# ADIPOSITY, SEX HORMONES, AND REPETITIVE ELEMENT DNA METHYLATION IN HEALTHY POSTMENOPAUSAL WOMEN

by

Devon J. Boyne

A thesis submitted to the Department of Public Health Sciences

In conformity with the requirements for

the degree of Master of Science

Queen's University

Kingston, Ontario, Canada

August, 2016

# Abstract

**INTRODUCTION:** Low levels of methylation within repetitive DNA elements, such as long interspersed nuclear element-1 (LINE-1) and Alu repeats, are believed to epigenetically predispose an individual to cancer and other diseases. The extent to which lifestyle factors affect the degree of DNA methylation within these genomic regions has yet to be fully understood. Adiposity and sex hormones are established risk factors for certain types of cancer and other illnesses, particularly amongst postmenopausal women. The aim of the current investigation is to assess the impact of adiposity and sex hormones on LINE-1 and Alu methylation in healthy postmenopausal women.

**METHODS:** A cross-sectional study was conducted using baseline data from an ancillary study of the Alberta Physical Activity and Breast Cancer Prevention (ALPHA) Trial. Current adiposity was measured using a dual-energy x-ray absorptiometry (DXA) scan, computed tomography (CT) scan, and balance beam scale. Historical weights were self-reported in a questionnaire. Current endogenous sex hormone concentrations were measured in fasting blood serum. Estimated lifetime number of menstrual cycles was used as a proxy for cumulative exposure to ovarian sex hormones. Repetitive element methylation was quantified in white blood cells using a pyrosequencing assay. Linear regression was used to model the relations of interest while adjusting for important confounders.

**RESULTS:** Adiposity and serum estrogen concentrations were positively related to LINE-1 methylation but were not associated with Alu methylation. Cumulative ovarian sex hormone exposure had a "U-shaped" relation with LINE-1 regardless of folate intake and a negative

relation with Alu methylation amongst low folate consumers. Androgens were not associated with repetitive element DNA methylation in this population.

**CONCLUSION:** Adiposity and estrogens appear to play a role in maintaining high levels of repetitive element DNA methylation in healthy postmenopausal women. LINE-1 methylation may be a mechanism whereby estrogen exposure protects against cardiovascular and neurodegenerative illnesses. These results add to the growing body of literature showing how the epigenome is shaped by our lifestyle choices. Future prospective studies assessing the relation between levels of repetitive element DNA methylation in healthy individuals and subsequent disease risk are needed to better understand the clinical significance of these results.

## Co-Authorship

The conceptualization of this thesis project, the identification of relevant existing data, the review of the literature, the cleaning, linkage, and analysis of the data, and the preparation of the two manuscripts was done by Devon Boyne. Will King and Christine Friedenreich provided feedback about the design of the study, the creation of the variables, the analysis and interpretation of the data, the presentation of the results, and the draft of the entire thesis document. John McIntyre, Kerry Courneya, and Frank Stanczyk made substantial contributions to the underlying study with respect to data collection and study design and they reviewed the manuscripts included in this thesis.

# Acknowledgements

First and foremost, I must thank Will King for his guidance and supervision. He has and will continue to be a loyal mentor and role model. Secondly, none of this would have been possible without Christine Friedenreich whose work ethic is matched only by her organizational prowess. I look forward to continuing to work under her supervision during my doctoral studies. I would also like to thank my course instructors at Queen's University. In particular, a special thanks to Michael McIsaac for his willingness to discuss and teach concepts outside of the classroom. I am also indebted to Timothy Rosillo, Lee Watkins, and Gerri Lyman for their unwavering administrative and moral support. Lastly, I am grateful for the funding that I received through the Ontario Graduate Scholarship Program and the Queen's University Terry Fox Foundation Training Program in Transdisciplinary Cancer Research in partnership with the Canadian Institutes of Health Research.

