aberrant global DNA methylation has been suggested to play an etiologic role for the two most common chronic diseases: cancer and CVD (10,121-124).

2.2.7.1 Cancer

Global and LINE-1 DNA hypomethylation are common characteristics of human cancers (91). Evidence has supported the role of DNA methylation in the development of a wide variety of cancers such as colon, cervical, and prostate cancer, and also leukemia (10,84,125,126). Consistently, global DNA hypomethylation is recognized as an early event in neoplastic development (10,124). Methylation levels are currently being examined as potential biomarkers for cancer risk, detection, and prognosis (127).

The strongest evidence exists for global DNA hypomethylation as a risk factor for the transformation of normal to neoplastic colon tissue. This section will briefly discuss some of the experimental and epidemiological evidence supporting a relationship between global hypomethylation and colon cancer. Global DNA hypomethylation is a known characteristic of colon tumor cells (128) and has been observed in normal-appearing adjacent colon tissue from individuals with colon cancer (87,118,129). Leukocyte DNA from subjects with colorectal adenomas was less methylated as compared to subjects with a normal colonoscopy (87). There exists a strong association between global DNA hypomethylation and chromosomal instability in colon cancer (126) suggesting that this may be a component of the pathway through which decreased global methylation would lead to carcinogenesis. Lastly, common polymorphisms in enzymes of the One-Carbon Metabolism Cycle have been shown to influence the risk of colorectal neoplasia (87,130). Overall, evidence supports global DNA hypomethylation as an early event in colon cancer development.

2.2.7.2 Cardiovascular Disease (CVD)

The One-carbon Metabolism Cycle is also relevant to CVD. Elevated plasma homocysteine is an established risk factor for CVD (6). DNA methylation is becoming recognized as an important pathway for CVD. However, the area is new and many questions remain. Atherosclerotic lesions have been characterized as having evidence of global DNA hypomethylation (131-133). However, it has been suggested that hypomethylation in the aortic plaque may be a consequence of rapidly multiplying smooth muscle cells that are unable to maintain methylation patterns as opposed to contributing to the etiology of the plaques themselves (131). Interestingly, cell lines stimulated with lipids showed global DNA hypermethylation, suggesting that DNA hypermethylation may be an early event in atherogenesis (131,132).

Kim et al. examined the association between DNA methylation status and the prevalence of CVD in the Singapore Chinese Health Study and found that elevated DNA methylation levels in leukocytes were associated with an increased prevalence of CVD (myocardial infarction and stroke) and CVD predisposing conditions (hypertension and diabetes) in males (111). Additional support originates from a study which compared methylation levels of individuals with chronic kidney disease to healthy controls. Global DNA hypermethylation was found to be present in patients with inflammation and those with CVD compared to healthy controls and DNA hypermethylation was associated with overall CVD mortality (16). Another study found that leukocyte LINE-1 DNA hypomethylation was associated with ischemic heart disease and stroke (74).

The literature presents inconsistent findings; global hypomethylation was suspected to be associated with atherosclerosis (133) but more recent evidence suggests a relationship

between global hypermethylation and CVD. Further research is required to disentangle the role of DNA methylation levels in CVD etiology and mortality.

2.3 Potential Biologic Pathways

The regulation of methylation patterns and the mechanisms through which environmental exposures impact on DNA methylation levels are still being uncovered. This section will outline two potential biologic mechanisms through which obesity may have an impact on methylation levels: oxidative stress and inflammation depicted in Figure 2-4.

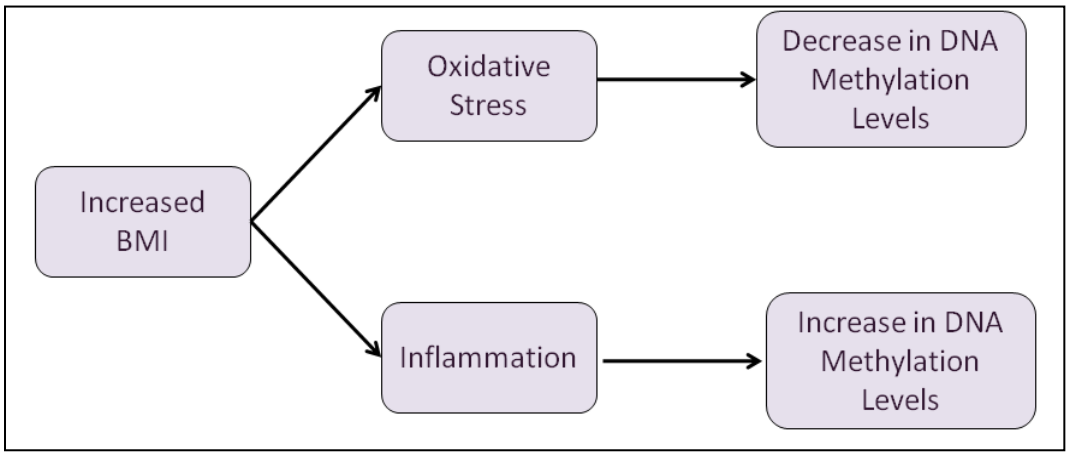

Figure 2-4. Potential Mechanisms between BMI and Methylation Levels. This figure illustrates that increased BMI may have an impact on DNA methylation levels through oxidative stress and/or inflammation.

Homocysteine is associated with obesity and the adverse health outcomes. Primarily, it is suggested that homocysteine causes DNA damage through oxidative stress. But methylation changes, specifically global DNA hypomethylation, may also play a critical role in homocysteine-related pathogenesis. This indicates a potential relationship between oxidative stress and DNA methylation changes.

2.3.1 Oxidative Stress

Oxidative stress refers to an imbalance of reactive oxygen species, a type of free radical, which can act to cause DNA damage. Oxidative stress is considered to contribute to the occurrence of carcinogenesis through genotoxic and epigenetic mechanisms (134). Experimentally, it has been shown that when a hydroxyl group was added to guanine nucleotide, adjacent DNA methylation was inhibited (135). Oxidative stress disrupts the binding of DNMTs to DNA, which would result in global DNA hypomethylation (17). Lastly, when methylated cytosines react with reactive oxygen species, the cytosines transition into unstable compounds that would deaminate or transition into a thymine or uracil nucleotide instead (136). Consequently, these mechanisms would result in a decrease in the presence of methylated cytosines (i.e.; hypomethylation).

2.3.2 Inflammation

Chronic low-grade inflammation is a characteristic of those who are obese. Evidence supports the role of inflammation in the etiology of chronic diseases. A few mechanisms have been suggested. Chronic inflammation has been associated with increased methylation at gene-specific locations (116,137). *In vitro*, cells exposed to the inflammatory cytokine, IL-6, exhibited significant increases in global DNA methylation levels (16). Similarly, *in vitro*, it appears that IL-6 may have an impact on the activity of DNMTs (138).

Inflammation has been shown to result in HOCl and perhaps HOBr production internally; these compounds have been detected in leukocytes and are known to react with DNA to form halogenated cytosines (139,140). *In vitro* studies have demonstrated that DNMTs were unable to discern halogenated cytosines from methylated cytosines and bound

to both with equal affinity (15,141). Thus the impact of inflammation and halogenated cytosine nucleotides may result in increased DNA methylation levels (15).

2.4 Literature Review of BMI and Methylation Levels

This section will examine the existing literature regarding the primary objective of this study, that is, to examine the relationship between BMI and DNA methylation levels.

2.4.1 BMI and the One-Carbon Metabolism Cycle

Studies have reported an association between BMI and components of the One-Carbon Metabolism Cycle, including homocysteine, SAM, and SAH. However, these studies have been inconsistent ranging from no association to BMI being a strong determinant of SAM and SAH (142-144). However, the majority of studies have found relationships between BMI and the components of the cycle.

2.4.2 BMI and Global DNA Methylation

Only a few studies in the literature have examined the relationship between BMI and global or repetitive element DNA methylation levels reporting mixed results. One study found that increasing BMI was associated with lower LINE-1 DNA methylation levels in healthy women (145), while a few studies have not found a relationship between BMI and methylation levels (16,94,112,146). However, Kim et al. observed that BMI was positively associated with methylation levels in Alu and Sat2 (111). The existing literature is inconsistent when examining the relationship between BMI and methylation. Furthermore, many of these studies were conducted in a study population that was burdened with disease and therefore it is difficult to determine the influence of BMI on methylation in healthy individuals.

2.4.3 Reverse Causality

Since DNA methylation plays a vital role in gene regulation, it has been suggested that methylation levels have the potential to regulate body weight. This research has mainly focused on gene-specific and not global DNA methylation levels. Genome-wide DNA methylation studies found an association between four methylation regions and BMI (147). Similarly, two birth cohorts supported an epigenetic component for metabolic disease based on the examination of umbilical tissue gene-specific methylation levels and found a few candidate genes that were significantly associated with percentage of body fat at 9 years of age (148). It appears that some genes may play a causal role in metabolic development and therefore BMI. However, the direction of causation and the impact of environmental exposures are still unclear. For example, a study of monozygotic twins with discordant BMIs did not show an association with methylation levels (149).

2.4.4 Study Hypotheses

The laboratory-based scientific literature provides evidence for two possible biologic mechanisms for BMI to influence DNA methylation levels: oxidative stress and inflammation. However, the proposed biologic mechanisms would result in different directions in the effect of elevated BMI on methylation levels. Specifically oxidative stress would result in decreased methylation levels and inflammation would result in increased methylation levels. Similarly, the epidemiologic literature has demonstrated inconsistent findings for the relationship between BMI and DNA methylation. Different epidemiologic studies have shown a positive relationship, a negative relationship, and no association between BMI and measures of global DNA methylation. Consequently, there is uncertainty surrounding the expected direction of the relationship between BMI and DNA methylation

levels and it is plausible for BMI to result in hypomethylation or hypermethylation. Therefore, it would not be appropriate to constrain this thesis's hypothesis to only one possible direction of association. The hypothesis is that BMI could result in either increases or decreases in DNA methylation levels.

Chapter 3

Study Design and Methods

3.1 Study Objectives

This thesis research had three objectives.

Objective 1: To quantify and describe LINE-1 DNA methylation in leukocytes in a large sample of healthy volunteers.

Objective 2: To examine the relationship between Body Mass Index (BMI) and LINE-1 DNA methylation levels in the study population.

Objective 3: To determine whether sex was an effect modifier of the relationship between BMI and LINE-1 DNA methylation levels.

3.2 Study Design

This cross-sectional study was nested within a larger study funded by the Canadian Institutes of Health Research (CIHR) that examined the relationship between environmental and lifestyle factors, and components of the One-Carbon Metabolism Cycle (Principal Investigator W. King). The larger study carried out recruitment and data collection on 663 subjects from 2007 to 2009. The subjects accrued were healthy volunteers from three cities: Kingston, Halifax, and Ottawa. Subjects provided a fasting blood sample (whole blood, plasma, and serum), completed a short questionnaire, and returned a home water sampling kit. Whole blood, serum, and plasma were stored at -80ºC prior to analysis. The analysis of the larger study focused on water disinfection byproducts and measures of the One-Carbon

Metabolism Cycle, which included biomarkers of DNA methylation capacity (S-Adenosyl-Methionine, S-Adenosyl-Homocysteine, and Homocysteine).

3.3 Source Population and Subjects

The target population was male and female subjects, aged 20 to 50, recruited in approximately equal numbers within 10-year age intervals. Subjects with health conditions that might be related to the outcome measures were excluded (outlined Section 3.3.1). Beyond these considerations, a representative population sample was not a priority because the study objectives were oriented towards understanding biologic mechanisms that are postulated to be consistent irrespective of population dynamics. Given the nature of the underlying relationships of interest, participation bias was unlikely and therefore response rates are a secondary consideration after recruitment of a sufficient sample.

Municipalities served by a public water treatment and distribution system or using a surface water source were selected to provide variability in disinfection by-product exposures. A large hospital with laboratory facilities was identified in each study area. Participants were recruited within the hospital adjacent to medical, research, and educational institutions by a study-coordinator employed at each centre. Advertisements on public bulletin boards within the hospital and adjacent medical and research institutes described the study and asked potential participants to contact the study coordinator. Those responding were contacted and an appointment was made to provide a blood sample at the hospital. The blood draw took place in the morning after a twelve-hour fast. The study coordinator met with the subject at the time of the blood draw in order to administer the questionnaire, measure blood pressure, and to give them the water sample kit. Subjects were contacted 3

days later to arrange for courier pickup of the water sample kit from their home. Compensation ($30) was given to participants who completed the study.

The Kingston centre was located in the Department of Community Health and Epidemiology. Subjects were recruited within adjacent institutions (Queen's University, Kingston General Hospital, Hotel Dieu Hospital, St. Mary's Hospital, Kingston Regional Cancer Centre) comprising an employment and student population of over 27,000. The Ottawa study centre was located at the Civic Campus of the Ottawa Hospital and the study coordinator recruited from the 3 campuses of the Ottawa Hospital (General, Civic, and Riverside) and adjacent medical institutions (Heart Institute, Cancer Centre) with a medical, research, and staff workforce of over 20,000. The Halifax study centre was located at the IWK Health Centre. Study participants were recruited from this and adjacent medical and educational institutions (Dalhousie University, the IWK Health Centre, and the QEII Health Sciences Centre). These Halifax institutions have a student/workforce of over 20,000.

3.3.1 Eligibility and Exclusion Criteria

For the larger study, potential subjects were screened for current or past history of medical conditions that may be related to the outcome of interest (DNA methylation), which included cancer, vascular disease, diabetes, or a pregnancy within the last year. These individuals were not eligible for the larger study sample. The age range from 20 to 50 years was appropriate for this study based on the fact that the chronic diseases of interest tend to develop after a long latency period. Since DNA methylation is considered to be an early step in the carcinogenic pathway for certain cancers and potentially an early event in CVD, the age range for study participants was meant to represent a meaningful time window for the early biologic changes.

Strategically for this thesis, subjects were excluded to maximize the validity of DNA methylation and BMI measurements. Subjects who did not have a fasting blood sample or biochemical analyses completed were excluded. Subjects were excluded from the study if their measurement of LINE-1 DNA methylation was considered unreliable, which is explained in Section 3.5.2.1. Also, all subjects who did not have complete BMI information (i.e. height and weight), or had underweight BMI values were excluded from this study because of their small numbers.

3.4 Data Collection

Participants were recruited from nearby hospitals and educational institutions detailed in Section 3.3. For the larger study, participants completed a self-administered questionnaire at a university or hospital located study centre. The study recruitment aimed to capture a similar population demographic from all three cities. The study questionnaire collected information regarding demographic and lifestyle characteristics. The relevant items will be detailed in the covariate section.

3.4.1 Exposure Measurement (BMI)

Height and weight were self-reported in either Metric or Imperial units. If a subject was unsure of his or her current weight or height, then an individual's weight or height was measured by the study research associate responsible for administering the questionnaire. Body Mass Index was calculated using the Quetelet's Index, where BMI is equal to weight in kilograms divided by height in metres squared (19). BMI was considered as a categorical variable. For the classification of BMI into categories, three divisions were used according to

the standard classification used by the World Health Organization: Normal (18.5 - < 25.0 kg/m^2), Overweight (25.0 - < 30.0kg/m^2), and Obese (≥ 30.00kg/m^2) (1).

3.4.2 Outcome Measurement (DNA Methylation)

3.4.2.1 Blood Sample Procurement

Participants completed a twelve-hour overnight fast prior to having their blood drawn from their median antecubital vein. The samples were immediately placed on ice and then centrifuged at 3300 rpm for 10 minutes to separate the plasma. The plasma and serum were aliquoted into vials and stored at -80ºC until analyzed. For the biochemical analyses, aliquots were sent to the Department of Pathology at Queen's University and to the Laboratory Medicine Centre at the Ottawa General Hospital. For the purpose of this study, aliquots of whole blood were sent to the Department of Pathology and stored at -80ºC.

3.4.2.2 Laboratory Procedures and Quality Control

This section will outline the steps involved in the DNA methylation measurement, starting with a blood sample and ending with a percent methylation level. Fundamental to the first study objective is the need to assess and document the validity and reliability of the measure of DNA methylation in the context of a novel biomarker with no gold standard for comparison. An overview of the procedures is provided in section 3.4.2.3 while a complete description is provided in sections 3.4.2.3.1 and 3.4.2.3.2.

3.4.2.3 Overview of DNA Methylation Measurement

DNA from white blood cells was isolated from a small sample of whole blood. DNA was bisulfite converted which converts the DNA to a form which permits the differentiation between methylated and unmethylated cytosines. DNA methylation was measured at 8 CpG

sites within the LINE-1 sequence which is a repetitive sequence that has been shown to be representative of global DNA methylation levels. For each qPCR experiment, five control standards with known levels of methylation (0%, 25%, 50%, 75%, and 100% methylated DNA) were included in order to generate a standard curve. The melting temperature and profile of samples were compared to the standard curves to determine the percent LINE-1 DNA methylation. Each sample was completed in triplicate in order to have more precision around the mean value, to permit the exclusion if a replicate is different from the other two, and to assess the reliability of the measurement.

3.4.2.3.1 DNA Isolation and Preparations

For this thesis the starting material available to the student investigator was an aliquot of 250µL of whole blood for each subject. Genomic DNA from leukocytes (white blood cells) was extracted from the blood sample using the 5-Prime ArchrivePure DNA Isolation Kit (Inter Medico, Markham, Ontario, Canada). The DNA isolation protocol used is provided in Appendix A. The concentration and purity of each DNA sample was measured with the NanoDrop 2000 UV-Vis Spectrophotometer (ThermoScientific, Wilmington, Delaware, United States of America). In preparation for the bisulfite conversion step, an aliquot of 2µg of DNA was made with a final concentration of 100µg/µL. If the DNA was at lower initial concentration after isolation, then the sample was concentrated and re-diluted to obtain a final concentration of 100µg/µL. This additional step was only necessary for a small number of samples. The DNA was stored at -20ºC during the processing and prior to bisulfite conversion.

A 2µg aliquot of DNA was used for the bisulfite conversion step. This step was conducted on each subject sample as well as standard human methylated and non-methylated

DNA standards purchased from Zymo Research. The Epitect Bisulfite Conversion Kit was used to convert the 2µg of genomic DNA. The protocol for the bisulfite conversion process is located in Appendix B. This process converts all of the unmethylated cytosine nucleotides to uracil nucleotides and all methylated cytosine nucleotides remain intact. Variation in the composition of the base pairing permits the detection of different methylation states. The bisulfite converted DNA was stored at -20ºC. All samples were thawed once and quantified on the NanoDrop 2000. Then, from each converted DNA sample, two aliquots of 8µL each at a concentration of 1ng/µL were prepared. To avoid multiple freeze-thaw cycles, the aliquots were made within twenty-four hours of DNA quantification. All stock and aliquots were stored at -20ºC until DNA methylation measurement.

3.4.2.3.2 DNA Methylation Measurement Procedures

The purchased non-methylated standard (i.e. 0% methylated) and the methylated standard (i.e. 100% methylated) were combined in different proportions to generate standard samples with known methylation levels. For every experiment a standard curve was generated using these standards with methylation levels of 0%, 25%, 50%, 75%, and 100%. In addition, on every experimental run an internal control was included, which permits the inter-assay variation to be assessed. This internal control was from a one-time blood draw from a volunteer associated with the study research group.