# Table of Contents

# List of Tables

# List of Abbreviations

**ALPHA:** Alberta Physical Activity and Breast Cancer Prevention

**BIA:** Bioelectrical Impedance Analysis

**BMI:** Body Mass Index

**CIHR:** Canadian Institutes of Health Research

**CpG:** Cytosine-Phosphate-Guanine

**CT:** Computed Tomography

**DFE:** Dietary Folate Equivalent

**DXA:** Dual x-ray absorptiometry

**LINE-1:** Long Interspersed Nuclear Element-1

**LNMC:** Lifetime Number of Menstrual Cycles

**MET:** Metabolic Equivalent

**PCR:** Polymerase chain reaction

**SE:** Standard Error

**SHBG:** Sex hormone-binding globulin

# Chapter 1

# Introduction

The prevention of cancer begins with an understanding of the factors that underlie its initiation. Such an understanding encompasses knowledge about the duration, intensity, type, and timing of the exposure as well as information about who is most susceptible to its effects. A thorough understanding of the relation between an exposure and a disease is needed in order to implement effective and efficient public health policies, guidelines, and interventions.

Molecular epidemiology is an approach used to study the etiology of cancer and other illnesses. The core tenet of this type of research is that there is, "…a continuum of molecular/genetic alterations leading to cancer that can be accessed using biomarkers." (1) Exposure to a toxin does not immediately result in the formation of a tumour. Rather, sufficient exposure to a carcinogenic agent triggers a series of molecular changes that eventually result in a mutation from which a malignant neoplasm arises. Molecular epidemiologists study this chain of events in human populations by combining traditional epidemiologic methods with biomarkers developed by basic scientists and translational researchers. The biomarkers used by molecular epidemiologists can be roughly divided into three overlapping categories: markers of an internal or biologically effective dose of an exposure; markers of a preclinical or early carcinogenic effect and; markers of susceptibility to the effects of the exposure. (1)

Incorporating biomarkers into epidemiologic studies has several advantages. (1,2) For example, some environmental agents have multiple routes of exposure (e.g. people can be exposed to polycyclic aromatic hydrocarbons through air pollution, diet, and tobacco smoke) (3) or are difficult to measure using questionnaire data (e.g. dietary acrylamide exposure from

multiple food sources). (4) In such circumstances, using a biomarker of an internal dose can provide a researcher with a more accurate and precise measure of the exposure in question. (1,2) The use of genetic polymorphisms as markers of susceptibility can help to better identify and characterize populations who are particularly vulnerable to the effects of a risk factor. (5) A biomarker of a preclinical effect can allow for the prospective assessment of an intervention (e.g. an aerobic exercise program) that would have not been feasible to conduct using a clinical endpoint (e.g. a cancer diagnosis) because of the cost, large number of participants, and long period of time required to carry out such a study. (1,2) Studying the relation between an exposure and an early biomarker of a carcinogenic effect can also help to strengthen inferences about causality by allowing researchers to measure the effects of an exposure with more sensitivity and specificity and by providing scientists with information about the underlying mechanisms through which an exposure affects an outcome in a human population. (2,6)

It is within this larger research context that the current thesis project is situated. The thesis is comprised of two manuscripts (Chapters 4 and 5), each of which examines the relation between an established risk factor for breast and other types of cancer (i.e. adiposity and endogenous sex hormones) and a biomarker of an early carcinogenic effect (i.e. repetitive element DNA methylation). The goal of this investigation is to provide information about the frequency, intensity, type, and timing of these exposures with respect to their impact on cancer risk and to determine if repetitive element DNA methylation is a potential mechanism through which these exposure impact cancer risk. The two manuscripts are preceded by a review of the literature (Chapter 2). An additional methods section (Chapter 3) was added to provide the reader with a rationale for some of the methods used in the subsequent two manuscripts. The final chapter (Chapter 6) briefly summarizes the findings of the two manuscripts, elaborates on the

strengths and limitations of this investigation, discusses the implications of this study, and suggests future research directions.

## References


1. Vineis P, Perera F. Molecular epidemiology and biomarkers in etiologic cancer research: The new in light of the old. *Cancer Epidemiol Biomarkers Prev* 2007; 16(10): 1954–65.
2. Merlo DF, Sormani MP, Bruzzi P. Molecular epidemiology: New rules for new tools? *Mutat Res Mol Mech Mutagen* 2006; 600(1): 3–11.
3. White AJ, Bradshaw PT, Herring AH, Teitelbaum SL, Beyea J, Stellman SD, et al. Exposure to multiple sources of polycyclic aromatic hydrocarbons and breast cancer incidence. *Environ Int* 2016: 89: 186-192.
4. Pelucchi C, Bosetti C, Galeone C, La Vecchia C. Dietary acrylamide and cancer risk: An updated meta-analysis. *Int J Cancer* 2015; 136(12): 2912–22.
5. Perera FP. Environment and cancer: who are susceptible? *Science* 1997; 278(5340): 1068–73.
6. Demetriou CA, Vineis P. Carcinogenicity of ambient air pollution: Use of biomarkers, lessons learnt and future directions. *J Thorac Dis* 2015; 7(1): 67–95.