High-resolution melt (HRM) analysis is a PCR-based method that requires a quantitative real-time PCR (qPCR) equipment. For this project the Roche LightCycler 480 (Roche Applied Science, Laval, Quebec, Canada) was used. Primers specific to the LINE-1 region, shown to be representative of global DNA methylation levels (12), were designed by Dr. M. Yat Tse. Development of the assay is described in M. Y. Tse et al. (150). This thesis

will highlight some of the important aspects of the assay. The primers contained three methylation sites and the amplicon was 141 base pairs long and contained eight methylation sites. For each experiment, a 96-well plate designed for the Roche LightCycler 480 was used. Each reaction well contained a total reaction volume of 12µL, including: 2µL of 1ng/µL of bisulfite converted DNA template, 2µL of $MgCl_2$, 2µL of primers, and 6µL of the LightCycler 480 High Resolution Master Mix. Therefore, each well had 2ng of DNA template to amplify, 0.2µM of each primer, 3.0mM of $MgCl_2$, and High Resolution Master mix. The only exception were the three no template control (NTC) wells which did not contain any DNA template. These wells served as negative controls for each experiment.

A fundamental premise of quantitative PCR (qPCR) is that an enzyme in the master mix fluoresces when intercalated with double stranded DNA. The application of this method permits the quantification of the amount of DNA. The qPCR reaction protocol was optimized as described in Tse et al., and described briefly in section 3.5.2.1, to the following conditions: 10 minutes at 95ºC to activate the Taq polymerase (the enzyme responsible to replicate the DNA), then 40 cycles of amplification, which was followed by a high resolution melt (HRM). The amplification cycle had 3 steps, beginning with a hold at 95ºC for 10 seconds to denature the DNA, followed by 48ºC for 10 seconds for the primers to anneal to the DNA sequences, and finally 72ºC for 15 seconds for the extension or replication of the template. During every cycle the level of fluorescence was acquisitioned. After 40 cycles were completed all samples should have reached their plateau. At this point an HRM is conducted, which consists of a run protocol of 95ºC for 1 minute in order to denature all the DNA and then 40ºC for 1 minute in order for annealing to occur, and then slow warming from 55ºC to

95ºC at a rate of 0.02ºC/second with the fluorescence level being measured 25 times every second (150).

For each set of triplicates multiple steps are involved in the analysis. First the crossing point (Cp) value, which is the number of cycles it has taken for a sample to reach a specified threshold value or quantity, is examined. Pre-determined cutoffs were made by an examination of a dilution assay to examine the limits of the assay. This is explained in section 3.5.2.1. Then the melt profiles of each triplicate were examined. The shape of the melt profile and the peak temperature for the melting curve provided information about the methylation level. Figure 3-1 depicts the melt profiles of the Zymo Research methylation standards in triplicates. Both melt profiles have a Gaussian distribution which demonstrates that multiple products are amplified.

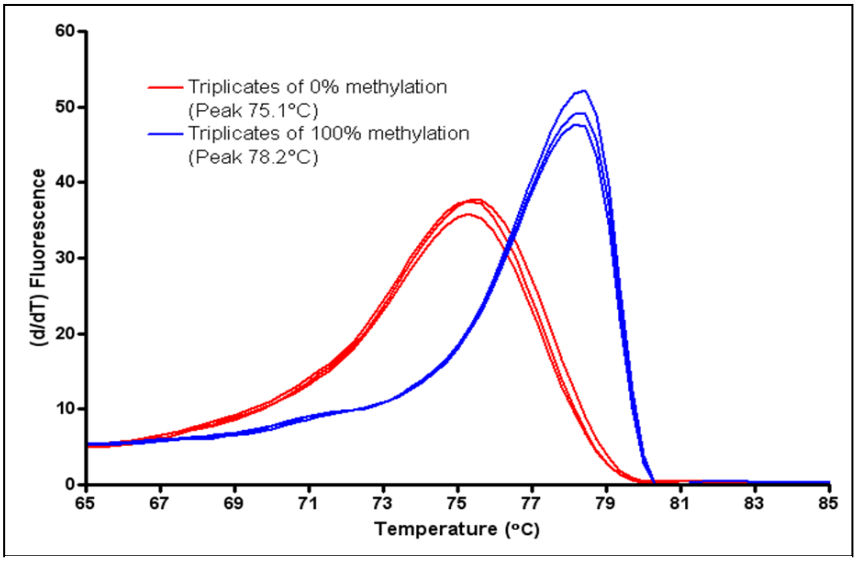

Figure 3-1. Melt Profiles for Methylation Standards. The 0% methylated DNA (in red) and 100% methylated DNA (in blue) standards melt at different temperatures.

The narrower shape of the melt profile for the 100% methylation standard compared to the 0% would indicate a smaller amount of variability in the methylation level. The distinct difference in the melt temperatures between the standards is demonstrated in Figure 3-1. The 0% methylation standard has a peak melting temperature over 3ºC lower than the peak melt temperature of the 100% methylation standard. The Gene Scanning software for the Roche LightCycler 480, designed for SNP (Single Nucleotide Polymorphisms), was adapted for the methylation analysis by Tse et al. (150). This process scales the fluorescence melt curves for intensity and a temperature shift is used to align the normalized melt curves. A difference curve is generated using the 0% methylation standard as the reference; this is shown in Figure 3-2. Data points for the area under the difference curve for all samples are exported into Microsoft Word Excel 2007 where they are summed to calculate the area under the difference curve.

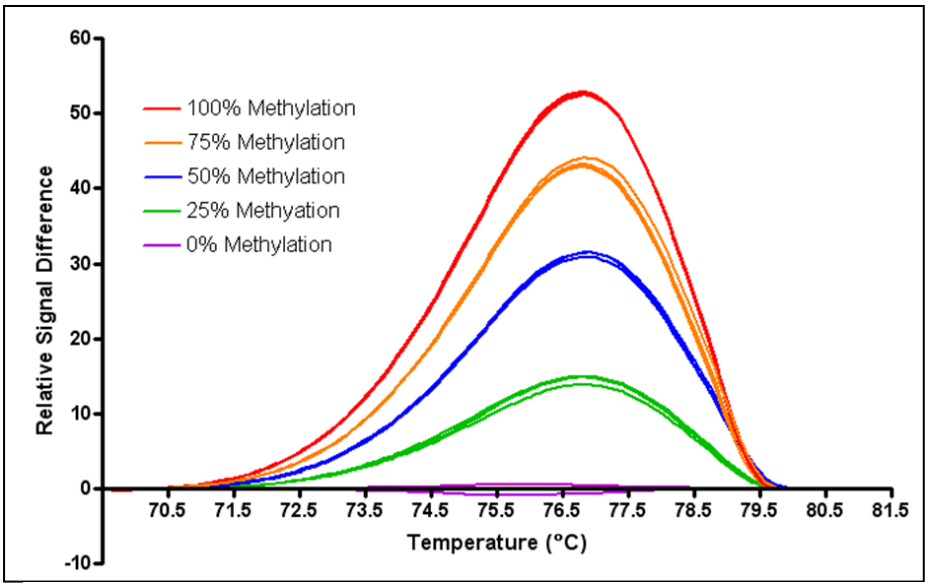

Figure 3-2. Difference Curve for Standard DNA Methylation Levels. Normalized melt curves are presented for triplicates of the each standard DNA methylation level.

Having obtained the values for the area under the curve for the methylation levels of the standards on each plate, a standard curve using linear regression was generated for each plate. An example of the standard curve is shown in Figure 3-3. This permits the quantification of the percent methylation values for the participant samples on the plate by interpolating the methylation levels from the area under the difference curve. This process results in triplicate measures of DNA methylation levels for each participant based on the high-resolution melt curve. For the statistical analysis, the average of the triplicates included for each participant was used.

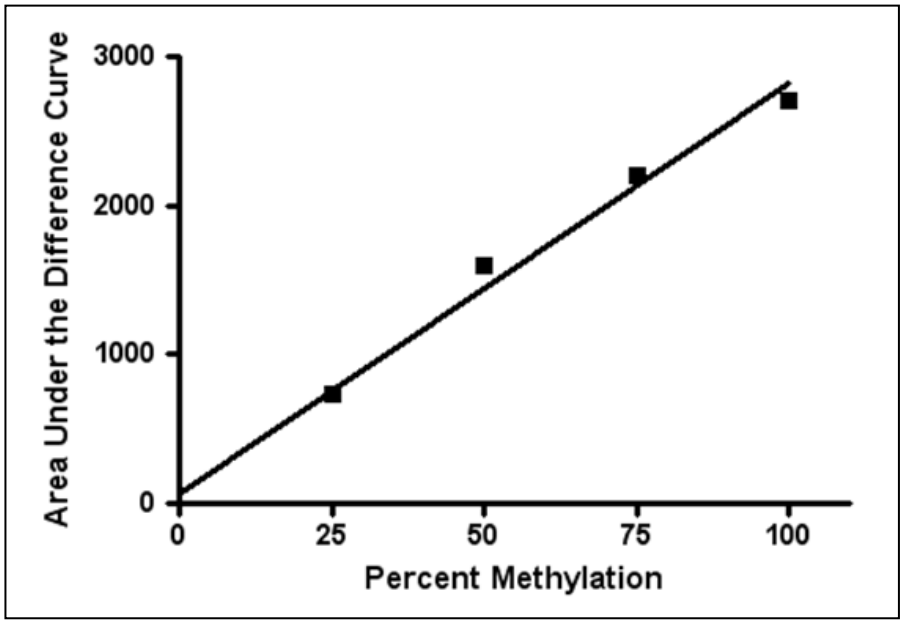

Figure 3-3. Standard Curve of Percent Methylation Levels.

Since the LINE-1 DNA methylation measurements were conducted in triplicates, this allowed for the potential exclusion of a replicate for various reasons. First, if there was interference in the fluorescence measurement of the melt profile then the single replicate

measure was excluded. Visually, the melt profile should appear as a smooth Gaussian distribution, while a disruption in the fluorescence appears as an inappropriate bump or shoulder in the profile. Every sample was checked for interference and excluded if this was evident. Second, if there was a clear shift in one of the peak melt temperatures compared to the other two replicates then the measure was excluded. Lastly, if the sample failed to amplify correctly, it was excluded. If more than one replicate had a technical reason to be excluded then it was planned to re-run the sample. In total, less than 40 individual triplicate measures were excluded for one of the above reasons out of over 1500 samples. No samples had more than one triplicate excluded.

Methylation measurements, the Cp values, and the plate number on which the sample was run were added to the database of the larger study. A statistical rule was applied to triplicates, which was if one of the triplicates differed by 5 percent or more from the other two replicates than it was excluded (150). Only for 9 samples out of 517 was one of the triplicates removed by this rule.

3.4.3 Covariate Measurements

Based on the literature review, covariates were selected to be considered as potential confounders. These included: age, sex, ethnicity, physical activity, smoking status, alcohol consumption, serum folate, and lipid levels. This section will outline how each covariate was measured and how it will be considered in the analysis. Because of the lack of literature in relation to DNA methylation levels the categories for each covariate were based on relationships with relevant health events observed in the literature.

3.4.3.1 Description of Potential Confounders

Age: Participants self-reported their age on the study questionnaire. To be eligible for this study subjects had to be between 20 and 50 years of age. Sampling fractions for the larger study were based on increments of ten years. Thus for this study, age was categorized into three divisions: 20 to 29 years of age, 30 to 39 years of age, and 40 to 49 years of age.

Sex: Information for a participant's sex was self-reported in the study questionnaire. The larger study aimed to obtain an approximately equal number of males and females in the study population.

Ethnicity: On the study questionnaire participants were asked to report the ethnicity of each parent. It was expected that the majority of participants would be of Caucasian decent. Ethnicity was divided into 3 categories: Caucasian, Asian, and other. If a parent was reported to be non-Caucasian then they were categorized according to the minority parent.

Physical Activity: To measure an individual's physical activity level the short International Physical Activity Questionnaire (IPAQ) was included in the study questionnaire. The guidelines for the IPAQ questionnaire were used to generate an overall Metabolic Equivalent (MET) physical activity score per week (151). One MET is the amount of energy that an individual would burn sitting each minute. For example, if an individual walked 40 minutes a day, seven days a week, he or she would be assigned 1,000 METs. Using the MET score, quartiles were generated to represent physical activity in the analysis. A distribution based method was used to categorize the MET score in order to achieve equal sized but meaningful categories.

Smoking Status: Smoking status was divided into three categories: current smoker, past smoker, and never smoked. The categorization was based on studies that examined lifestyle

factors and components of the One-Carbon Metabolism Cycle (152,153). 'Current smoker' was defined as smoking at least one cigarette a day for the past month. 'Past smoker' was defined as not being a current smoker but having had at least one cigarette a day for six months or more in the past. 'Never smoker' included all participants who did not meet the criteria to be considered a current or past smoker. If there was a sufficient number of smokers, over 15 percent, then categories would be created to represent smoking intensity and duration.

Alcohol Consumption: The quantity and frequency of alcohol consumption in the previous month was addressed in the questionnaire. From the information provided, the average number of drinks per week was calculated. Categories were created on the basis of logical and approximately equal-distant categories. Five categorizes were used to assess alcohol consumption: less than 1 drink per week, 1 to 3 drinks per week, 4 to 6 drinks per week, 7 to 9 drinks per week, and 10 or more drinks per week. The literature does not provide clear thresholds for categorizing alcohol consumption. These cut-offs were based on previous studies indicating that more than one drink per day or 7 drinks per week was associated with an increased incidence for multiple cancers (154) and evidence suggesting that 10 or more drinks per week was associated with heart disease and cancer (155).

Folate Level: Serum folate levels were measured in nmol/L according to established laboratory procedures (156). Quartiles were used to categorize folate levels.

Lipids: Lipids refers to a classification of normally occurring molecules that are hydrophobic small molecules. Biologically, lipids serve a large variety of biological functions. Triglycerides and two types of cholesterol (i.e. HDL and LDL) were the lipids examined in this study.

Triglycerides: Triglycerides are involved in the conversion of dietary components to fat storage and play a role in metabolism. High triglyceride levels have been associated with diabetes, atherosclerosis, and other cardiovascular diseases. Categories were based on clinical cut-offs from the American Heart Association (157,158). Triglycerides were categorized into four levels: optimal with values less than 1.3mmol/L, normal with values between 1.3mmol/L to 1.7mmol/L, moderate with values between 1.7mmol/L to 2.2 mmol/L, and high with values equal to or above 2.2mmol/L.

Cholesterol: High Density Lipoproteins (HDLs) and Low Density Lipoproteins (LDLs) were considered for this thesis. Both of these were measured in the unit mmol/L. HDL is commonly referred to as the 'good' cholesterol because it reduces the amount of plaque in arteries. LDL is referred to as the 'bad' cholesterol because it can form plaque in arteries. HDL levels were divided into four categories based on standard clinical cut-offs: very low were values less than 1mmol/L , low were values from 1mmol/L to less than 1.3mmol/L, medium were values from 1.3mmol/L to less than 1.5mmol/L, and high were values equal to or above 1.5mmol/L being high. LDL levels were divided into four categories based on standard clinical cut-offs. Low were values less than 2.6mmol/L, medium were values from 2.6mmol/L to less than 3.4mmol/L, borderline-high were values from 3.4mmol/L to values less than 4.1mmol/L, and high were values equal to or above 4.1mmol/L (157,158).

3.5 Data Validity

3.5.1 Overview of DNA Methylation Validity

Dr. Pang's laboratory group has been developing the HRM approach to measure DNA methylation levels. Reliability of the method is supported by the methodology used.

For each plate the methylation values are interpolated based on methylation levels of a series of standards with known methylation levels. In order to evaluate the reliability of this approach, the methodology used permitted the reliability within plates to be assessed and the variability between plates to be assessed. This is important to ensure the accuracy and reliability of the method to measure methylation of subject samples consistently within and between plate runs.

3.5.2 Outcome Validity

The standards used in this study were purchased from Zymo Research. Zymo Research conducts quality control to ensure that methylation standards adhere to certain criteria. The non-methylated human DNA is produced by using a cell line with genomic knock-out of DNMT1 and DNMT3. The methylation level of the non-methylated standard is quantified by two methods to ensure that the methylation level is less than 5%. Zymo Research generates their methylated human DNA by enzymatically attaching a methyl group to all CpG sites. Similarly, their methylated standard is subjected to rigorous quality control methods to ensure that the methylated level is greater than 95% (159).

3.5.2.1 Reliability of DNA Methylation Measurement

The reproducibility of the final DNA methylation level was measured using different amounts of genomic DNA for bisulfite conversion by Tse et al (150). It was shown that using 2µg of starting DNA for the bisulfite conversion compared to 1µg and 0.5µg had the least amount of variability (150). Therefore, for the bisulfite conversion, 2µg of starting DNA was used for all participants and for conversion of the DNA standards. Additionally, the primers

designed were specific to bisulfite converted DNA and did not amplify unconverted DNA template.

Bisulfite converted (BSC) DNA is less stable than unconverted DNA and it is recommended BSC DNA be stored at -20ºC and multiple freeze-thaw cycles be avoided. Bisulfite DNA samples for each subject were frozen directly following conversion. All bisulfite converted DNA samples were thawed once to quantify the amount of DNA for conversion and at the same unfrozen time samples were diluted and two aliquots were made for qPCR. Then one aliquot used for LINE-1 DNA methylation level measurement was unfrozen in the hours preceding the experiment. Therefore, actions were taken to minimize the number of times a sample thawed. A second aliquot of bisulfite converted DNA was stored at -20 ºC in case a sample had to be run a second time.

3.5.2.2 Reliability between Plate and Adjustments

Every plate has 96 sample wells and each experimental plate held a series of standards with known varying methylation levels to produce a standard curve, an internal control sample, 25 subject samples, and non-template-controls (NTCs) in triplicate. Due to the fact that only 25 participants could be run on a single plate, measures were taken to ensure a minimal amount of variability between plates and the same internal control was placed on every plate to permit a comparison between plates. First, all standard curves were generated from the same Zymo standards and the same bisulfite conversion reactions. The standards were diluted to 1ng/µL, aliquots for all plates were made at one time, and the 5 points for the standard curve were made. It was predetermined that if the internal control varied by more than 2.0% on a single plate from its average on all plates then that plate

would be re-run. This measure was taken to identify any plates that systematically differed in the percent methylation measurement.

To assess the reproducibility of the triplicate measures a coefficient of variation was calculated. The coefficient of variation measures the variability in methylation measurement due to the assay rather than true differences (160).

3.6 Data Management

Each sample was assigned a specific laboratory three-digit ID. These IDs were used throughout the laboratory procedures. For the methylation experiments, all plate layouts were recorded on plate template forms before the experiment and then entered into the RocheLight Cycler Program. Data for each experiment was saved on the computer designated for the Roche LightCycler 480. An Excel spreadsheet and a Prism file were generated for each plate with the standard curve and interpolated data saved. Two final spread sheets were generated, one with the methylation measurements for each sample, Cp values, and plate number, and the second spread sheet contained the internal control values for each plate. These values were double-checked from the original results prior to being entered into the database.