# Chapter 2

# Background and Literature Review

## Epigenetics and DNA methylation: A Brief Introduction

A growing body of literature is beginning to uncover how epigenetic mechanisms can mediate the relation between certain risk factors and disease. Epigenetics was recently defined by Deans et al. (2015) as "the study of phenomena and mechanisms that cause chromosome-bound, heritable changes to gene expression that are not dependent on changes to DNA sequence." (1) In other words, epigenetic alterations affect the expression of the genome and can be passed from one cell to another without changing the underlying base sequence. In her book, "The Epigenetic Revolution: How Modern Biology is Rewriting Our Understanding of Genetics, Disease, and Inheritance," Nessa Carey puts forth a useful analogy: epigenetics are like the director's notes on the script of a play in that these notes affect the interpretation of the script and the production of the play without changing the actual script itself. (2)

First discovered in 1948 (ref. 3), DNA methylation has become the most widely studied and best understood epigenetic mechanism to date. (4) The methylation of mammalian DNA is characterized by the covalent binding of a methyl group ($CH_3$) to a cytosine base within a cytosine-phosphate-guanine dinucleotide (CpG) to form 5-methylcytosine. (5) The consequences of DNA methylation depend upon the genomic region in which it takes place. (6) High levels of DNA methylation within a promoter region of a gene or within transposable elements suppresses or completely silences their expression. (6) The impact of DNA methylation in other genomic regions (e.g. gene bodies) or within non-CpG sites is currently not well known. (6) Although its function has yet to be fully understood, DNA methylation is recognized as having a crucial role

in a number of critical physiological processes such as the regulation of cellular differentiation, tissue-specific gene expression, genomic imprinting, and X-inactivation. (7)

At the molecular level, DNA methylation patterns are created, maintained, and erased by various biological mechanisms. Methyl groups are attached to cytosine bases by enzymes known as DNA methyltransferases (DNMTs). (8) DNMTs create new methylation patterns and transfer pre-existing methylation patterns from parent cells to daughter cells during cell division. (9) The methyl groups used in these methylation reactions are supplied through one-carbon metabolism in which S-adenosylmethionine (SAM) becomes S-adenosylhomocysteine (SAH) after it transfers its methyl group to a DNMT. (10) Passive de-methylation of the genome can occur when the activity of DNMTs are suppressed such that methylation patterns are not properly maintained during cellular division. (11,12) Recent findings suggest that ten-eleven translocation (TET) proteins can actively induce demethylation by hindering the activity of DNMTs and by promoting base excision through the oxidization of methylated cytosine bases. (12)

As implied by the previous paragraph, patterns of DNA methylation are not immutable. Rather, the methylation of certain genomic regions can change in response to ageing and certain environmental insults. (4,13) Referring back to Nessa Carrey's analogy, a director's notes can be revised and edited in response to certain feedback. One can imagine how such alterations might affect the interpretation of the script in a way that would be detrimental to the production of the play. Herein lies the focus of this thesis project – to describe the relation between certain behavioral factors and epigenetic changes that are believed to impact the risk of cancer and other diseases.

**Repetitive Element DNA Methylation – A Biomarker of an Early Carcinogenic Effect**

Epigenetic aberrations, particularly changes in the levels of DNA methylation, have an important role in the etiology of a wide variety of diseases including cancer, cardiovascular disease, neurodegenerative disorders, autoimmune disease, and mental illness. (14,15) Cancer has been the primary focus of most disease-related epigenetic research to date. The link was first established in 1983 when Feinberg and Vogelstein discovered that cancer cells had abnormal methylation patterns compared to adjacent normal cells taken from the same individual. (16) It is now recognized that nearly all cancer cells feature aberrant DNA methylation characterized by the hypermethylation of the promoter regions of tumour suppressor genes and the hypomethylation of repetitive elements. (17) Emerging evidence suggests that some epigenetic aberrations, particularly repetitive element DNA hypomethylation, are not merely consequences of cancer but are early carcinogenic events which help to drive the initiation of a tumour.