3.7 Overview of Data Analysis

The goal of this analysis was to examine the relationship between BMI and LINE-1 DNA methylation levels. The main analysis strategy used linear regression with percent methylation as the outcome and BMI as the exposure of primary interest. Predictors of LINE-1 DNA methylation levels were controlled for in the adjusted model. Figure 3-4 is a diagram of the potential confounders considered in the model between BMI and LINE-1 DNA methylation.

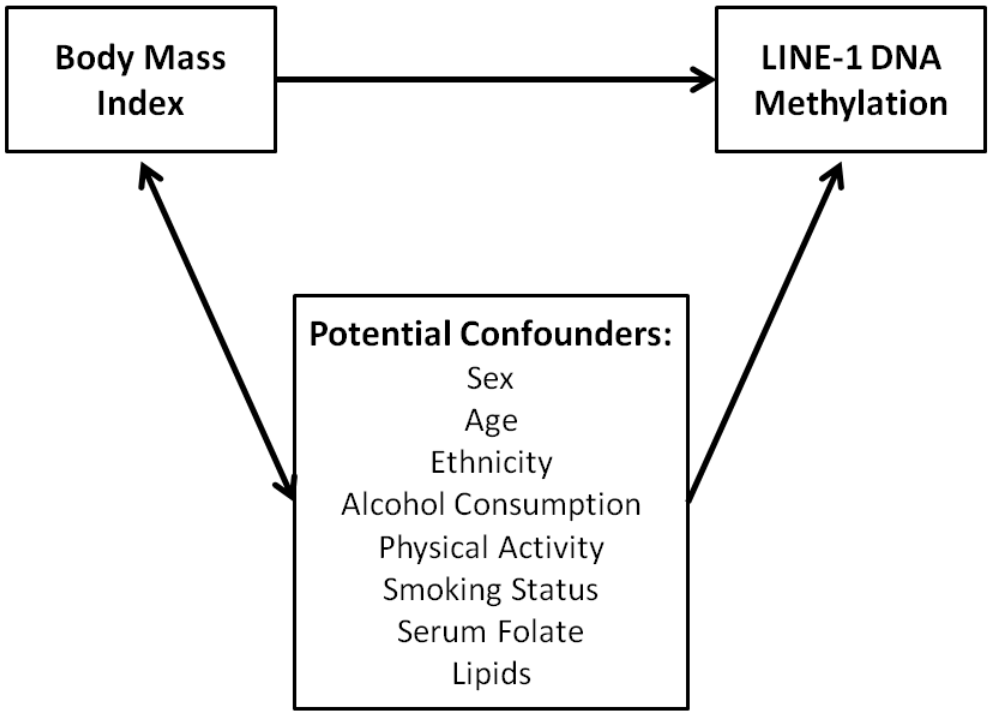

Figure 3-4. Causal Diagram of Potential Confounders of Relationship of Interest.

The initial step was an examination of the distribution of the outcome (LINE-1 DNA methylation), exposure (BMI), and potential confounders. In order to understand the potential for confounding, bivariate analysis was conducted between the outcome and potential confounders, and between the exposure and potential confounders. A covariate model identifying predictors of the outcome at a liberal p-value of 0.20 was created. The main analysis examined the relationship between BMI and percent LINE-1 DNA methylation controlling for predictors of LINE-1 methylation identified by the covariate model. Secondary analysis considered potential interaction with sex and consideration of a dichotomous outcome of DNA methylation levels.

3.7.1 Descriptive Statistics

Descriptive statistics for the exposure, outcome, and all covariates considered in this study were carried out. Percent DNA methylation, the outcome of interest, was described as a continuous variable. The mean, standard deviation, and the 25^{th}, 50^{th} and 75^{th} percentile were calculated and provided. In addition, the distribution of DNA methylation was tested to see if it was normally distributed. For categorical variables the frequencies of each category were determined. A summary table outlining the demographic information of the study population was generated.

Bivariate analyses were conducted between the exposure and outcome with all covariates being considered. To examine the relationship between the continuous measures and categorical variables an F-statistic and T-test, and corresponding p-values were used. A chi-square statistic was used to examine the relationship between categorical variables.

3.7.2 Model Selection

A least squares regression model was used to evaluate the relationship between BMI and LINE-1 DNA methylation. The assumptions of a least squares regression model were assessed and included in the analysis section. BMI, if considered as a categorical variable, permits a non-linear relationship between BMI and methylation levels. Coefficients in the model can be interpreted as a difference in percent DNA methylation for obese and overweight individuals compared to those with normal BMI.

3.7.2.1 Control Strategy for Potential Confounding

The basic strategy was to create a parsimonious model (from among potential covariates) predicting the outcome. The potential confounders that were considered were: age, sex, physical activity, smoking status, alcohol consumption, ethnicity, serum folate, and

lipids. The covariate model was developed by using a step-wise backwards elimination with a liberal p-value of 0.2. The use of a liberal p-value ensures the inclusion of any potential confounders in the covariate model (161,162). The sub-set of covariates that were retained in the model using this approach was included as potential confounders in each multivariate analysis. The step-wise backwards elimination method was selected over a backward elimination or change-in-estimate method because of the existing study power and the small number of potential covariates. The main advantage of the change-in-estimate approach is the elimination of covariates which are not truly confounding the relationship. Given a small number of potential confounders there is no advantage to this approach in this study. The main analysis examined the relationship between BMI and percent LINE-1 DNA methylation controlling for predictors of LINE-1 methylation identified by the covariate model.

3.7.2.2 Potential for Effect Modification

Sex had the potential to modify the relationship of interest based on the fact that the distribution of adipose tissue varies by sex and that BMI as a measure may not have the same meaning for different sexes. In the main regression model, an interaction term between sex and BMI was tested.

3.7.2.3 Sensitivity Analysis

Physical activity and lipid levels could potentially be on the biologic pathway between obesity and DNA methylation changes. Physical activity could mediate or precede weight gain in the process of becoming overweight or obese. Triglycerides, LDL, and HDL also have the potential to fall on the pathway between obesity and changes in methylation

levels. Thus, if necessary, sensitivity analyses were planned to be carried out to compare the relationship of interest with and without these covariates.

Little is known about percent DNA methylation as a pathologic entity. As a result it is not possible to determine the most appropriate representation of this outcome in a statistical analysis. A sensitivity analysis explored dichotomous outcome variables representing hypomethylation and hypermethylation. The lowest 15% of LINE-1 DNA methylation values in each sex were considered as events of DNA hypomethylation. The highest 15% of LINE-1 DNA methylation values in each sex were considered as events of DNA hypermethylation. Similar to the main analysis strategy, covariate models were generated using a backward elimination with a liberal p-value of 0.2 to control for potential confounders and sex was considered as an effect modifier of the relationship. Logistic regressions modeled BMI as the exposure of interest and hypomethylation or hypermethylation as a dichotomous event outcome, adjusted for the covariate model and with an interaction term for sex and BMI. The odds ratio and 95% confidence interval were used to assess the relationship between BMI categories and hypomethylation or hypermethylation.

3.8 Power and Detectable Effect

The power of the study was considered in terms of a detectable change in percent DNA methylation with the overweight and obese BMI categories in comparison to those with normal BMI. *A priori* the power and detectable effect were calculated based on a sample of 155 subjects with normal BMI values, 155 subjects with overweight BMI values, and 90 subjects with obese BMI values. It was determined that the study would have 99.3% power to detect a half a standard deviation difference in percent DNA methylation between those categorized as overweight compared to normal and 96.4% power to detect a half a standard

deviation difference between those categorized as obese compared to normal (with an alpha of 0.05). However, the number of subjects included in the study is considerably different from the *a prior* estimation. So post-hoc the power was calculated. With the total sample size of 502 subjects, with 284 subjects with a normal BMI value, 144 with an overweight BMI value, and 74 with obese BMI values, this study had 97.0% power to detect a half a standard deviation difference in percent LINE-1 DNA methylation between those categorized as obese compared to normal and 90.1% power to detect a third of a standard deviation difference between those categorized as overweight compared to normal (163).

3.9 Design Effect

The study sample for this thesis was selected from the larger study sample, which utilized cluster sampling to recruit subjects from three institutions. Observations may be more similar within a clustered sample compared to random sample (164,165). Cluster sampling is not expected to be a concern as neither the exposure values or outcome values are likely to be associated with the study centre. However, variables representing the study centre were included as a random effects parameter in the multivariate regression model, to account for the possibility that observations will be more similar in this cluster sampling design.

3.10 Ethical Considerations

This research, specifically measuring LINE-1 DNA methylation levels, was consistent with the consent given by participants and ethics approval that were obtained for the larger study. Ethics approval for this thesis project was obtained from the Health Sciences and Affiliated Teaching Hospitals Research Ethics Board (REB) at Queen's University and had the study code EPID-316-10. A copy of the approval is attached in Appendix C. The

application to the REB outlined that the information in the database and blood samples must not have personal identifiers and that only ID numbers were accessible and used by the student investigator.

Chapter 4

Results and Analysis

4.1 Validity of DNA Methylation Measurement

The first objective of this thesis was to describe LINE-1 DNA methylation levels in a large sample. There were several challenges in this regard, including the application of a novel method, a large number of participant samples, and a small amount of starting DNA. In addition, there was no opportunity for a comparison with a gold standard DNA methylation measurement. Therefore an assessment of the validity of the LINE-1 DNA methylation measurement was critical to this objective. In this section, validity was assessed in terms of evaluation of the validity of the method and reliability of sample analysis.

4.1.1 Evaluation of the Validity of the Methylation Measurement Method

A consistent methodology was used to measure LINE-1 DNA methylation levels. Experiments were conducted to evaluate the efficiency of the assay using varying amounts of starting template. A dilution curve, using 6 different quantities of starting bisulfite converted (BSC) DNA template was run on a plate to calculate the efficiency of the assay and to evaluate the amount of starting template that should be included. On two different dilution assays (an example can be seen in Figure 4-1), which contained 10ng, 5ng, 2ng, 1ng, 0.5ng, and 0.1ng of starting BSC DNA template, the efficiencies were 90% and 93% respectively.

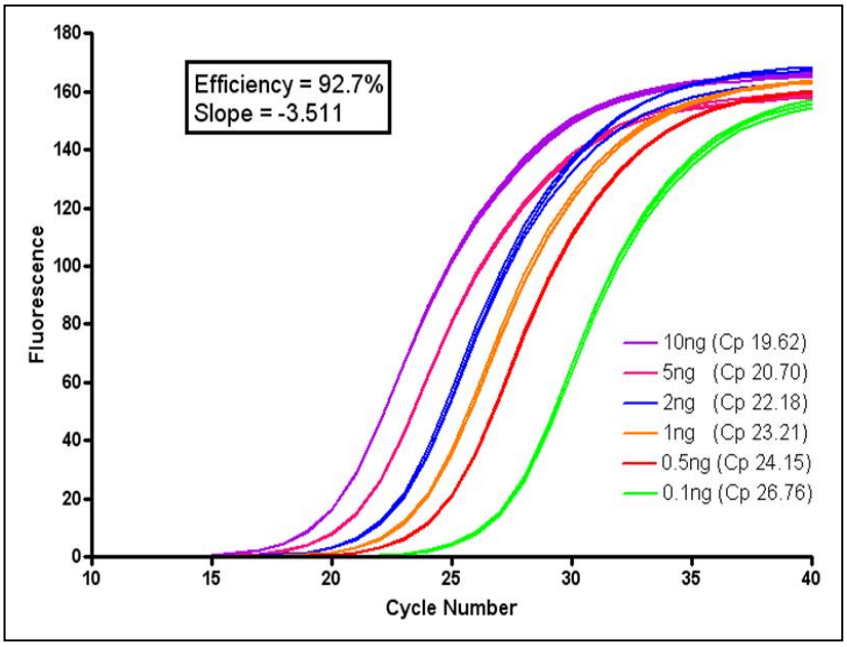

Figure 4-1. Dilution Curve Efficiency. The amplification of different starting quantities (0.1, 0.5, 1.0, 2.0, 5.0, and 10.0ng of bisulfite converted DNA).

The initial amount of BSC DNA template in each reaction well was evaluated using the combined results from the two dilution assays in order to have a greater number of replicates for each point. This can be seen in Figure 4-2 and it was determined that there were no significant differences in percent methylation between 10.0ng, 5.0ng, 2.0ng, 1.0ng, and 0.5ng. However, when only using 0.1ng of template, the methylation level was significantly different from 2.0ng and the variability in the measurement increased. Therefore it was determined that the starting quantity for each sample template would be 2.0ng of bisulfite converted DNA in every well to minimize variation in methylation measures. A large benefit of this assay was that small variations in the starting quantity of BSC DNA did not affect the overall methylation measurement unless the value was approaching 0.1ng of starting BSC

DNA template. A cutoff value of a Cp greater than 26.5 was used to exclude any samples with an unreliable amount of starting BSC DNA template.

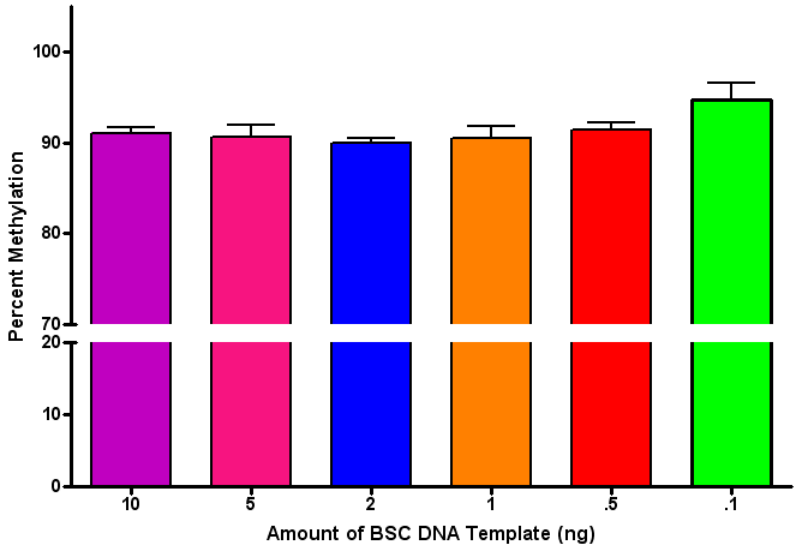

Figure 4-2. Percent LINE-1 Methylation Levels for Dilution Assay. The results of the dilution assay show similar LINE-1 methylation levels for all starting amounts of BSC DNA except for 0.1ng, which also has a larger standard error.

PCR bias occurs when one product is preferentially amplified over another. Since the primers designed for the LINE-1 region were designed complementary to the bisulfite converted methylated DNA, it was important to optimize the assay to ensure that both the unmethylated and methylated products were amplified with equal efficiency. The annealing temperature was adjusted to permit a degree of mismatch base-pairing when the primer bound to a complementary region. Optimization of the annealing temperature was completed between 54.0ºC and 46.0ºC by 2.0ºC temperature changes in order to determine the optimal annealing temperature. In Figure 4-3 the unmethylated and methylated standards show very similar Cp values and thus are amplifying with the same efficiency. This indicates that there

is no PCR bias due to the fact that the primers are binding equally to the unmethylated and methylated BSC DNA templates.

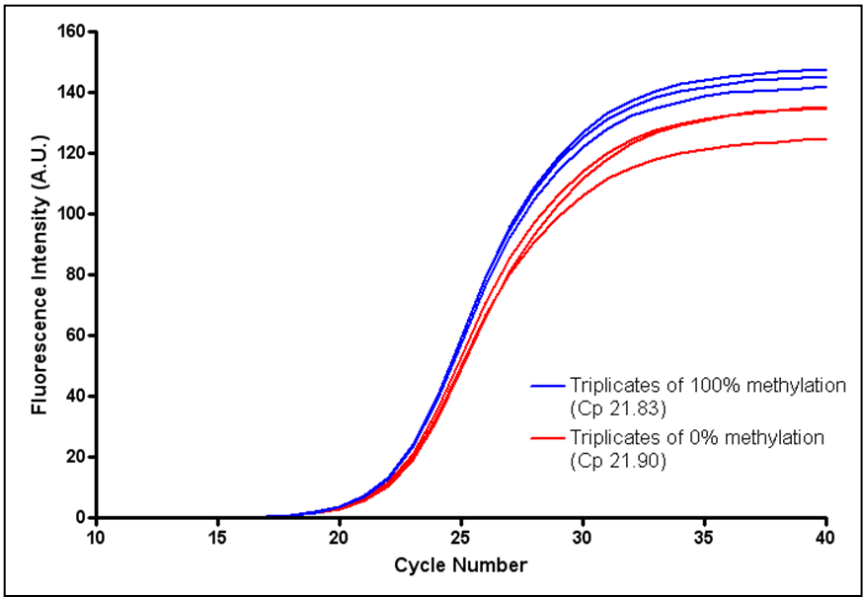

Figure 4-3. The Crossing Point (Cp) Values for the Methylated and Unmethylated Standards. The Cp values were similar indicating that both standards are amplifying with the same efficiency, and therefore no PCR bias.

4.1.1.1 Statistical Assessment of Initial Amount of DNA Template

Although the amount of bisulfite converted DNA template was optimized and demonstrated to be constant across the amount of template, further calculations were done to assess whether there was any PCR bias towards the methylated DNA. Thus the Cp value was assessed as a possible determinant of DNA methylation levels in a statistical model and it was not significant. The parameter estimate was -0.128, with a p-value for the Cp value to predict DNA methylation level of 0.36 indicating further that PCR bias was not a concern in the study.

4.1.2 Reliability of the Method on Subject Samples

The ability of the newly developed method to reliably measure DNA methylation levels is critical to this thesis. Two aspects of reliability were assessed. First, the reproducibility of triplicate values for subject samples on a plate (intra-assay reproducibility) was assessed with a one-way analysis of variance (166,167) to calculate the coefficient of variation. Second, inter-assay reproducibility was assessed based on methylation values of the internal control (168). Therefore the reproducibility of the methylation measurement was evaluated by examining the coefficient of variation between replicates and between plates.

4.1.2.1 Reliability of Replicated within Subjects and Between Plates

For the 502 participants included in the study the coefficient of variation between replicates (intra-assay reproducibility) was 1.71% (Table 4-1). Each plate (n=23) contained an identical internal control sample. It was predetermined that if any value for the internal control varied by more than 2.0% from the mean then the plate would be redone. The coefficient of variation for the internal control on 23 plates was 0.87%. For only 4 of 23 plates was the internal control value more than 1.0% different than the overall mean. The internal control DNA methylation value did not reach the pre-determined rule for repeating a plate (e.g. a difference of more than 2.0% from the overall internal control mean) for any of the plates. Therefore, no plates needed to be repeated. This illustrates that the assay was replicable and that there were only small differences in the methylation values for the internal control between plates permitting an appropriate comparison of all samples without any adjustments of values.

Table 4-1. Reproducibility of LINE-1 DNA Methylation Measurement.

	Number of Triplicate Samples	Root Mean Square Error	Grand Mean	Coefficient of Variation (%)
Intra-assay variation	502	1.447	84.52	1.71
Inter-assay variation	23	0.812	93.44	0.87

4.2 Descriptive Statistics

4.2.1 Study Sample

The larger study collected information on 663 participants. Participants were excluded from this analysis due to: non-fasting blood sample (n=80), missing or inadequate blood sample for DNA analysis (n=24), unreliable DNA methylation (n=33), missing BMI information (n=8), BMI value considered underweight (n=7), and missing covariate information (n=9). All analysis presented are based on the resulting sample of 502 participants with complete DNA methylation values, BMI, and covariate information.

4.2.2 Distribution of Outcome, Exposure and Covariate

4.2.2.1 Distribution of LINE-1 DNA Methylation

LINE-1 DNA methylation was measured as a continuous variable (i.e. percent methylation). The distribution of LINE-1 DNA methylation is presented in Figure 4-4 with a normal curve super-imposed onto the histogram. LINE-1 DNA methylation values had a mean of 84.52%, a standard deviation of 3.19%, and a median of 84.32%. LINE-1 DNA methylation values varied by 22.28% with a range of values from 71.19% to 93.47%. Notably, the interquartile range was only 4.25%, with a 25^{th} quartile value of 82.45% and a 75^{th} quartile value of 86.70%.

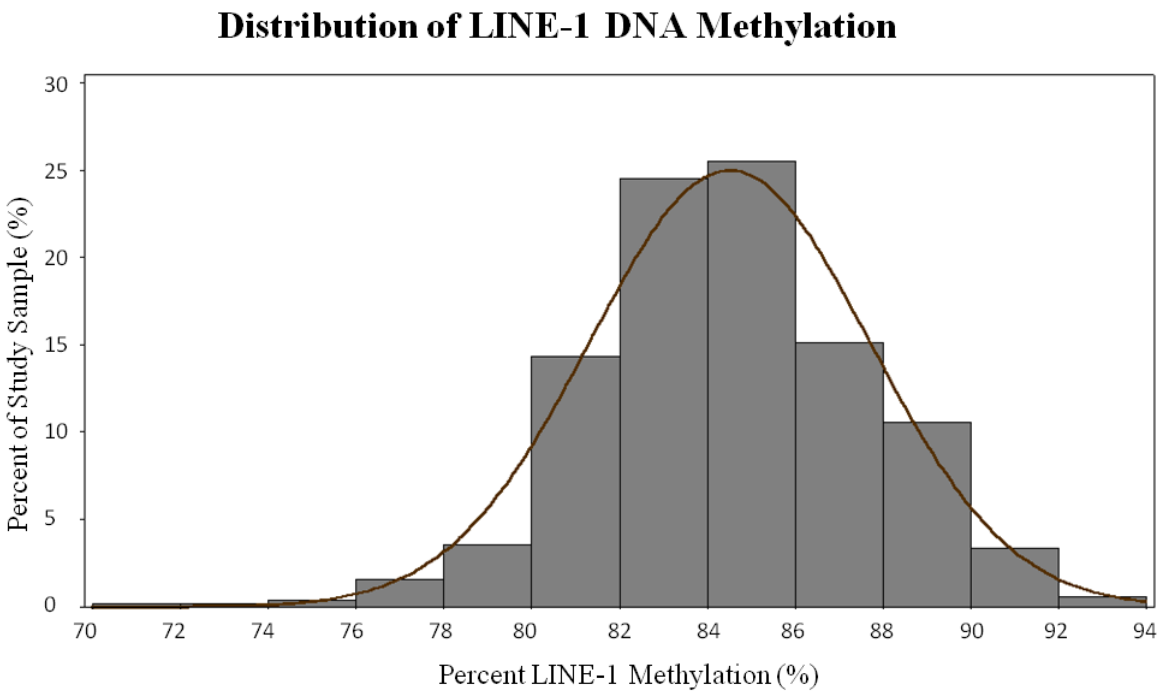

Figure 4-4. Histogram of LINE-1 DNA Methylation Levels in Study Sample. A normal curve is superimposed over the distribution of percent LINE-1 DNA Methylation.

The Kolmogorov-Smirnov goodness-of-fit test for normal distribution had a p-value of 0.12. Therefore, the distribution of LINE-1 DNA methylation can be considered to be normally distributed. However, the distribution is slightly skewed to the left, having a skewness measure of -0.15.

4.2.2.2 Distribution of Body Mass Index

Body Mass Index was calculated for each participant. Subjects with BMI values below 18.5kg/m^2 were not included in the study. The distribution of BMI is presented in Figure 4-5. A gamma curve was superimposed over the histrogram because it appeared to best fit the distribution.

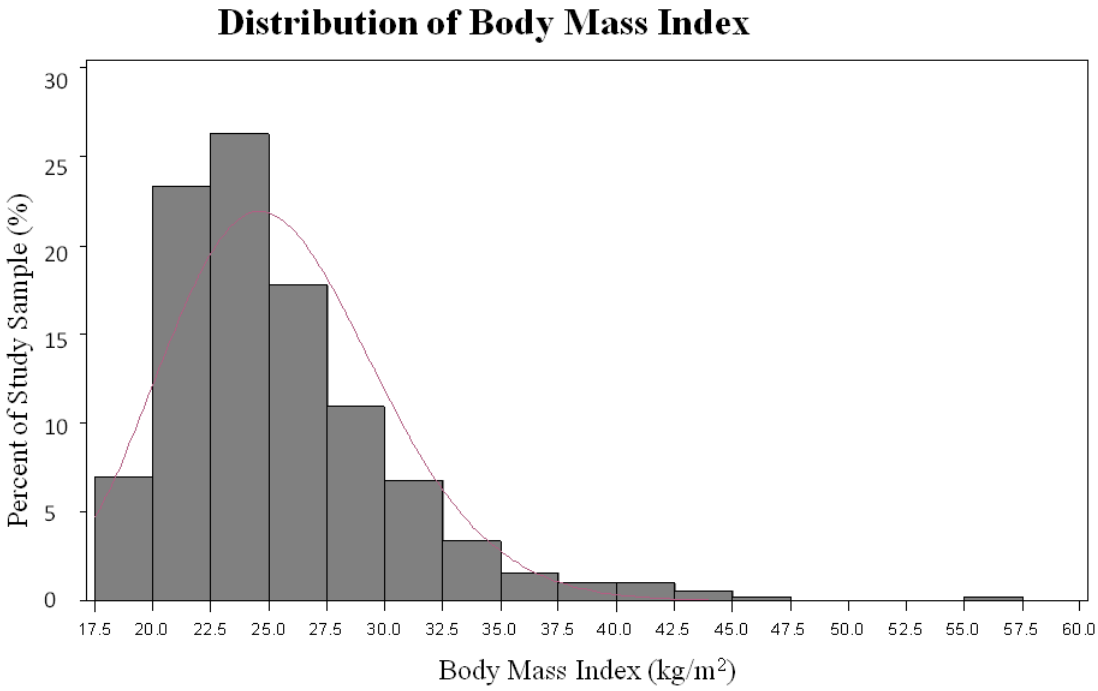

Figure 4-5. Histogram of BMI in Study Sample. A gamma curve is superimposed over the distribution of percent LINE-1 DNA Methylation.

Subsequently, categories for BMI were created. The percentage of the study sample categorized as normal, overweight, and obese were 56.6%, 28.7%, and 14.7% respectively, presented in Table 4-2.

Table 4-2. Frequency of BMI Categories in Study Sample.

Variable	Categories	N	Percentage (%)
BMI	Normal (18.5 – < 25.0kg/m^2)	284	56.6
	Overweight (25.0 – < 30.0kg/m^2)	144	28.7
	Obese (≥ 30.0kg/m^2)	74	14.7

4.2.2.3 Distribution of Covariates

All of the covariates were either fundamentally categorical or categories were created from continuous variables (see Section 3.4.3). For each variable the frequency of each category and percentages are presented in Table 4-3.

Subject recruitment was designed to result in an approximately balanced distribution by sex and age. The study population for this thesis had a greater proportion of females (59%) and a relatively balanced age distribution. The majority (88%) of the study sample described their parents' ethnicity as Caucasian. Twenty percent were past smokers and 14% current smokers. Almost a third of participants consumed less than 1 drink of alcohol per week, and 22% consumed over 7 drinks per week. Categories for physical activity and serum folate levels were based on quartiles.

For the lipid variables, 6% of the study population was in the highest category for triglyceride level, 20% were in the very low category for HDL level, and 9% had high LDL levels.

Table 4-3. Frequency and Percentages of Categorical Variables

Variable	Categories	N	Percentage (%)
Sex	Male	207	42.2
	Female	295	58.8
Age	20 – 29 years of age	181	36.1
	30 – 39 years of age	147	29.3
	40 – 49 years of age	174	34.7
Ethnicity	Caucasian	438	87.8
	Asian	40	8.0
	Other	24	4.8
Smoking Status	Never	330	65.7
	Past	101	20.1
	Current	71	14.1
Alcohol Consumption	Less than 1 drink/week	157	31.3
	1 – 3 drinks/week	162	32.3
	4 – 6 drinks/week	70	13.9
	7 – 9 drinks/week	52	10.4
	10 or more drinks/week	61	12.2
Physical Activity	Quartile 1 ($\leq$ 1117 METs)	125	24.9
	Quartile 2 (1118 – 2273 METs)	125	24.9
	Quartile 3 (2274 – 4398 METs)	126	25.1
	Quartile 4 ($\geq$ 4399 METs)	126	25.1
Serum Folate Level	Quartile 1 ($\leq$ 21.4 nmol/L)	125	24.9
	Quartile 2 (21.5 – 28.1 nmol/L)	128	25.5
	Quartile 3 (28.2 – 35.5 nmol/L)	123	24.5
	Quartile 4 ($\geq$ 35.6 nmol/L)	126	25.1
Triglycerides	Optimal (< 1.3 mmol/L)	355	70.7
	Normal (1.3 – < 1.7 mmol/L)	82	16.3
	Moderate (1.7 – < 2.2 mmol/L)	34	6.8
	High ($\geq$ 2.2 mmol/L)	31	6.2
HDL	Very Low (< 1.0 mmol/L)	98	19.5
	Low (1.0 – < 1.3 mmol/L)	197	39.2
	Medium (1.3 – < 1.5 mmol/L)	111	22.1
	High ($\geq$ 1.5 mmol/L)	96	19.1

Table 4-3 Continued

Variable	Categories	N	Percentage (%)
LDL	Low (< 2.6 mmol/L)	181	36.1
	Medium (2.6 – < 3.4 mmol/L)	196	39.0
	Borderline High (3.4 – < 4.1 mmol/L)	79	15.7
	High (≥ 4.1 mmol/L)	46	9.2

4.3 Bivariate Analysis of the Relation of Covariates to Exposure and Outcome

This section describes the relationships between BMI and all covariates, as well as, the relationship between LINE-1 DNA methylation and all covariates. The intent of these analyses was to inform on potential confounders of the relationship of interest in terms of the direction and strength of relationships. In addition, these analyses were used to identify problematic distributions for subsequent regression analysis (e.g. low cell counts).

4.3.1 Relationship between BMI and Covariates

The relationship between BMI and covariates of interest was assessed using a chi-square test statistic and these relationships are presented in Table 4-4. Where a cell count was 5 or below a Fisher's exact test was employed.

Statistically significant associations (p-value <0.05) were observed for the relationship of BMI with sex, age, ethnicity, smoking status, serum folate level, triglycerides, HDL, and LDL. There were no statistically significant associations between BMI with physical activity and alcohol consumption.

Patterns of the associations of BMI with covariates are represented with row percentages in Table 4-4. Males were more likely to be overweight and obese than females. Those in the youngest age group had the healthiest distribution of BMI and were less likely to

be overweight or obese compared to the older age groups. With increasing age, there was a corresponding increase in the percentage of individuals in the obese BMI category. The distribution of BMI varied considerably by ethnicity. Individuals of Asian descent were more likely to have a healthy BMI with a very low percentage categorized as obese compared to Caucasians. Individuals categorized as 'Other' (predominantly comprised as African American or Aboriginals) had the largest percentage of individuals with obese BMI values. Physical activity was not significantly associated with BMI. The first three quartiles of physical activity showed a pattern of increasing activity related to decreased likelihood of being obese and increased likelihood of having a healthy BMI. However, the highest quartile of physical activity did not follow this pattern with a higher percentage of individuals in the obese BMI category reporting higher physical activity levels compared to the second and third quartiles. Alcohol consumption was not significantly associated with BMI. The two extreme categories, consumption of less than 1 drink per week and consumption of 10 or more drinks per week, had the lowest percentages of subjects with a normal BMI compared to the other categories.

Serum folate was significantly associated with BMI. The bottom two quartiles of serum folate levels had a much lower percentage of individuals with normal BMI values compared to the two highest quartiles of serum folate levels. As would be expected from the literature, triglycerides showed a clear and significant relationship with BMI. Those with optimal triglyceride levels were most likely to have a normal BMI category and least likely to be obese. With increasing triglyceride levels there was an increasing percentage of individuals categorized as overweight and obese and decreasing percentages of individuals with healthy BMI values. High Density Lipoprotein (HDL) is the 'good' cholesterol and the

HDL level was inversely related to BMI. Those with very low HDL levels were more likely to be overweight or obese compared to a normal BMI. With increasing HDL levels the percentages in the normal BMI category increased and the percentages in the obese category decreased. Low Density Lipoprotein (LDL) was significantly associated with BMI. Individuals in the lowest LDL category were the most likely to have a normal BMI and the least likely to be obese compared to all other LDL categories.

Table 4-4. Associations between BMI and Covariates (Row percentages are presented)

Variable	Categories	N	BMI – Normal (N = 284)	BMI- Overweight (N = 144)	BMI- Obesity (N = 74)	Overall P-value*
Sex	Male	207	45.9%	35.8%	18.4%	
	Female	295	64.1%	23.7%	12.2%	< 0.01
Age	20 – 29 years	181	69.1%	21.0%	9.9%	
	30 – 39 years	147	49.7%	34.7%	15.7%	
	40 – 49 years	174	49.4%	31.6%	19.0%	< 0.01
Ethnicity	Caucasian	438	55.3%	29.5%	15.3%	
	Asian	40	77.5%	20.0%	2.5%	
	Other	24	45.8%	29.1%	25.0%	0.02[+]
Physical Activity	Quartile 1	125	48.8%	32.0%	19.2%	
	Quartile 2	125	58.4%	29.6%	12.0%	
	Quartile 3	126	62.7%	26.2%	11.1%	
	Quartile 4	126	56.4%	27.0%	16.7%	0.33
Smoking Status	Never	330	63.6%	23.9%	12.4%	
	Past	101	43.6%	38.6%	17.8%	
	Current	71	42.3%	36.6%	21.1%	< 0.01
Alcohol Consumption (drinks/week)	Less than 1	157	51.6%	29.9%	18.5%	
	1 – 3	162	59.9%	28.4%	11.7%	
	4 – 6	70	57.1%	25.7%	17.1%	
	7 – 9	52	61.5%	23.1%	15.4%	
	10 or more	61	55.7%	34.4%	9.8%	0 .58

* Chi square p-value

[+] Fisher's exact test was used due to low cell count numbers

Table 4-4 Continued

Variable	Categories	N	BMI – Normal (N = 284)	BMI- Overweight (N = 144)	BMI- Obesity (N = 74)	Overall P-value*
Serum Folate Level	Quartile 1	125	48.8%	34.4%	16.8%	
	Quartile 2	128	48.4%	34.4%	17.2%	
	Quartile 3	123	65.9%	21.1%	13.0%	
	Quartile 4	126	63.5%	24.6%	11.9%	0.03
Triglycerides	Optimal	355	66.2%	24.2%	9.6%	
	Normal	82	39.0%	36.6%	24.4%	
	Moderate	34	38.2%	41.2%	20.6%	
	High	31	12.9%	45.2%	41.9%	< 0.01
HDL	Very Low	98	28.6%	41.8%	29.6%	
	Low	197	49.2%	32.5%	18.3%	
	Medium	111	71.2%	24.3%	4.5%	
	High	96	83.3%	12.5%	4.2%	< 0.01[+]
LDL	Low	181	69.6%	21.6%	8.8%	
	Medium	196	51.5%	31.1%	17.4%	
	Borderline- High	79	44.3%	40.5%	15.2%	
	High	46	47.8%	26.1%	26.1%	< 0.01

* Chi square p-value
[+] Fisher's exact test was used due to low cell count numbers

4.3.2 Relationship between LINE-1 DNA methylation and covariates

The relationship between LINE-1 DNA methylation levels and covariates of interest was assessed using an F-test and a T-test. An F-test statistic was used to assess the overall association. The T-test was used to compare the mean percent LINE-1 methylation for each covariate category with the referent category. The corresponding means, difference in means with their 95% confidence interval and p-values are presented in Table 4-5. LINE-1 DNA methylation levels had a statistically significant association with sex and HDL. Females had

lower LINE-1 DNA methylation levels than males. Those with very low and low levels of HDL had higher LINE-1 DNA methylation in comparison with those with medium or high HDL levels.

Overall, there were not statistically significant associations between LINE-1DNA methylation levels with physical activity and triglycerides. However, there were significant differences between one group compared to the referent group for LINE-1 DNA methylation level differences in physical activity and triglyceride levels. For physical activity, those in the highest quartile of METs had higher LINE-1 DNA methylation levels compared to those in the lowest quartile of METs. Participants in the middle two quartiles of physical activity had LINE-1 DNA methylation levels between those in the lowest and highest quartiles of physical activity, demonstrating an increasing pattern in LINE-1 methylation levels with increasing activity levels. LINE-1 DNA methylation levels were similar across the optimal, normal, and moderate levels of triglycerides. However, those in the high triglyceride category had a statistically significant higher LINE-1 DNA methylation level compared to the other categories. Overall, there were no statistically significant associations between LINE-1DNA methylation levels and age, ethnicity, smoking, alcohol consumption, serum folate, and LDL.

Table 4-5. Bivariate Relationships between LINE-1 DNA Methylation and Covariates.