More than half of the human genome consists of highly repetitive base sequences, the most common type of which are interspersed repetitive elements. (18) Interspersed repetitive elements can be classified according to their length: short-interspersed repetitive elements (SINEs) are usually less than 500 base pairs whereas long-interspersed repetitive elements (LINEs) are typically greater than 5000 base pairs. (19) The two most common types of LINEs and SINEs active in human DNA are long interspersed element-1 (LINE-1) and Alu elements respectively. (20) There are around 500,000 copies of LINE-1 elements and 1,000,000 copies of Alu elements which correspondingly account for roughly 20% and 14% of the human genome. (20) Most CpG sites are found within these repetitive elements and these genomic regions are typically heavily methylated such that they account for a substantial proportion of global (i.e. the

total amount of) DNA methylation. (21) This high level of methylation helps to silence the activity of these retrotransposons. (5)

Accumulating research suggests that low levels of LINE-1 and Alu methylation can increase an individual's risk for developing certain types of cancer and possibly other diseases. (22,23) Hypomethylation within LINE-1 and Alu repeats can result in an increase in their expression which in turn can have harmful cellular consequences such as the induction of chromosomal instability and disruption of tumour suppressor and oncogene expression. (20,24–28) The causal role of genomic hypomethylation in cancer initiation is supported by experimental evidence in which induced hypomethylation in mice has been shown to result in the formation of tumours. (29) These results have been mirrored in prospective epidemiologic studies showing that healthy individuals with lower LINE-1 and Alu methylation have a significantly increased risk of developing cancer. (30,31) For example, DeRoo et al. (2014) conducted a nested case-control study in which they compared the risk of developing breast cancer across quartiles of LINE-1 methylation as measured at baseline using peripheral blood samples from healthy women aged 35 to 74 years of age. (32) DeRoo and colleagues reported a significant negative dose-response relationship between LINE-1 methylation and breast cancer risk and found that women in the lowest quartile of methylation were 75% more likely to develop breast cancer in comparison to women in the highest quartile of methylation (HR = 1.75; 95% CI: 1.19, 2.59). Lower levels of DNA methylation within LINE-1 and Alu repeats may therefore serve as a biomarker of an early carcinogenic effect. New findings suggests that the harmful effects of LINE-1 and Alu hypomethylation are not specific to cancer in that they may also be related to the risk of heart disease (33) and certain neurodegenerative disorders. (34) Although the effects of aberrantly high levels of LINE-1 and Alu methylation are unknown, an

association with cancer and heart disease risk has been reported in a small number of studies. (35–37)

**Determinants of DNA Methylation**

Although the literature is limited, several suspected determinants of DNA methylation have been identified which include age, race, sex, diet, smoking, alcohol consumption, obesity, hormones, physical activity, environmental chemical exposures, morbidity, and genetics. (13,38,39)

***Age ~*** Aging is the single biggest risk factor for cancer and other chronic diseases. (40) Epidemiologic studies have shown that the epigenome is not impervious to the aging process – repetitive elements that are heavily methylated tend to become hypomethylated as individuals grow older. (4) Such epigenetic changes are believed to be caused by accumulating damage from environmental insults and random errors made during cellular replication. (41) Age may also modify the effects of certain exposures – some researchers argue that the epigenome is most susceptible to the effects of environmental factors in utero and during early development although this has yet to be conclusively demonstrated. (42)

***Sex, Race, and Genetics ~*** Sex and possibly racial differences in DNA methylation patterns exist amongst human populations. Many epidemiologic studies have found that women tend to have lower levels of repetitive element methylation in relation to men. (43–45) Such findings could be attributable to differences in genetics and/or exposure to sex hormones. (45) Some epidemiologic studies have reported significant racial differences in DNA methylation after controlling for various confounders, (44,46,47) however, such findings are inconsistent. (48)

Polymorphisms affecting the expression of epigenetic enzymes such as DNMTs or proteins involved in the one-carbon metabolism (e.g. MTHFR) have been found to impact DNA methylation and the risk of cancer. (49,50)