Variable	Categories	Mean DNA Methylation Level (%)	Mean Change in Percent Methylation (95% CI)	Categorical P-value*	Overall P-value+
Sex	Male	85.92	Referent		
	Female	83.52	-2.39 (-2.92, -1.86)	< 0.01	< 0.01
Age	20 – 29 years	84.58	Referent		
	30 – 39 years	84.37	-0.21 (-0.91, 0.49)	0.55	
	40 – 49 years	84.55	-0.03 (- 0.70, 0.63)	0.92	0.82
Ethnicity	Caucasian	84.48	Referent		
	Asian	84.83	0.35 (-0.69, 1.39)	0.51	
	Other	84.57	0.09 (-1.22, 1.41)	0.89	0.80
Physical Activity	Quartile 1	84.01	Referent		
	Quartile 2	84.55	0.53 (-0.26, 1.33)	0.19	
	Quartile 3	84.50	0.49 (-0.30, 1.28)	0.23	
	Quartile 4	84.98	0.97 (0.18, 1.76)	0.02	0.12
Smoking Status	Never	84.67	Referent		
	Past	84.10	-0.57 (-1.29, 0.14)	0.12	
	Current	84.34	-0.33 (-1.56, 0.49)	0.42	0.26
Alcohol Consumption (drinks/week)	Less than 1	84.53	Referent		
	1 – 3	84.40	-0.12 (-0.83, 0.58)	0.73	
	4 – 6	84.75	0.23 (-0.68, 1.13)	0.62	
	7 – 9	84.41	-0.11 (-1.12, 0.89)	0.82	
	10 or more	84.56	0.03 (-0.92, 0.98)	0.95	0 .96
Serum Folate Level	Quartile 1	84.76	Referent		
	Quartile 2	84.49	-0.27 (-1.06, 0.53)	0.51	
	Quartile 3	84.49	-0.27 (-1.07, 0.53)	0.51	
	Quartile 4	84.30	-0.45 (-1.24,0.34)	0.26	0.74
Triglycerides	Optimal	84.41	Referent		
	Normal	84.46	0.05 (-0.72, 0.81)	0.91	
	Moderate	84.48	0.07 (-1.05,1.19)	0.90	
	High	85.78	1.37 (0.20, 2.54)	0.02	0.74

* P-values based on T-test compared to referent category

+ P-values based on the F-statistic for the overall association

Table 4-5 Continued

Variable	Categories	Mean DNA Methylation Level (%)	Mean Change in Percent Methylation (95% CI)	Categorical P-value*	Overall P-value+
HDL	Very Low	85.31	Referent		
	Low	84.69	-0.62 (-1.38, 0.15)	0.12	
	Medium	83.97	-1.34 (-2.19, -0.47)	< 0.01	
	High	83.95	-1.36 (-2.25, -0.47)	< 0.01	< 0.01
LDL	Low	84.25	Referent		
	Medium	84.52	0.27 (-0.37, 0.92)	0.40	
	Borderline- High	85.04	0.79 (-0.05, 1.63)	0.07	
	High	84.56	0.31 (-0.73, 1.34)	0.56	0.33

* P-values based on T-test compared to referent category
+ P-values based on the F-statistic for the overall association

4.4 Relationship between BMI and LINE-1 DNA methylation

A primary objective of this thesis was to examine the relationship between Body Mass Index and LINE-1 DNA methylation levels. The main analysis for this objective utilized a least squares regression model with LINE-1 DNA methylation regressed on a categorical representation of BMI. The normal BMI group was used as the referent category. Coefficients from this model represented the difference in mean percent LINE-1 DNA methylation between overweight and obese categories, and the normal category. Potential confounders of this relationship were identified by creating a parsimonious model from among the covariates. The final analysis examined the BMI-LINE-1 DNA methylation relationship while controlling for potential confounders. The design effect of recruitment centre was taken into account by including a random effects parameter for centre in each model.

4.4.1 Assessment of Confounding

To assess confounding, a covariate model was generated, as describe in section 3.7.2.1, using the step-wise backwards elimination method with a liberal p-value of 0.2. Table 4-6 shows the backward elimination process, where HDL was the first variable removed and ethnicity the last. The only covariate that met the criteria was sex with a p-value of <0.01. Therefore the covariate model only included the variable sex.

Table 4-6. Covariate Model Selection Process.

Variable	F-Statistic	P-Value	Order of Elimination
HDL	0.13	0.94	1st
Triglycerides	0.20	0.90	2nd
Serum Folate	0.50	0.69	3rd
Age	0.46	0.63	4th
LDL	0.70	0.56	5th
Physical Activity	0.72	0.54	6th
Alcohol	0.91	0.46	7th
Ethnicity	1.51	0.22	8th
Sex	6.18	<0.01	

4.4.2 Modeled relationship of BMI and LINE-1 DNA Methylation

The main analysis of the relationship between BMI and LINE-1 DNA methylation is presented in Table 4-7. BMI was not associated with LINE-1 DNA methylation levels while controlling for sex. Adjusted estimates for the overweight and obese BMI categories were small, less than one-tenth of the standard deviation and went in opposing directions compared

to the normal BMI category. These results did not support a relationship between BMI and LINE- DNA methylation levels.

Sex was a strong predictor of LINE-1 DNA methylation levels with females having an average DNA methylation level 2.35% lower than males. An R^2 value of 0.16 was primarily due to the relationship between sex and LINE-1 DNA methylation

Table 4-7. Linear Regression Model with BMI and Sex. The crude model did not control for any covariates. The adjusted model controlled for the effect of sex, and the estimate, standard error, and p-values are provided.

Variable	Categories	Crude Model*	Adjusted Model*	Standard Error	F-Statistic	P-Value
Intercept		84.81	86.43	0.40		
BMI	Normal	Referent	Referent			
	Overweight	0.097	- 0.31	0.31		
	Obese	0.619	0.22	0.39	0.90	0.41
Sex	Male	Referent	Referent			
	Female	- 2.33	- 2.35	0.27	74.79	<0.01

* The models include a random effects parameter representing recruitment centre

4.4.3 Assessment of Interaction

To assess the third objective of the study, to determine whether there was an interaction between sex and BMI, an interaction term was added to the model. Results of this analysis are presented in Table 4-8. Each cell had a sufficient number of study subjects. The smallest cells had 36 and 38 subjects. The effect of overweight and obesity on LINE-1 DNA methylation is presented for males and females from this model. All coefficients are small (e.g. below 1%) and non-significant. Among males, overweight was associated with a 0.7% decrease in LINE-1 DNA methylation. Among females, obesity was associated with a 0.5%

increase in LINE-1 DNA methylation. Although different patterns were observed for males and females, all the estimates were close to the null value. The interaction term was not significant (p-value = 0.50) demonstrating that sex did not modify the relationship between BMI and DNA methylation levels.

Table 4-8. Change in LINE-1 DNA Methylation Levels by BMI according to Sex

Variable		Number of Males	Males	P-value	Number of Females	Female	P-value
BMI	Normal	95	Referent		189	Referent	
	Overweight	74	- 0.66	0.15	70	- 0.04	0.92
	Obese	38	- 0.16	0.77	36	0.51	0.34

4.4.4 Assumption Tests for Model

The following analyses tested the basic assumptions of linear regression for this data. Homoscedasticity is the assumption of equal variance. The studentized residuals were plotted against the BMI categories (Figure 4-6). The variance was fairly equal across groups with slightly more variability in methylation values in the normal BMI category.

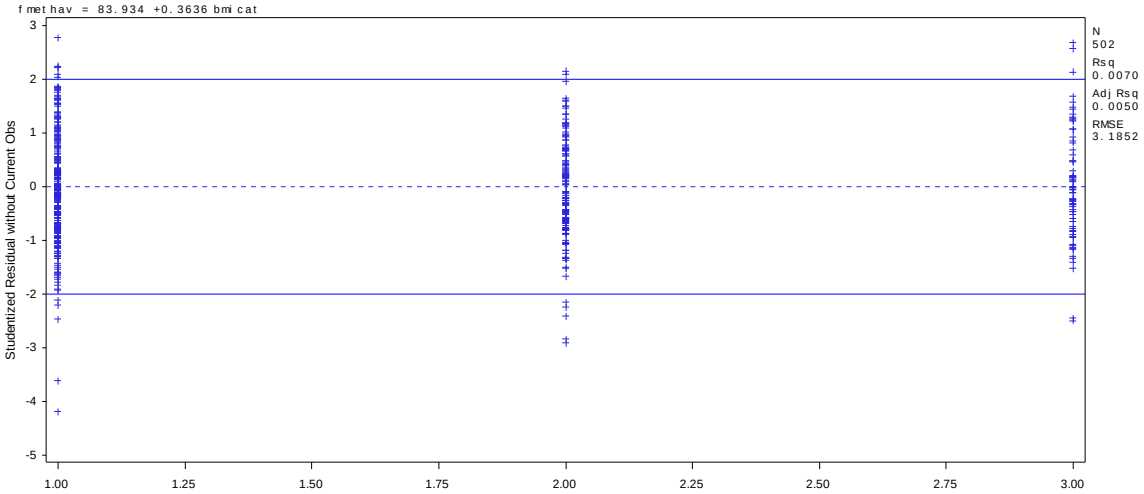

Figure 4-6. Homoscedasticity Assessment. Residuals were plotted against BMI categories to assess the homoscedasticity assumption.

The normality assumption states that the error is normally distributed. A Q-Q Plot, a plot of the residual quantiles verses normal probability quantiles, was generated. The plot showed a relatively straight diagonal line which indicates that the normality assumption is most likely met.

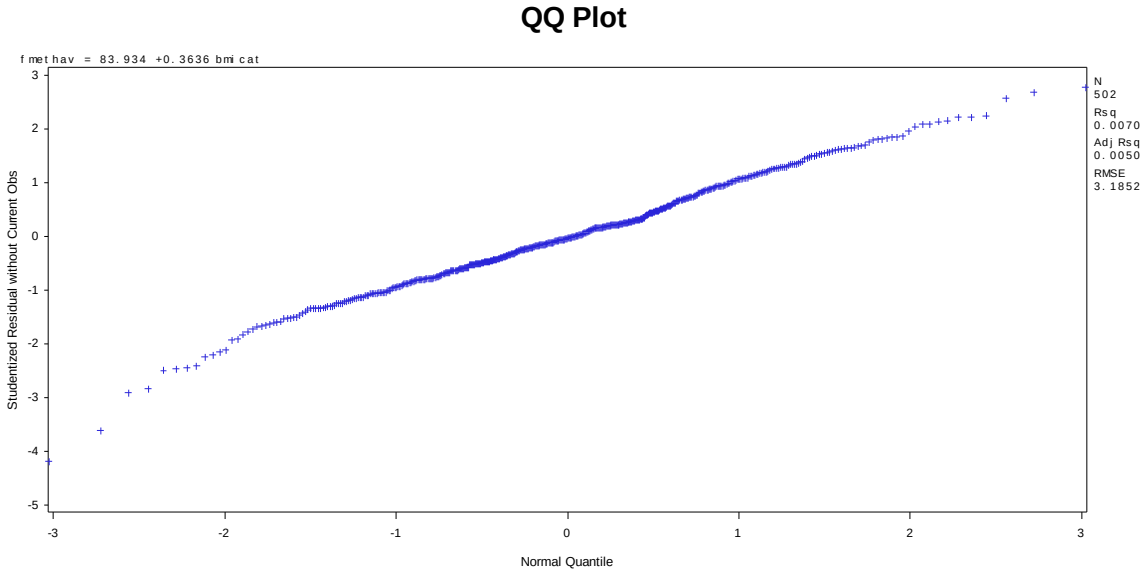

Figure 4-7. Normality Assessment. Q-Q plot was generated to assess the normality assumption.

Cook's D plot was generated to check for outliers. All of the values for Cook's D were considerably below 1; the highest value was less than 0.035. This indicates that no single observation was driving the parameter estimate.

4.5 Sensitivity Analyses

4.5.1 Logistic Regression with LINE-1 DNA Methylation as a Dichotomous Variable

The objective was to examine BMI as a potential determinant of LINE-1 DNA hypomethylation and hypermethylation because increased risk for adverse health events may act through increased or decreased methylation levels. Although there is great interest in DNA methylation as a determinant of chronic disease, there is still limited understanding of DNA methylation as a pathologic entity. As a result, there is no definitive way to conceptualize this outcome for analysis. The sensitivity analyses considered dichotomous representations of LINE-1 DNA methylation for hypomethylation and hypermethylation. Sex specific cut-points for the lowest and highest 15 percent of methylation were used to define events for each dichotomous outcome. The bottom 15 percent of methylation values for each sex defined DNA hypomethylation, representing values below cut-points of 82.28 percent for males and 80.88 percent for females LINE-1 DNA methylation. The top 15 percent of methylation values for each sex defined DNA hypermethylation, representing values above cut-points of 89.29 percent for males and 86.07 percent for females LINE-1 DNA methylation. Since sex was a potential effect modifier of the relationship between BMI and LINE-1 DNA methylation levels, an interaction term was included in the model to assess the sex specific relationship. The odds ratio was used as a measure of effect and logistic regression employed to control for potential confounders.

A covariate model was generated using backwards elimination for the sensitivity analysis for DNA hypomethylation. HDL met the criteria of determinants of hypomethylation and was included in the logistic model as a potential confounder. A logistic model was generated with HDL, sex, and an interaction term between sex and BMI. The effect of overweight and obesity on LINE-1 DNA methylation is presented for males and females from this model in Table 4-9. Although the interaction between BMI and sex was not statistically significant (p-value= 0.15) the odds ratios were in different directions. Males in the overweight and obese categories were more likely to have DNA hypomethylation outcome (OR= 1.28 and 1.44 respectively) compared to those with a normal BMI. While females in the overweight and obese categories were less likely to have hypomethylation (OR = 0.92 and 0.15 respectively).

Table 4-9. Logistic Regression Model for LINE-1 DNA Hypomethylation

Variable		Odds Ratio for Males (95% CI)*	P-value	Odds Ratio for Females (95% CI)*	P-value
BMI	Normal	Referent		Referent	
	Overweight	1.28 (0.53, 3.08)	0.63	0.92 (0.42, 2.01)	0.84
	Obese	1.44 (0.50, 4.13)	0.42	0.15 (0.02, 1.18)	0.08

*Model adjusted for HDL and effects of centre

A covariate model was generated using backwards elimination for the sensitivity analysis for DNA hypermethylation. Smoking status met the criteria of determinants of LINE-1 DNA hypermethylation and was included in the logistic model as a potential confounder. A logistic model was generated with smoking status, sex, and an interaction term between sex and BMI. The effect of overweight and obesity on LINE-1 DNA methylation is presented for males and females from this model in Table 4-10. Results showed a pattern of

increasing effect of BMI on LINE-1 DNA hypermethylation for females but not males. Males in the overweight category were less likely to have hypermethylation (OR= 0.63) and males in the obese categories were more likely have hypermethylation outcome (OR= 1.38) compared to those with a normal BMI. While females in the overweight and obese categories were more likely to be hypermethylated (OR = 1.23 and 1.92 respectively) compared to those with a normal BMI. Although differences were observed for males and females, the interaction term was not statistically significant (p-value= 0.43).

Table 4-10. Logistic Regression Model for LINE-1 DNA Hypermethylation

Variable		Odds Ratio for Males (95% CI)*	P-value	Odds Ratio for Females (95% CI)*	P-value
BMI	Normal	Referent		Referent	
	Overweight	0.63 (0.28, 1.39)	0.26	1.23 (0.55, 2.76)	0.61
	Obese	1.38 (0.61, 3.12)	0.43	1.92 (0.81, 4.57)	0.13

*Adjusted for smoking status and effect of centre

4.5.2 Covariates Potentially in Biologic Pathway

A priori, it was determined that if a significant association was observed between BMI and LINE-1 DNA methylation, covariates of interest, potentially on the biologic pathway between BMI and DNA methylation would be assessed further. However, this assessment was not necessary based on the fact that these potential covariates were not related to DNA methylation. For example, when comparing the crude to the adjusted logistic regression results there was no difference between the risk estimates when HDL was included suggesting that HDL was not influencing the relationship of interest.

Chapter 5

Discussion and Conclusions

5.1 Summary of Findings

The objectives of this study were to quantify and describe LINE-1 DNA methylation in leukocytes in a large population sample, to examine the relationship between Body Mass Index (BMI) and LINE-1 DNA methylation levels in this study sample, and to investigate sex as a potential effect modifier of the relationship of interest.

Validity of the novel method used to measure LINE-1 DNA methylation measurement was considered. Optimization of annealing temperature for the LINE-1 DNA methylation assay ensured that the primers were not preferentially replicating methylated versus unmethylated DNA template. Further, a statistical analysis of the Cp values demonstrated that the amplification cycle did not predict percent LINE-1 DNA methylation, confirming that PCR bias was not a concern. Since it was difficult to quantify the precise amount of bisulfite converted DNA used for each methylation analysis it was important to assess whether the starting quantity of BSC DNA template influenced the methylation results. Experiments demonstrated that variability in the starting amount of BSC DNA template did not influence the percent LINE-1 DNA methylation level obtained.

Statistical assessment of LINE-1 DNA methylation measurements showed that it was a reliable measure. The inter-assay variation was relatively small with a coefficient of variation of 1.71 percent and the intra-assay variation was also relatively small with a coefficient of variation of 0.87 percent.

LINE-1 DNA methylation levels were normally distributed around a mean of 84.52% and a standard deviation of 3.19%. For the main analysis, a step-wise backward elimination approach with a liberal p-value was used to assess potential confounders with sex being the only significant predictor of DNA methylation. BMI was not significantly associated with LINE-1 DNA methylation levels after controlling for sex. The adjusted linear regression model did not show any trend in response to increasing BMI. The differences in mean percent LINE-1 DNA methylation levels between BMI categories were small for the overweight (-0.31%) and obese (0.22%) compared to the referent normal BMI category.

Sex was not found to be an effect modifier of the relationship between BMI and LINE-1 DNA methylation levels in the linear regression model. All coefficients in the model were small (e.g. below 1%) and non-significant.

In addition, sensitivity analyses were conducted with the bottom 15% of methylation values for each sex considered to be hypomethylated and the top 15% of methylation values for each sex considered to be hypermethylated. The sensitivity analysis for DNA hypomethylation used an adjusted logistic regression model controlling for HDL and did not find a statistically significant relationship between BMI and DNA hypomethylation. The pattern of adjusted odds ratios for DNA hypomethylation showed decreasing odds ratios for females and increasing odds ratios for males with increasing BMI categories. The sensitivity analysis for DNA hypermethylation used an adjusted logistic regression model to control for smoking status and did not find a statistically significant relationship between BMI and DNA hypermethylation. The pattern of odds ratios for DNA hypermethylation showed increasing odds ratios for females with increasing BMI categories but no pattern in the odds ratios for

males. These sensitivity analyses lend support to the hypothesis that the influence of BMI on DNA hypomethylation and hypermethylation may be modified by sex.

5.1.1 Comparison to Literature – DNA Methylation Levels

Aberrant DNA methylation as an important epigenetic modification of DNA has recently emerged as a topic of increasing interest in epidemiologic studies. Published studies vary in the methods used to measure DNA methylation, the type and location of repetitive sequences studied, and the number of CpG sites examined. As a result, direct comparisons of average DNA methylation levels across studies are problematic. The mean LINE-1 DNA methylation was higher in this study sample compared to other studies with a comparable standard deviation. However, the coefficient of variation for the intra-assay reproducibility was much smaller in this study compared to the literature.

When comparing results, consideration should be given to the use of different repetitive sequences measured, for example whether a study used the repetitive sequence LINE-1 or Alu. Both of these types of repeats may represent global DNA methylation levels but they are not equivalent measures. Even within the repetitive sequences themselves, different regions may be methylated at different locations and contain varied numbers of CpG sites. The exact CpG sites being measured could potentially impact upon the findings. Also, the varying methods used to measure DNA methylation levels make comparisons more difficult since each method may be more susceptible to different forms of bias and have different levels of reliability.