***Diet and Alcohol ~*** Dietary factors associated with one-carbon metabolism – the pathway responsible for the supply of methyl groups used in various cellular processes – may be important determinants of DNA methylation. (11) Epidemiologic and experimental studies have demonstrated that dietary nutrient intake and alcohol consumption have the potential to influence DNA methylation and other epigenetic processes. (51–54) The functioning of one-carbon metabolism is dependent upon several dietary micro-nutrients which include methionine, choline, betaine, vitamin B2, B6, B12, and most importantly folate (B9). (53) Similarly, alcohol consumption can impact one-carbon metabolism through various pathways such as reducing the bioavailability of SAM or hindering the metabolism of folate. (52) Results from epidemiologic investigations regarding the impact of dietary folate and alcohol consumption on levels of repetitive element methylation have been inconsistent. (52,53,55) The impact of macronutrients other than alcohol (e.g. proteins, fats, or carbohydrates) and dietary patterns (e.g. Western or Mediterranean) on the epigenome has also yet to be fully understood. (53,54)

***Environmental Toxins and Smoking ~*** A growing body of literature supports the etiologic role of environmental chemicals in determining DNA methylation profiles. (43,56) Levels of DNA methylation have been shown to vary in response to a wide range of environmental insults which include air pollution (e.g. benzene, particulate matter, and black carbon), persistent organic pesticides (e.g. organochlorine pesticides and polychlorinated biphenyls (PCBs)), endocrine disruptors (e.g. diethylstilbestrol (DES) and bisphenol A (BPA)),

heavy metals (e.g. arsenic, cadmium, chromium, lead, and nickel), water disinfection by-products (e.g. chloroform and trihalomethanes), and radiation. (39,56–60)

Although tobacco has been shown to affect DNA methylation within specific genes, there is no evidence for an association between tobacco smoke and repetitive element DNA methylation. (39)

***Physical Activity*** **~** A review of all known studies investigating the impact of physical activity and exercise on DNA methylation was recently published by Horsburgh and colleagues (2015). (61) The primary mechanism through which exercise is believed to influence the epigenome is by reducing chronic inflammation which has been shown to regulate the expression of DNMTs and has been linked to aberrant DNA methylation. (61) Although some studies have found that higher levels of physical activity are associated with higher levels of repetitive element DNA methylation, such findings are inconsistent and firm conclusions cannot be made from the current state of evidence. (61)

**Sex Hormones, Adiposity, and DNA Methylation: Biological Plausibility**

The purpose of this study is to determine the impact of adiposity and estrogens on DNA methylation. Epidemiologic evidence has convincingly demonstrated that obesity and sex hormones causally contribute towards the initiation of postmenopausal breast cancer and other diseases. (62–65) Aberrant DNA methylation may be one mechanism whereby exposure to adiposity and estrogens affect disease risk.

***Estrogens*** **~** Several factors associated with exposure to estrogens have been shown to affect the risk of breast and other types of cancer, particularly amongst postmenopausal women. These factors include endogenous sex hormone levels (both estrogens and androgens), levels of

sex hormone-binding globulin (which affects the bioavailability of endogenous hormones), reproductive characteristics (i.e. parity, age at first birth, and breastfeeding), menstrual history (i.e. age at menarche and age at menopause), and exogenous hormone use (i.e. oral contraceptive or hormone replacement therapy use). (62,66) In contrast to the cancer literature, estrogen exposure has been shown to be protective against cardiovascular and neurodegenerative diseases. (67,68) Understanding the impact of estrogens on repetitive element DNA methylation may provide insight into the molecular mechanisms at play which might help to explain this heterogeneity with respect to disease risk.

Although few studies to date have investigated the association between sex hormones and DNA methylation, the relation is biologically plausible and is supported by analogous evidence. Sex hormones have been shown to have an impact on various components of one-carbon metabolism. In particular, exposure to estrogens has been shown to negatively affect the metabolism of folate, methyl groups, and homocysteine. (69,70) The activity of DNMT and methyl CpG binding proteins has also been shown to be repressed by estrogens. (71) In addition, sex hormones may indirectly influence DNA methylation by inducing other epigenetic changes such as histone modifications. (71) By analogy, studies in humans have shown that soy isoflavones (which affect endogenous estrogen levels) and endocrine disruptors (i.e. chemicals that simulate or interfere with endogenous hormones) can impact DNA methylation which supports the hypothesis that DNA methylation is susceptible to hormonal influence. (72,73) Given the negative impact that estrogens are believed to have on the activity of epigenetic enzymes and supply of methyl groups, one would expect that higher levels of estrogen exposure would be associated with lower levels of repetitive element DNA methylation.