In the literature, many studies do not provide descriptions of DNA methylation levels as a continuous variable. The most appropriate comparison for LINE-1 DNA methylation levels was with the Zhu et al. paper which was a combined analysis of 5 studies. Zhu et

al.(94) provided descriptive information of LINE-1 DNA methylation levels measured using pyrosequencing in leukocyte DNA. The means for the 5 studies varied from 71.7% to 80.2% and the standard deviations were provided. The smallest standard deviation was 1.0 and the largest standard deviation was 6.2 (94). Also, methylation measurements were completed in duplicates and the within sample (intra-assay) coefficient of variation was 5.3 % (94). In comparison, the mean DNA methylation level for this thesis project was slightly higher at 84.52% and the standard deviation (3.19) was in the middle of the range found by Zhu et al. However, the intra-assay coefficient of variation for this thesis was much smaller at 1.71%, indicating better agreement between triplicate values.

A strength of this study is the reliability of the High-Resolution Melt (HRM) method used to measure percent LINE-1 DNA methylation. The intra-assay variation had a coefficient of variation of 1.71% which implies precision with this method for quantifying LINE-1 DNA methylation. Consequently, this study would have a smaller degree of misclassification making it easier to be able to detect whether or not an association truly exists. In the Zhu et al. study the coefficient of variation of 5.3% is fairly large in comparison to the standard deviations. Unfortunately, large coefficients of variation are currently common in the literature. Thus, even if previous studies have not found a significant association between potential determinants and DNA methylation levels this may have resulted from the imprecision in DNA methylation measurement. This study is contributing to the existing literature by examining the relationship between BMI and percent LINE-1 DNA methylation using a more precise and reliable measure of LINE-1 DNA methylation.

5.1.2 Comparison to Literature – Relationship between BMI and LINE-1 DNA Methylation

Overall, this study did not find a relationship between BMI and LINE-1 DNA methylation levels, which is consistent with the existing literature. Five studies have examined the relationship between BMI and global DNA methylation using either a genome-wide measure or a measure of repetitive element methylation levels. Previous findings in the literature were inconsistent with study results ranging from a significant association between BMI (along with waist circumference and body fat percentage) and DNA hypomethylation (145) to a positive relationship between higher BMI and higher DNA methylation level (111). However, most studies did not find a significant relationship between BMI and DNA methylation levels (16,94,113)

These five previous studies included healthy subjects comparable to the study population of this thesis. However, different techniques were used to measure DNA methylation levels of different repetitive sequences in these studies. Three studies used pyrosequencing to measure methylation in LINE-1 DNA, one used LUMA to measure methylation in LINE-1 DNA, and one study used MethyLight to measure methylation in the Alu and Sat2 repetitive sequences (16,94,111,113,145). It is possible that findings may have varied based on the repetitive elements that were examined. The study that reported a positive relationship between BMI and DNA methylation levels measured DNA methylation levels of Alu repetitive sequences (111). It is possible that BMI may have a different relationship with DNA methylation dependent on the repetitive element studied. Overall, this study is consistent with the literature which has not found an association between BMI and DNA methylation levels in cross-sectional study designs.

The null findings in this study may be due to the fact that the specific LINE-1 DNA methylation region used did not represent the critical location of global DNA methylation detection. For this project only 8 CpG sites within the LINE-1 sequence were used to quantify percent methylation. A sufficient number of methylation CpG sites or the most relevant CpG sites may not have been included in the methylation measurement. Furthermore, LINE-1 repetitive sequences might not adequately represent global DNA methylation levels. Ideally, a study might include Alu and Sat2 repetitive sequences as well.

5.2 Methodological Considerations

The potential for measurement error, uncontrolled confounding, and bias to influence the study observations are addressed in this section. Both the exposure of interest, BMI, and the outcome, LINE-1 DNA methylation levels, will be discussed. Overall, the exposure and outcome are susceptible to the effects of non-differential misclassification which would result in the effect estimates being biased toward the null.

5.2.1 Study Population

The prevalence of obesity in the study sample was lower than the prevalence of obesity in the general Canadian population. Since the study participants were healthy volunteers recruited from three university or hospital sites, who reported no history of diseases such as diabetes, it was expected that the prevalence of obesity in the study sample would be lower than the general population. In this study where participants were aged 20 to 50 the prevalence of obesity was 15% which is lower than the prevalence estimated based on the Canadian Health Measures Survey, where the prevalence of obesity was approximately 20% for those 20 to 39 years of age (21). Volunteers were likely to be healthier than the

general population. Therefore it was not anticipated that the BMI distribution in the study population would be representative of the general Canadian population. Furthermore, the distribution of LINE-1 DNA methylation found in this study may not have represented the distribution of DNA methylation in the general population. However, since the underlying relationship being examined is a biologic relationship and subjects would not have prior knowledge of their LINE-1 DNA methylation level, it is unlikely that the volunteer nature of the participants would bias the results of the primary hypothesis.

However, it is recognized that a reduction in the variability of both BMI and LINE-1 DNA methylation due to the healthy population sample (in comparison with the general population) may have reduced the statistical power of this study to identify relationships at the extremes of these distributions.

5.2.2 Exposure Assessment

The larger study collected information regarding each participant's current height and weight in order to calculate BMI. It is unknown whether their BMI was stable over time and what the ideal time window to capture BMI would have been in relation to LINE-1 DNA methylation. However, it was assumed that BMI likely remained fairly stable and if changes had occurred it would most likely have resulted in random misclassification.

A primary concern with self-reported data, such as height and weight used to derive BMI, is its validity. Validation studies have shown that self-reported and measured height and weight are highly correlated (Spearman $r>0.9$). Also, self-reported height and weight have a high specificity and sensitivity when compared with measured BMI values (30). This study sample is comprised of a younger and healthier sample of adults for whom self-reported measures would be considered valid. However, there is the potential for non-

differential systematic bias to occur when self-reported height and weight are used to calculate BMI as a measure of body composition. It is likely that there would have been a systematic bias due to the tendency for individuals to under-report their weight. This would result in non-differential misclassification of exposure (BMI) since underreporting of weight is unlikely to be related to DNA methylation levels. The consequence of this misclassification would tend to bias the effect estimates towards the null.

Since BMI is an easily obtainable and widely used measure it is beneficial to examine the relationship of interest using BMI. Ideally, a better measure of body fat such as body fat percentage and even the type of body fat would permit a more rigorous investigation of the potential health consequences associated with being overweight or obese. Using BMI as a proxy measure of body composition is practical but does not necessarily have the sensitivity to prevent misclassification. This misclassification would result in the effect estimates between categories appearing more similar and may prevent observing a relationship when one truly exists.

5.2.3 Outcome Assessment

The method used to quantify DNA methylation levels was assessed to determine the reliability within individuals and between plates. With a coefficient of variation of 1.71% within subjects (intra-assay variation) and 0.87% for the internal between plates (inter-assay variation) this data represents a highly reproducible and precise assay to measure methylation levels. There is a small amount of random error in the measurement that is considerably lower than other studies have presented.

In the design and optimization of the HRM method used to quantify DNA methylation, there were attempts to ensure that the method was not biased. Tse et al.

developed the primers for a region that was a conserved sequence of 8 CpG sites within the LINE-1DNA sequence (150). This is an important feature to ensure that all products are the same length because in order that the change in the melt profile can be attributed to different base-pairings and not due to different sized products. Optimization of the annealing temperature permitted the primers to bind with no preference to complementary methylated or unmethylated template. The dilution assay of the starting quantity of DNA template showed that this did not influence the percent methylation level obtained. This permitted the acceptance of a small amount of variability in the starting DNA template. Furthermore, in order to ensure that this variability in the amount of DNA was not associated with methylation levels, Cp values were assessed as a potential determinant of methylation levels. Cp values were not significantly associated with methylation levels.

Samples of whole blood were used as starting material for the laboratory experiments. All blood samples were drawn between 8 and 10am after an overnight fast. This prevented any dietary habits immediately preceding blood draw from impacting on the DNA methylation measurement. Leukocyte DNA was extracted from the whole blood sample. Leukocyte DNA is composed of DNA from many different types of white blood cells. It is possible that not using one cell type, such as lymphocytes, or controlling for the ratio of different cell types may have contributed to misclassification. One study assessed the impact of white blood cell type on DNA methylation levels and found that a small portion of the methylation heterogeneity between individuals was attributable to the overall cell composition of leukocytes (73).

Since DNA methylation is an emerging field of research, the determinants of DNA methylation levels and factors that may alter DNA methylation levels are still largely

unknown. Therefore it is possible that there are unknown confounders of the relationship of interest that were not controlled for in the study.

5.3 Limitations

There are several limitations to this study which may have impacted on the validity of this investigation of the relationship between BMI and LINE-1 DNA methylation. These include the underlying study design, available biologic medium, validity of the DNA methylation measure, and uncontrolled confounding.

An inherent limitation is based on the cross-sectional study design. With a cross-sectional study design temporality could not be assessed. However, the study design was appropriate in this context for a number of reasons. First, there is no evidence to suggest that LINE-1 DNA methylation levels would influence BMI. Second, an individual's BMI is suggested to remain fairly stable over a number of years (169,170) and tends to increase gradually in adulthood. Since this is one of the first studies of this nature it is appropriate to determine whether or not a relationship exists prior to investigating the association in a prospective manner.

The accessibility of blood makes leukocytes an ideal location to detect an association. Leukocytes are composed of all types of white blood cells. It is possible that there is misclassification of DNA methylation levels based on differences in the proportion of white blood cell types comprising the blood sample between individuals. Further, it is unknown whether LINE-1 DNA methylation levels in leukocytes would be representative of methylation status of other body tissues, such as the colon. Although a significant relationship between BMI and LINE-1 DNA methylation levels in leukocytes did not exist in

this study population, it is possible that BMI would be a predictor of DNA methylation levels in other tissues or white blood cells.

The determinants of DNA methylation levels are still largely unknown. There is the potential for uncontrolled confounding if this study did not account for unknown strong predictors of DNA methylation levels.

DNA methylation measurement is a new and rapidly expanding area of research. Although the technique used to measure DNA methylation (HRM) was reliable, the validity of the method used for this study was not tested against a gold standard. Also, DNA methylation is not well understood as a pathologic entity. It is unknown what would constitute a meaningful change in DNA methylation and what would be the most appropriate categorization or statistical analysis to examine methylation levels.

There was a moderate amount of variability in LINE-1 DNA methylation levels within this volunteer sample. However, given that these were all healthy subjects, it may be that a general population sample would have had a greater range of DNA methylation values which would have provided a more robust analysis of this thesis topic.

5.4 Strengths

The objective of the larger study was to examine lifestyle determinants of the One-Carbon Metabolism Cycle. A strength of this thesis study was that it included the assessment of many potential determinants of DNA methylation and that materials were collected in a thorough manner to assess methylation capacity.

Epigenetics is an emerging field of research, in which there is much interest in DNA methylation in order to understand biologic mechanisms regarding how environmental

exposures influence disease etiology. This study will contribute to the current literature that consists of only a few studies examining determinants of DNA methylation levels.

A benefit of investigating a biologic mechanism is that the response rate and random exclusion of participants from the larger study are not important given that selection bias was not a primary concern because individuals were unaware of their DNA methylation levels. For selection bias to occur in this study, a subject's decision to participate in the study would have to be related to both exposure and outcome status. This is not possible because DNA methylation levels were unknown to potential participants.

Traditional epidemiologic studies are limited in their ability to identify small relative risks associated with environmental and lifestyle exposures, and disease outcomes, particularly diseases with long latencies between exposure and outcome. The examination of an intermediate event such as global DNA methylation, that follows an exposure and precedes an outcome on a causal biologic pathway, provides several critical advantages in the investigation of exposure-outcome relationships. These include: a shorter latency period between exposure and the intermediate event (as opposed to the health outcome), an event that may be more frequent than the outcome, a study population of relatively healthy subjects, and a continuous outcome measure providing greater statistical power (171). Thus, if a causal exposure-outcome relationship exists that is mediated through an intermediate endpoint, then the observed relationship will be stronger between the exposure-intermediate as compared to the association observed for a corresponding exposure-outcome relationship (171).

5.5 Generalizability

A benefit of studying biologic mechanisms is that the findings are generalizable to similar populations. If this was a valid study, then the relationship would likely hold in most populations with a similar range of ages and BMI values. One potential challenge to the generalizability of these findings is that there may be an interaction with a factor that differs across populations, for example ethnicity. In terms of varying ethnic populations, a different relationship may exist between BMI and DNA methylation due to the influence of genetic factors on this relationship.

5.6 Conclusion

Biologic mechanisms linking obesity to increased incidence of adverse health events, such as cancer and CVD, are not well understood. The aim of this study was to investigate whether DNA methylation level changes may be a mechanism through which this increased risk was conferred. The research is relevant because changes in DNA methylation levels are recognized as a component in the etiology of various cancers and coronary artery disease. A significant relationship between BMI and LINE-1 DNA methylation levels was not observed in this study using either linear or logistic regression. The sensitivity analysis for DNA hypomethylation showed a pattern of decreasing odds ratios in females and a pattern of increasing odds ratios in males with increasing BMI categories. However, the interaction variable did not reach statistical significance. The sensitivity analysis for DNA hypermethylation showed a pattern of increasing odds ratios in females and no pattern in the odds ratios for males with increasing BMI categories. Overall, further investigation is warranted to examine if BMI as a proxy for adipose tissue acts as a determinant of DNA methylation levels and whether sex is an effect modifier of the relationship.

5.7 Contributions and Implications

Contributions of this thesis research include the addition of LINE-1 DNA methylation measurements to the existing database for the larger study. Further analyses can be undertaken to examine genetic or molecular factors in relation to the LINE-1 DNA methylation values obtained.

This thesis provides insight into a potential determinant of DNA methylation (i.e., BMI) which may mediate the risk for adverse outcomes. DNA methylation is an emerging field of research and unraveling the environmental determinants will assist in understanding underlying biologic mechanism for disease occurrence. Using a novel method to measure DNA methylation, this study was able to capture variability in LINE-1 DNA methylation levels between individuals. This research will contribute to the existing scientific literature by providing a description of LINE-1 DNA methylation levels, quantified using a reliable and precise method.

Very few studies have examined lifestyle factors as determinants of methylation levels. This study provided information regarding the relationship between BMI and global DNA methylation levels, through the measurement of LINE-1 DNA methylation levels, which can be applied to understanding determinants of DNA methylation levels. Since DNA methylation is modifiable, eventually, clinical applications may be possible to prevent or reverse changes in DNA methylation levels. Additionally, if changes in methylation levels are detectable in blood cells, they may represent a potential biomarker to identify individuals at high risk for certain diseases.

5.8 Future Directions

Methods to accurately measure global and gene-specific DNA methylation are evolving, which will allow epidemiologic studies to investigate determinants of changes in DNA methylation levels both globally and at gene-specific locations. Additionally, changes in DNA methylation should be examined longitudinally in order to examine the extent to which DNA methylation levels are modifiable and hereditary. Global DNA methylation may have potential predictive and prognostic capability. However, much more research is required to understand the biologic mechanisms and utility of DNA methylation levels in these settings. For example, examination of the correlations and patterns of DNA methylation levels of various organs, tissues, and progenitor cells would assist in the comprehension of cellular differentiation and when cells might be susceptible to environmental influences.

Global DNA methylation is one component of an epigenetic profile which interacts with an individual's genetic profile and environmental exposures. Histone type and acetylation levels interplay with DNA methylation in the conformation of chromosomal structure. Gene-specific DNA methylation acts in the regulation of gene expression that may interact with polymorphisms. Basic science and epidemiological research in the field of epigenetics needs to continue to begin to unravel the complexity of disease etiology.

References

(1) Obesity: preventing and managing the global epidemic. Report of a WHO consultation. World Health Organ Tech Rep Ser 2000;894:i-xii, 1-253.

(2) Physical status: the use and interpretation of anthropometry. Report of a WHO Expert Committee. World Health Organ Tech Rep Ser 1995;854:1-452.

(3) Calle EE, Kaaks R. Overweight, obesity and cancer: epidemiological evidence and proposed mechanisms. Nat Rev Cancer 2004 Aug;4(8):579-591.

(4) Statistics Canada. Table 105-0507 - Measured adult body mass index (BMI), by age group and sex, household population aged 18 and over excluding pregnant females, Canada (excluding territories), occasional (number unless otherwise noted), CANSIM (database). 2009; Available at: http://cansim2.statcan.gc.ca/cgi-win/cnsmcgi.exe?Lang=E&CNSM-Fi=CII/CII_1-eng.htm. Accessed September, 2010.

(5) Wilson AS, Power BE, Molloy PL. DNA hypomethylation and human diseases. Biochim Biophys Acta 2007 Jan;1775(1):138-162.

(6) Jakubowski H. Pathophysiological consequences of homocysteine excess. J Nutr 2006 Jun;136(6 Suppl):1741S-1749S.

(7) Duthie SJ. Folate and cancer: how DNA damage, repair and methylation impact on colon carcinogenesis. J Inherit Metab Dis 2010 Jun 11.

(8) Giovannucci E. Alcohol, one-carbon metabolism, and colorectal cancer: recent insights from molecular studies. J Nutr 2004 Sep;134(9):2475S-2481S.

(9) Niculescu MD, Zeisel SH. Diet, methyl donors and DNA methylation: interactions between dietary folate, methionine and choline. J Nutr 2002 Aug;132(8 Suppl):2333S-2335S.

(10) Dunn BK. Hypomethylation: one side of a larger picture. Ann N Y Acad Sci 2003 Mar;983:28-42.

(11) Singal R, Ginder GD. DNA methylation. Blood 1999 Jun 15;93(12):4059-4070.

(12) Weisenberger DJ, Campan M, Long TI, Kim M, Woods C, Fiala E, et al. Analysis of repetitive element DNA methylation by MethyLight. Nucleic Acids Res 2005 Dec 2;33(21):6823-6836.

(13) John BJ, Irukulla S, Abulafi AM, Kumar D, Mendall MA. Systematic review: adipose tissue, obesity and gastrointestinal diseases. Aliment Pharmacol Ther 2006 Jun 1;23(11):1511-1523.

(14) Trayhurn P, Wood IS. Adipokines: inflammation and the pleiotropic role of white adipose tissue. Br J Nutr 2004 Sep;92(3):347-355.

(15) Valinluck V, Sowers LC. Inflammation-mediated cytosine damage: a mechanistic link between inflammation and the epigenetic alterations in human cancers. Cancer Res 2007 Jun 15;67(12):5583-5586.

(16) Stenvinkel P, Karimi M, Johansson S, Axelsson J, Suliman M, Lindholm B, et al. Impact of inflammation on epigenetic DNA methylation - a novel risk factor for cardiovascular disease J Intern Med 2007 May;261(5):488-499.

(17) Franco R, Schoneveld O, Georgakilas AG, Panayiotidis MI. Oxidative stress, DNA methylation and carcinogenesis. Cancer Lett 2008 Jul 18;266(1):6-11.

(18) Revicki DA, Israel RG. Relationship between body mass indices and measures of body adiposity. Am J Public Health 1986 Aug;76(8):992-994.

(19) Garrow JS, Webster J. Quetelet's index (W/H2) as a measure of fatness. Int J Obes 1985;9(2):147-153.

(20) National Institutes of Health. Clinical guidelines on the indentification, evaluation, and treatment of overweight and obesity in adults: The Evidence Report. Obes 1998(6).

(21) Shields M, Tjepkema M. Trends in adult obesity. 2006;Vol. 17 Number 3:53.

(22) Shields M, Tremblay M, Laviolette M, Craig CL, Janssen I, Gorber SC. Fitness of Canadian adults: Results from the 2007-2009 Canadian Health Measures Survey. 2010.

(23) Deurenberg P, Weststrate JA, Seidell JC. Body mass index as a measure of body fatness: age- and sex-specific prediction formulas. Br J Nutr 1991 Mar;65(2):105-114.

(24) Wells JC, Treleaven P, Cole TJ. BMI compared with 3-dimensional body shape: the UK National Sizing Survey. Am J Clin Nutr 2007 Feb;85(2):419-425.

(25) Borkan GA, Hults DE, Gerzof SG, Robbins AH, Silbert CK. Age changes in body composition revealed by computed tomography. J Gerontol 1983 Nov;38(6):673-677.

(26) Romero-Corral A, Somers VK, Sierra-Johnson J, Thomas RJ, Collazo-Clavell ML, Korinek J, et al. Accuracy of body mass index in diagnosing obesity in the adult general population. Int J Obes (Lond) 2008 Jun;32(6):959-966.

(27) World Cancer Research Fund. Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Prespective. 2007.

(28) Willett W. Nutritional Epidemiology. Second Edition ed.: Oxford University Press; 1998.

(29) Kuczmarski MF, Kuczmarski RJ, Najjar M. Effects of age on validity of self-reported height, weight, and body mass index: findings from the Third National Health and Nutrition Examination Survey, 1988-1994. J Am Diet Assoc 2001 Jan;101(1):28-34; quiz 35-6.

(30) Spencer EA, Appleby PN, Davey GK, Key TJ. Validity of self-reported height and weight in 4808 EPIC-Oxford participants. Public Health Nutr 2002 Aug;5(4):561-565.

(31) Nyholm M, Gullberg B, Merlo J, Lundqvist-Persson C, Rastam L, Lindblad U. The validity of obesity based on self-reported weight and height: Implications for population studies. Obesity (Silver Spring) 2007 Jan;15(1):197-208.

(32) Kershaw EE, Flier JS. Adipose tissue as an endocrine organ. J Clin Endocrinol Metab 2004 Jun;89(6):2548-2556.

(33) Trayhurn P, Bing C, Wood IS. Adipose tissue and adipokines--energy regulation from the human perspective. J Nutr 2006 Jul;136(7 Suppl):1935S-1939S.

(34) Hopfl G, Ogunshola O, Gassmann M. HIFs and tumors--causes and consequences. Am J Physiol Regul Integr Comp Physiol 2004 Apr;286(4):R608-23.

(35) Renehan AG, Tyson M, Egger M, Heller RF, Zwahlen M. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. Lancet 2008 Feb 16;371(9612):569-578.

(36) Frezza EE, Wachtel MS, Chiriva-Internati M. Influence of obesity on the risk of developing colon cancer. Gut 2006 Feb;55(2):285-291.

(37) Stein CJ, Colditz GA. Modifiable risk factors for cancer. Br J Cancer 2004 Jan 26;90(2):299-303.

(38) Hill MJ. Physical activity and human cancer. Eur J Cancer Prev 1999 Dec;8(6):475-477.

(39) Hill MJ. Diet, physical activity and cancer risk. Public Health Nutr 1999 Sep;2(3A):397-401.

(40) Fehily AM, Phillips KM, Yarnell JW. Diet, smoking, social class, and body mass index in the Caerphilly Heart Disease Study. Am J Clin Nutr 1984 Oct;40(4):827-833.

(41) Tavani A, Negri E, La Vecchia C. Determinants of body mass index: a study from northern Italy. Int J Obes Relat Metab Disord 1994 Jul;18(7):497-502.

(42) Lukasiewicz E, Mennen LI, Bertrais S, Arnault N, Preziosi P, Galan P, et al. Alcohol intake in relation to body mass index and waist-to-hip ratio: the importance of type of alcoholic beverage. Public Health Nutr 2005 May;8(3):315-320.

(43) Gruchow HW, Sobocinski KA, Barboriak JJ, Scheller JG. Alcohol consumption, nutrient intake and relative body weight among US adults. Am J Clin Nutr 1985 Aug;42(2):289-295.

(44) Slattery ML, McDonald A, Bild DE, Caan BJ, Hilner JE, Jacobs DR,Jr, et al. Associations of body fat and its distribution with dietary intake, physical activity, alcohol, and smoking in blacks and whites. Am J Clin Nutr 1992 May;55(5):943-949.

(45) Kuk JL, Lee S, Heymsfield SB, Ross R. Waist circumference and abdominal adipose tissue distribution: influence of age and sex. Am J Clin Nutr 2005 Jun;81(6):1330-1334.

(46) de Onis M, Habicht JP. Anthropometric reference data for international use: recommendations from a World Health Organization Expert Committee. Am J Clin Nutr 1996 Oct;64(4):650-658.

(47) Lahti-Koski M, Vartiainen E, Mannisto S, Pietinen P. Age, education and occupation as determinants of trends in body mass index in Finland from 1982 to 1997. Int J Obes Relat Metab Disord 2000 Dec;24(12):1669-1676.

(48) Sarlio-Lahteenkorva S, Lahelma E. The association of body mass index with social and economic disadvantage in women and men. Int J Epidemiol 1999 Jun;28(3):445-449.

(49) Baum CL,2nd, Ruhm CJ. Age, socioeconomic status and obesity growth. J Health Econ 2009 May;28(3):635-648.

(50) Calle EE, Rodriguez C, Walker-Thurmond K, Thun MJ. Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. N Engl J Med 2003 Apr 24;348(17):1625-1638.

(51) Health Canada. Canadian Guidelines for Body Weight Classification in Adults. 2003.

(52) McMillan DC, Sattar N, McArdle CS. ABC of obesity. Obesity and cancer. BMJ 2006 Nov 25;333(7578):1109-1111.

(53) Eckel RH, Krauss RM. American Heart Association call to action: obesity as a major risk factor for coronary heart disease. AHA Nutrition Committee. Circulation 1998 Jun 2;97(21):2099-2100.

(54) Sowers JR. Obesity as a cardiovascular risk factor. Am J Med 2003 Dec 8;115 Suppl 8A:37S-41S.

(55) Grundy SM. Obesity, metabolic syndrome, and coronary atherosclerosis. Circulation 2002 Jun 11;105(23):2696-2698.

(56) Rimm EB, Stampfer MJ, Giovannucci E, Ascherio A, Spiegelman D, Colditz GA, et al. Body size and fat distribution as predictors of coronary heart disease among middle-aged and older US men. Am J Epidemiol 1995 Jun 15;141(12):1117-1127.

(57) Cooper R, Cutler J, Desvigne-Nickens P, Fortmann SP, Friedman L, Havlik R, et al. Trends and disparities in coronary heart disease, stroke, and other cardiovascular diseases in the United States: findings of the national conference on cardiovascular disease prevention. Circulation 2000 Dec 19;102(25):3137-3147.

(58) Gensini GF, Comeglio M, Colella A. Classical risk factors and emerging elements in the risk profile for coronary artery disease. Eur Heart J 1998 Feb;19 Suppl A:A53-61.

(59) Christman JK. Separation of major and minor deoxyribonucleoside monophosphates by reverse-phase high-performance liquid chromatography: a simple method applicable to quantitation of methylated nucleotides in DNA. Anal Biochem 1982 Jan 1;119(1):38-48.

(60) Chiang PK, Gordon RK, Tal J, Zeng GC, Doctor BP, Pardhasaradhi K, et al. S-Adenosylmethionine and methylation. FASEB J 1996 Mar;10(4):471-480.

(61) Alberts B, Johnson A, Lewis J, Roberts K, Walter P editors. Molecular Biology of The Cell. Fourth Edition ed.: Garland Science; 2002.

(62) Yi P, Melnyk S, Pogribna M, Pogribny IP, Hine RJ, James SJ. Increase in plasma homocysteine associated with parallel increases in plasma S-adenosylhomocysteine and lymphocyte DNA hypomethylation. J Biol Chem 2000 Sep 22;275(38):29318-29323.

(63) Friso S, Choi SW, Girelli D, Mason JB, Dolnikowski GG, Bagley PJ, et al. A common mutation in the 5,10-methylenetetrahydrofolate reductase gene affects genomic DNA methylation through an interaction with folate status. Proc Natl Acad Sci U S A 2002 Apr 16;99(8):5606-5611.

(64) Geisel J, Schorr H, Bodis M, Isber S, Hubner U, Knapp JP, et al. The vegetarian lifestyle and DNA methylation. Clin Chem Lab Med 2005;43(10):1164-1169.

(65) Fowler B. Disorders of homocysteine metabolism. J Inherit Metab Dis 1997 Jun;20(2):270-285.

(66) Finkelstein JD, Martin JJ. Homocysteine. Int J Biochem Cell Biol 2000 Apr;32(4):385-389.

(67) Friso S, Choi SW. Gene-nutrient interactions and DNA methylation. J Nutr 2002 Aug;132(8 Suppl):2382S-2387S.

(68) Stern LL, Mason JB, Selhub J, Choi SW. Genomic DNA hypomethylation, a characteristic of most cancers, is present in peripheral leukocytes of individuals who are homozygous for the C677T polymorphism in the methylenetetrahydrofolate reductase gene. Cancer Epidemiol Biomarkers Prev 2000 Aug;9(8):849-853.

(69) Takai D, Jones PA. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. Proc Natl Acad Sci U S A 2002 Mar 19;99(6):3740-3745.

(70) Bird A. DNA methylation patterns and epigenetic memory. Genes Dev 2002 Jan 1;16(1):6-21.

(71) Hansen KD, Timp W, Bravo HC, Sabunciyan S, Langmead B, McDonald OG, et al. Increased methylation variation in epigenetic domains across cancer types. Nat Genet 2011 Jun 26.

(72) Taberlay PC, Jones PA. DNA methylation and cancer. Prog Drug Res 2011;67:1-23.

(73) Talens RP, Boomsma DI, Tobi EW, Kremer D, Jukema JW, Willemsen G, et al. Variation, patterns, and temporal stability of DNA methylation: considerations for epigenetic epidemiology. FASEB J 2010 May 14.

(74) Baccarelli A, Tarantini L, Wright RO, Bollati V, Litonjua AA, Zanobetti A, et al. Repetitive element DNA methylation and circulating endothelial and inflammation markers in the VA normative aging study. Epigenetics 2010 Apr 1;5(3).

(75) Rocha MS, Castro R, Rivera I, Kok RM, Smulders YM, Jakobs C, et al. Global DNA methylation: comparison of enzymatic- and non-enzymatic-based methods. Clin Chem Lab Med 2010 Dec;48(12):1793-1798.

(76) Friso S, Choi SW, Dolnikowski GG, Selhub J. A method to assess genomic DNA methylation using high-performance liquid chromatography/electrospray ionization mass spectrometry. Anal Chem 2002 Sep 1;74(17):4526-4531.

(77) Eads CA, Danenberg KD, Kawakami K, Saltz LB, Blake C, Shibata D, et al. MethyLight: a high-throughput assay to measure DNA methylation. Nucleic Acids Res 2000 Apr 15;28(8):E32.

(78) Wojdacz TK, Dobrovic A, Hansen LL. Methylation-sensitive high-resolution melting. Nat Protoc 2008;3(12):1903-1908.

(79) Vossen RH, Aten E, Roos A, den Dunnen JT. High-resolution melting analysis (HRMA): more than just sequence variant screening. Hum Mutat 2009 Jun;30(6):860-866.

(80) Stanzer S, Balic M, Strutz J, Heitzer E, Obermair F, Hauser-Kronberger C, et al. Rapid and reliable detection of LINE-1 hypomethylation using high-resolution melting analysis. Clin Biochem 2010 Dec;43(18):1443-1448.

(81) Cordaux R, Batzer MA. The impact of retrotransposons on human genome evolution. Nat Rev Genet 2009 Oct;10(10):691-703.

(82) Weiner AM. SINEs and LINEs: the art of biting the hand that feeds you. Curr Opin Cell Biol 2002 Jun;14(3):343-350.

(83) Deininger PL, Moran JV, Batzer MA, Kazazian HH,Jr. Mobile elements and mammalian genome evolution. Curr Opin Genet Dev 2003 Dec;13(6):651-658.

(84) Ehrlich M. DNA methylation in cancer: too much, but also too little. Oncogene 2002 Aug 12;21(35):5400-5413.

(85) Kazazian HH,Jr. Mobile elements: drivers of genome evolution. Science 2004 Mar 12;303(5664):1626-1632.

(86) Ehrlich M, Gama-Sosa MA, Huang LH, Midgett RM, Kuo KC, McCune RA, et al. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. Nucleic Acids Res 1982 Apr 24;10(8):2709-2721.

(87) Pufulete M, Al-Ghnaniem R, Leather AJ, Appleby P, Gout S, Terry C, et al. Folate status, genomic DNA hypomethylation, and risk of colorectal adenoma and cancer: a case control study. Gastroenterology 2003 May;124(5):1240-1248.

(88) Choi JY, James SR, Link PA, McCann SE, Hong CC, Davis W, et al. Association between global DNA hypomethylation in leukocytes and risk of breast cancer. Carcinogenesis 2009 Nov;30(11):1889-1897.

(89) Lim U, Flood A, Choi SW, Albanes D, Cross AJ, Schatzkin A, et al. Genomic methylation of leukocyte DNA in relation to colorectal adenoma among asymptomatic women. Gastroenterology 2008 Jan;134(1):47-55.

(90) Wilhelm CS, Kelsey KT, Butler R, Plaza S, Gagne L, Zens MS, et al. Implications of LINE1 methylation for bladder cancer risk in women. Clin Cancer Res 2010 Mar 1;16(5):1682-1689.

(91) Chalitchagorn K, Shuangshoti S, Hourpai N, Kongruttanachok N, Tangkijvanich P, Thong-ngam D, et al. Distinctive pattern of LINE-1 methylation level in normal tissues and the association with carcinogenesis. Oncogene 2004 Nov 18;23(54):8841-8846.

(92) Hou L, Wang H, Sartori S, Gawron A, Lissowska J, Bollati V, et al. Blood leukocyte DNA hypomethylation and gastric cancer risk in a high-risk Polish population. Int J Cancer 2010 Oct 15;127(8):1866-1874.

(93) Terry MB, Delgado-Cruzata L, Vin-Raviv N, Wu HC, Santella RM. DNA methylation in white blood cells: Association with risk factors in epidemiologic studies. Epigenetics 2011 Jul 1;6(7).

(94) Zhu ZZ, Hou L, Bollati V, Tarantini L, Marinelli B, Cantone L, et al. Predictors of global methylation levels in blood DNA of healthy subjects: a combined analysis. Int J Epidemiol 2010 Sep 15.

(95) Ross JP, Rand KN, Molloy PL. Hypomethylation of repeated DNA sequences in cancer. Epigenomics 2010 Apr;2(2):245-269.

(96) El-Maarri O, Becker T, Junen J, Manzoor SS, Diaz-Lacava A, Schwaab R, et al. Gender specific differences in levels of DNA methylation at selected loci from human total blood: a tendency toward higher methylation levels in males. Hum Genet 2007 Dec;122(5):505-514.

(97) Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML, et al. Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci U S A 2005 Jul 26;102(30):10604-10609.

(98) Kok RM, Smith DE, Barto R, Spijkerman AM, Teerlink T, Gellekink HJ, et al. Global DNA methylation measured by liquid chromatography-tandem mass spectrometry: analytical technique, reference values and determinants in healthy subjects. Clin Chem Lab Med 2007;45(7):903-911.

(99) Jiricny J, Menigatti M. DNA Cytosine demethylation: are we getting close? Cell 2008 Dec 26;135(7):1167-1169.

(100) Ooi SK, Bestor TH. The colorful history of active DNA demethylation. Cell 2008 Jun 27;133(7):1145-1148.

(101) Tost J. DNA methylation: an introduction to the biology and the disease-associated changes of a promising biomarker. Methods Mol Biol 2009;507:3-20.

(102) Brena RM, Huang TH, Plass C. Quantitative assessment of DNA methylation: Potential applications for disease diagnosis, classification, and prognosis in clinical settings. J Mol Med 2006 May;84(5):365-377.

(103) Chen RZ, Pettersson U, Beard C, Jackson-Grusby L, Jaenisch R. DNA hypomethylation leads to elevated mutation rates. Nature 1998 Sep 3;395(6697):89-93.

(104) Perry AS, Watson RW, Lawler M, Hollywood D. The epigenome as a therapeutic target in prostate cancer. Nat Rev Urol 2010 Dec;7(12):668-680.

(105) Ng HH, Bird A. DNA methylation and chromatin modification. Curr Opin Genet Dev 1999 Apr;9(2):158-163.

(106) Cheng X, Blumenthal RM. Coordinated chromatin control: structural and functional linkage of DNA and histone methylation. Biochemistry 2010 Apr 13;49(14):2999-3008.

(107) Bollati V, Schwartz J, Wright R, Litonjua A, Tarantini L, Suh H, et al. Decline in genomic DNA methylation through aging in a cohort of elderly subjects. Mech Ageing Dev 2009 Apr;130(4):234-239.

(108) Pilsner JR, Liu X, Ahsan H, Ilievski V, Slavkovich V, Levy D, et al. Genomic methylation of peripheral blood leukocyte DNA: influences of arsenic and folate in Bangladeshi adults. Am J Clin Nutr 2007 Oct;86(4):1179-1186.

(109) Bjornsson HT, Sigurdsson MI, Fallin MD, Irizarry RA, Aspelund T, Cui H, et al. Intra-individual change over time in DNA methylation with familial clustering. JAMA 2008 Jun 25;299(24):2877-2883.

(110) Zhang FF, Morabia A, Carroll J, Gonzalez K, Fulda K, Kaur M, et al. Dietary patterns are associated with levels of global genomic DNA methylation in a cancer-free population. J Nutr 2011 Jun;141(6):1165-1171.

(111) Kim M, Long TI, Arakawa K, Wang R, Yu MC, Laird PW. DNA methylation as a biomarker for cardiovascular disease risk. PLoS One 2010 Mar 15;5(3):e9692.

(112) Kim KC, Friso S, Choi SW. DNA methylation, an epigenetic mechanism connecting folate to healthy embryonic development and aging. J Nutr Biochem 2009 Dec;20(12):917-926.

(113) Zhang FF, Cardarelli R, Carroll J, Fulda KG, Kaur M, Gonzalez K, et al. Significant differences in global genomic DNA methylation by gender and race/ethnicity in peripheral blood. Epigenetics 2011 May;6(5):623-629.

(114) Figueiredo JC, Grau MV, Wallace K, Levine AJ, Shen L, Hamdan R, et al. Global DNA hypomethylation (LINE-1) in the normal colon and lifestyle characteristics and dietary and genetic factors. Cancer Epidemiol Biomarkers Prev 2009 Apr;18(4):1041-1049.

(115) Bonsch D, Lenz B, Reulbach U, Kornhuber J, Bleich S. Homocysteine associated genomic DNA hypermethylation in patients with chronic alcoholism. J Neural Transm 2004 Dec;111(12):1611-1616.

(116) Ushijima T, Okochi-Takada E. Aberrant methylations in cancer cells: where do they come from? Cancer Sci 2005 Apr;96(4):206-211.

(117) Detich N, Hamm S, Just G, Knox JD, Szyf M. The methyl donor S-Adenosylmethionine inhibits active demethylation of DNA: a candidate novel mechanism for the pharmacological effects of S-Adenosylmethionine. J Biol Chem 2003 Jun 6;278(23):20812-20820.

(118) Pufulete M, Emery PW, Sanders TA. Folate, DNA methylation and colo-rectal cancer. Proc Nutr Soc 2003 May;62(2):437-445.

(119) Finkelstein JD. Pathways and regulation of homocysteine metabolism in mammals. Semin Thromb Hemost 2000;26(3):219-225.

(120) Basten GP, Duthie SJ, Pirie L, Vaughan N, Hill MH, Powers HJ. Sensitivity of markers of DNA stability and DNA repair activity to folate supplementation in healthy volunteers. Br J Cancer 2006 Jun 19;94(12):1942-1947.

(121) Ordovas JM, Smith CE. Epigenetics and cardiovascular disease. Nat Rev Cardiol 2010 Sep;7(9):510-519.

(122) Laird PW, Jaenisch R. The role of DNA methylation in cancer genetic and epigenetics. Annu Rev Genet 1996;30:441-464.

(123) Jones PA, Laird PW. Cancer epigenetics comes of age. Nat Genet 1999 Feb;21(2):163-167.

(124) Feinberg AP. An epigenetic approach to cancer etiology. Cancer J 2007 Jan-Feb;13(1):70-74.

(125) Das PM, Singal R. DNA methylation and cancer. J Clin Oncol 2004 Nov 15;22(22):4632-4642.

(126) Matsuzaki K, Deng G, Tanaka H, Kakar S, Miura S, Kim YS. The relationship between global methylation level, loss of heterozygosity, and microsatellite instability in sporadic colorectal cancer. Clin Cancer Res 2005 Dec 15;11(24 Pt 1):8564-8569.

(127) Kitkumthorm N, Mutirangura A. Long interspersed nuclear element-1 hypomethylation in cancer: biology and clinical applications. Clinical Epigenetics 2011;2(2):315.

(128) Feinberg AP. Alterations in DNA methylation in colorectal polyps and cancer. Prog Clin Biol Res 1988;279:309-317.

(129) Cravo M, Fidalgo P, Pereira AD, Gouveia-Oliveira A, Chaves P, Selhub J, et al. DNA methylation as an intermediate biomarker in colorectal cancer: modulation by folic acid supplementation. Eur J Cancer Prev 1994 Nov;3(6):473-479.

(130) Pufulete M, Al-Ghnaniem R, Rennie JA, Appleby P, Harris N, Gout S, et al. Influence of folate status on genomic DNA methylation in colonic mucosa of subjects without colorectal adenoma or cancer. Br J Cancer 2005 Mar 14;92(5):838-842.

(131) Zaina S, Lindholm MW, Lund G. Nutrition and aberrant DNA methylation patterns in atherosclerosis: more than just hyperhomocysteinemia? J Nutr 2005 Jan;135(1):5-8.

(132) Lund G, Andersson L, Lauria M, Lindholm M, Fraga MF, Villar-Garea A, et al. DNA methylation polymorphisms precede any histological sign of atherosclerosis in mice lacking apolipoprotein E. J Biol Chem 2004 Jul 9;279(28):29147-29154.

(133) Dong C, Yoon W, Goldschmidt-Clermont PJ. DNA methylation and atherosclerosis. J Nutr 2002 Aug;132(8 Suppl):2406S-2409S.

(134) Cerda S, Weitzman SA. Influence of oxygen radical injury on DNA methylation. Mutat Res 1997 Apr;386(2):141-152.

(135) Weitzman SA, Turk PW, Milkowski DH, Kozlowski K. Free radical adducts induce alterations in DNA cytosine methylation. Proc Natl Acad Sci U S A 1994 Feb 15;91(4):1261-1264.

(136) Zuo S, Boorstein RJ, Teebor GW. Oxidative damage to 5-methylcytosine in DNA. Nucleic Acids Res 1995 Aug 25;23(16):3239-3243.

(137) Shen L, Ahuja N, Shen Y, Habib NA, Toyota M, Rashid A, et al. DNA methylation and environmental exposures in human hepatocellular carcinoma. J Natl Cancer Inst 2002 May 15;94(10):755-761.

(138) Hodge DR, Peng B, Cherry JC, Hurt EM, Fox SD, Kelley JA, et al. Interleukin 6 supports the maintenance of p53 tumor suppressor gene promoter methylation. Cancer Res 2005 Jun 1;65(11):4673-4682.

(139) Henderson JP, Byun J, Williams MV, Mueller DM, McCormick ML, Heinecke JW. Production of brominating intermediates by myeloperoxidase. A transhalogenation pathway for generating mutagenic nucleobases during inflammation. J Biol Chem 2001 Mar 16;276(11):7867-7875.

(140) Badouard C, Masuda M, Nishino H, Cadet J, Favier A, Ravanat JL. Detection of chlorinated DNA and RNA nucleosides by HPLC coupled to tandem mass spectrometry as potential biomarkers of inflammation. J Chromatogr B Analyt Technol Biomed Life Sci 2005 Nov 15;827(1):26-31.

(141) Valinluck V, Liu P, Kang JI,Jr, Burdzy A, Sowers LC. 5-halogenated pyrimidine lesions within a CpG sequence context mimic 5-methylcytosine by enhancing the binding of the methyl-CpG-binding domain of methyl-CpG-binding protein 2 (MeCP2). Nucleic Acids Res 2005 May 25;33(9):3057-3064.

(142) Hirsch S, Ronco AM, Guerrero-Bosagna C, de la Maza MP, Leiva L, Barrera G, et al. Methylation status in healthy subjects with normal and high serum folate concentration. Nutrition 2008 Nov-Dec;24(11-12):1103-1109.

(143) van Driel LM, Eijkemans MJ, de Jonge R, de Vries JH, van Meurs JB, Steegers EA, et al. Body mass index is an important determinant of methylation biomarkers in women of reproductive ages. J Nutr 2009 Dec;139(12):2315-2321.

(144) Poirier LA, Wise CK, Delongchamp RR, Sinha R. Blood determinations of S-adenosylmethionine, S-adenosylhomocysteine, and homocysteine: correlations with diet. Cancer Epidemiol Biomarkers Prev 2001 Jun;10(6):649-655.

(145) Piyathilake CJ, Badiga S, Johanning GL, Alvarez RD, Partridge EE. Predictors and health consequences of epigenetic changes associated with excess body weight in women of child-bearing age. Cancer Epidemiol.Biomarkers Prev. 2011;20(719).

(146) Zhang FF, Cardarelli R, Carroll J, Zhang S, Fulda KG, Gonzalez K, et al. Physical activity and global genomic DNA methylation in a cancer-free population. Epigenetics 2011 Mar 1;6(3):293-299.

(147) Feinberg AP, Irizarry RA, Fradin D, Aryee MJ, Murakami P, Aspelund T, et al. Personalized epigenomic signatures that are stable over time and covary with body mass index. Sci Transl Med 2010 Sep 15;2(49):49ra67.

(148) Godfrey KM, Sheppard A, Gluckman PD, Lillycrop KA, Burdge GC, McLean C, et al. Epigenetic gene promoter methylation at birth is associated with child's later adiposity. Diabetes 2011 May;60(5):1528-1534.

(149) Souren NY, Tierling S, Fryns JP, Derom C, Walter J, Zeegers MP. DNA methylation variability at growth-related imprints does not contribute to overweight in monozygotic twins discordant for BMI. Obesity (Silver Spring) 2011 Jul;19(7):1519-1522.

(150) Tse MY, Ashbury JE, Zwingerman N, King WD, Taylor S, Pang SC. A refined, rapid and reproducible high resolution melt (HRM) - based method suitable for quantification of global LINE-1 repetitive element methylation. BMC Research Notes 2011;4:565.

(151) Ainsworth BE, Leon AS, Richardson MT, Jacobs DR, Paffenbarger RS,Jr. Accuracy of the College Alumnus Physical Activity Questionnaire. J Clin Epidemiol 1993 Dec;46(12):1403-1411.

(152) Saw SM, Yuan JM, Ong CN, Arakawa K, Lee HP, Coetzee GA, et al. Genetic, dietary, and other lifestyle determinants of plasma homocysteine concentrations in middle-aged and older Chinese men and women in Singapore. Am J Clin Nutr 2001 Feb;73(2):232-239.

(153) Chrysohoou C, Panagiotakos DB, Pitsavos C, Zeimbekis A, Zampelas A, Papademetriou L, et al. The associations between smoking, physical activity, dietary habits and plasma homocysteine levels in cardiovascular disease-free people: the 'ATTICA' study. Vasc Med 2004 May;9(2):117-123.

(154) Allen NE, Beral V, Casabonne D, Kan SW, Reeves GK, Brown A, et al. Moderate alcohol intake and cancer incidence in women. J Natl Cancer Inst 2009 Mar 4;101(5):296-305.

(155) Roerecke M, Haydon E, Giesbrecht N. Alcohol and Chronic Disease: An Ontario Perspective. 2007.

(156) Walker MC, Smith GN, Perkins SL, Keely EJ, Garner PR. Changes in homocysteine levels during normal pregnancy. Am J Obstet Gynecol 1999 Mar;180(3 Pt 1):660-664.

(157) Miller M, Stone NJ, Ballantyne C, Bittner V, Criqui MH, Ginsberg HN, et al. Triglycerides and cardiovascular disease: a scientific statement from the American Heart Association. Circulation 2011 May 24;123(20):2292-2333.

(158) National Heart, Lung, and Blood Institute. High Blood Cholesterol. National Institute of Health .

(159) Zymo Research. Human Methylated and Non-methylated DNA Set: Product Information.

(160) Bland JM, Altman DG. Measurement error and correlation coefficients. BMJ 1996 Jul 6;313(7048):41-42.

(161) Budtz-Jorgensen E, Keiding N, Grandjean P, Weihe P. Confounder selection in environmental epidemiology: assessment of health effects of prenatal mercury exposure. Ann Epidemiol 2007 Jan;17(1):27-35.

(162) Vittinghoff E, Glidden DV, Shiboski SC, McCullech CE. Regression Methods in Biostatistics. 2008.

(163) Dupont WD, Plummer WD,Jr. Power and sample size calculations. A review and computer program. Control Clin Trials 1990 Apr;11(2):116-128.

(164) Donner A, Klar N. Pitfalls of and controversies in cluster randomization trials. Am J Public Health 2004 Mar;94(3):416-422.

(165) Klar N, Donner A. Current and future challenges in the design and analysis of cluster randomization trials. Stat Med 2001 Dec 30;20(24):3729-3740.

(166) Altman DG, Bland JM. Comparing several groups using analysis of variance. BMJ 1996 Jun 8;312(7044):1472-1473.

(167) Bland JM, Altman DG. Calculating correlation coefficients with repeated observations: Part 1--Correlation within subjects. BMJ 1995 Feb 18;310(6977):446.

(168) Bland JM, Altman DG. Calculating correlation coefficients with repeated observations: Part 2--Correlation between subjects. BMJ 1995 Mar 11;310(6980):633.

(169) Livshits G, Malkin I, Williams FM, Hart DJ, Hakim A, Spector TD. Longitudinal study of variation in body mass index in middle-aged UK females. Age (Dordr) 2011 Aug 19.

(170) Katzmarzyk PT, Perusse L, Malina RM, Bouchard C. Seven-year stability of indicators of obesity and adipose tissue distribution in the Canadian population. Am J Clin Nutr 1999 Jun;69(6):1123-1129.

(171) Merlo LM, Pepper JW, Reid BJ, Maley CC. Cancer as an evolutionary and ecological process. Nat Rev Cancer 2006 Dec;6(12):924-935.

Appendix A
DNA Isolation Protocol

This isolation protocol is for whole blood comprised samples. The isolation protocol was based on the 5PRIME ArchivePure DNA Purification System.

Sample Preparation:

1. Remove sample whole blood 250µL from freezer and thaw quickly. Once thawed, immediately place samples in ice until proceeding with DNA isolation.

Cell Lysis:

2. Add 750µL of Red Blood Cell (RBC) Lysis Solution to 250µL of whole blood in a 1.5ml microcentrifuge tube. Invert to mix and incubate 5 minutes at room temperature. Invert again at least once during incubation.
3. Centrifuge at 3,500 x g for 5 minutes.
4. Pour off supernatant leaving behind a visible brown pellet and approximately 100µL of residual liquid.
5. Vortex tube vigorously to resuspend the cells in the residual liquid.
6. Dispense 250µL of Cell Lysis Solution and 2µL of Protinase K into each tube of resuspended cells. Vortex at high speed for 10 seconds.
7. Incubate in a shaker at 56ºC overnight.

Protein Precipitation:

8. Cool sample by placing tube into an ice bath for 5 minutes.
9. Dispense 112.5µL of Protein Precipitation Solution into each tube.
10. Vortex samples vigorously at high speed for 20 seconds, in order to create a homogeneous solution.

11. Centrifuge at 10,000 x g for 5 minutes. The precipitated proteins will form a tight dark brown pellet.

DNA Precipitation:

12. Pour the supernatant containing the DNA (leaving behind the precipitated proteins) into a clean 1.5ml tube containing 250µL of 100% isopropanol.
13. Mix the sample by inverting gently in a tube rack 50 times.
14. Centrifuge at 14,000 x g for 5 minutes the DNA will be visible as a white pellet.
15. Pour off supernatant and drain tube briefly on clean absorbent paper.
16. Dispense 250µL of 70% ethanol into each tube and invert gently to wash the DNA pellet.
17. Centrifuge at 14,000 x g for 1 minute. Carefully pour off the enthanol.
18. Invert and drain the tube on clean absorbent paper and dry on heating block for 5 minutes.

DNA Hydration:

19. Add 60µL of Hydration Solution and gently vortex.
20. Rehydrate DNA by incubating at 56ºC overnight in a shaker. The tube is gently vortexed a couple of times to aid in the dispersing of the DNA.
21. Sample is centrifuges and then stored at -20ºC.

Appendix B

Bisulfite Conversion Protocol

This bisulfite conversion protocol is based on the sodium bisulfite conversion of unmethylated cystosines in DNA from the EpiTect Bisulfite Handbook from Qiagen.

Bisulfite Conversion:

1. To a 200µL PCR tube, add 20µL of DNA at 100ng/µL. Then add 85µL of Bisulfite Mix (stored at -20ºC) and then 35µL of DNA Protect Buffer (stored at 2-4ºC). Note: when the DNA Protect Buffer is added it should turn from green to blue.
2. Close the tubes and centrifuge the reactions thoroughly.
3. Place tubes (maximum 20) in thermal cycle and use program 'bisulfite'. This program is outlined in the following table:

Step	Time	Temperature
Denaturation	5 minutes	95ºC
Incubation	25 minutes	60ºC
Denaturation	5 minutes	95ºC
Incubation	85 minutes	60ºC
Denaturation	5 minutes	95ºC
Incubation	175 minutes	60ºC
Hold	Indefinite	20ºC

Cleanup of converted DNA:

4. Briefly centrifuge the tubes and transfer the complete bisulfite reaction into a 1.5ml microcentrifuge tube.

5. Add 560µL of BL Buffer to each sample. Vortex for 10 seconds and then centrifuge at 14,000 x g for 1 minute.
6. Label EpiTect columns and transfer mix into column. Centrifuge the spin column at 12,000 x g for 1 minute and discard the flow-through.
7. Add 500µL of BW Buffer to each column and then centrifuge at 12,000 x g for 1 minute. Discard the flow-through.
8. Dispense 500µL of BD (stored at 2-4ºC) buffer without transferring any particulates and incubate spin columns at room temperature for 15 minutes.
9. Centrifuge columns at 14,000 x g for 1 minute and discard the flow-through.
10. Dispense 500µL of BW Buffer, centrifuge at 14,000 x g for 1 minute, and discard flow-through.
11. Repeat step 10.
12. Centrifuge the spin columns for 1 minute to remove any residual liquid.
13. Spin columns were placed open into a lidless 1.5ml microcentrifuge tube onto a 56ºC heating block and incubated for 5 minutes.
14. Spin columns are placed into new labeled 1.5ml microcentrifuge tubes.
15. Dispense 20µL of EB Buffer onto the membrane of the column and centrifuge at 14,000 x g for 1 minute.
16. Repeat step 15 twice to increase the potential yield.
17. Store at -20ºC.

Appendix C

Research Ethics Board Approval

QUEEN'S UNIVERSITY HEALTH SCIENCES & AFFILIATED TEACHING HOSPITALS RESEARCH ETHICS BOARD

June 8, 2010

This Ethics Application was subject to:

☐ Full Board Review
Meeting Date:
☒ Expedited Review

Dr. Will King
Department of Community Health and Epidemiology
Room 211
Carruthers Hall
Queen's University

Dear Dr. King,

Study Title: The Influence of Overweight and Obesity on Global DNA Methylation Levels as an Intermediate Event for Colon Cancer
Co-Investigators: Ms. Nora Zwingerman, Dr. Stephen Pang

I am writing to acknowledge receipt of your recent ethics submission. We have examined the protocol for your project (as stated above) and consider it to be ethically acceptable. This approval is valid for one year from the date of the Chair's signature below. This approval will be reported to the Research Ethics Board. Please attend carefully to the following list of ethics requirements you must fulfill over the course of your study:

- **Reporting of Amendments**: If there are any changes to your study (e.g. consent, protocol, study procedures, etc.), you must submit an amendment to the Research Ethics Board for approval. (see http://www.queensu.ca/vpr/reb.htm).
- **Reporting of Serious Adverse Events**: Any unexpected serious adverse event occurring locally must be reported within 2 working days or earlier if required by the study sponsor. All other serious adverse events must be reported within 15 days after becoming aware of the information.
- **Reporting of Complaints**: Any complaints made by participants or persons acting on behalf of participants must be reported to the Research Ethics Board within 7 days of becoming aware of the complaint. Note: All documents supplied to participants must have the contact information for the Research Ethics Board.
- **Annual Renewal**: Prior to the expiration of your approval (which is one year from the date of the Chair's signature below), you will be reminded to submit your renewal form along with any new changes or amendments you wish to make to your study. If there have been no major changes to your protocol, your approval may be renewed for another year.

Yours sincerely,

Albert Clark
Chair, Research Ethics Board

June 9, 2010
Date

ORIGINAL TO INVESTIGATOR - COPY TO DEPARTMENT HEAD- COPY TO HOSPITAL(S) /P&T (If appropriate) - FILE COPY

Study Code: EPID-316-10

- **Investigators please note that if your trial is registered by the sponsor, you must take responsibility to ensure that the registration information is accurate and complete**

ENTERED
JUL 15 2010

ORR
Feb 2011

ENTERED OCT 12 2010
SHIPPED OCT 12 2010