***Adiposity ~*** The continued increase in obesity rates across the globe and its relation to a number of chronic diseases make it a public health priority. (63) Among postmenopausal women, excess body fat is an established risk factor for breast and other types of cancers including endometrial, pancreatic, and colorectal cancers. (65) There are three primary ways in which body fatness is conceptualized in the literature: 1) overall adiposity; 2) body fat distribution; 3) changes in weight over time. (74) Most epidemiologic studies have relied upon body mass index (BMI) as a proxy measure of overall adiposity. (75) The distribution of body fat is also of biological relevance – some studies have found that central adiposity (i.e. body fat located within and around the abdomen or trunk) can have more of a detrimental health effect in comparison to adipose tissue found in more proximal regions (e.g. the arms or the legs) independent of overall adiposity. (75) Central adiposity is commonly assessed in the literature by waist circumference or waist-to-hip ratio. (75) More accurate measures of these constructs that account for the amount of lean body mass (e.g. body fat percent or abdominal fat area) can be obtained using medical imaging devices such as a dual energy x-ray absorptiometry (DXA) or computed tomography (CT) scan. (75) The use of such imaging devices also provides a means to further explore the distribution of body fat by allowing central adiposity to be segmented into body fat surrounding the organs (i.e. intraabdominal or visceral fat) and adipose tissue found near the surface of the skin (i.e. subcutaneous fat). (75) Although the impact of weight loss and weight cycling (i.e. sequential patterns of weight loss and gain) have yet to be established, weight gain has been consistently linked to cancer risk. (65,74) Although there is evidence that the relation between adiposity and cancer risk differs according to time (e.g. childhood obesity has been found to be protective against breast cancer), the biological windows during which

women are most susceptible to the effects of excess body fat and the underlying mechanisms driving such differences have yet to be fully uncovered. (76)

Although the precise mechanisms whereby higher levels of body fat may impact DNA methylation in postmenopausal women has yet to be established, there are several plausible pathways. After menopause, aromatase found in adipose tissue becomes the primary source of estrogen production. (77) Epidemiologic studies have found that higher levels of adiposity have been positively associated with higher levels of circulating estrogen concentrations in postmenopausal women. (78) As previously described, exposure to estrogens may have an impact on DNA methylation. Obesity is also known to cause chronic low-grade inflammation which has been shown to impact the activity of DNMT and interfere with one-carbon metabolism. (61,79,80) Excess adiposity is also known to increase oxidative stress by triggering the production of reactive oxygen species (ROS). (81) Oxidative stress can induce aberrant DNA methylation by affecting the normal functioning of TET proteins, DNMTs, and the one-carbon metabolism.(82–84) Body fat may therefore alter DNA methylation within postmenopausal women by affecting their exposure to estrogens, chronic inflammation, and oxidative stress. Since there are multiple potential pathways through which adiposity can affect DNA methylation, it is difficult to form an *a priori* hypothesis about the direction of the relation between body fat and repetitive element DNA methylation. However, if one assumes that this relation is largely mediated by estrogen in postmenopausal women, then higher levels of body fat would be expected to be related to lower levels of repetitive element DNA methylation.

**DNA Methylation as a Biomarker of an Early Carcinogenic Effect**

In the current study, repetitive element DNA methylation is being conceptualized as a biomarker of an early carcinogenic effect to determine if it is a possible mechanism through which adiposity and sex hormone exposure impact cancer risk. Various informal and formal sets of criteria have been developed to assess the appropriateness of an intermediate endpoint with respect to its ability to predict the impact of an exposure on a clinical end point. (85,86) Although a formal evaluation is beyond the scope of this review, the appropriateness of repetitive element DNA methylation as a biomarker of an early carcinogenic effect will be briefly discussed with respect to its biological relevance, relation to cancer incidence, laboratory assay's reliability and validity, and intra-individual variability over time and across tissues. (85)

If a intermediate biomarker is not on the causal pathway between an exposure and an outcome, changes in the biomarker may not reflect changes in the clinical endpoint (e.g., the relation between the biomarker and clinical endpoint may be confounded). (87) Epigenetic aberrations, particularly repetitive element DNA hypomethylation, are now believed to play a causal role in the initiation of cancer. (17,88) Therefore, LINE-1 and Alu methylation are biologically relevant endpoints since they are thought to exist on the causal pathway. It is important to note that adiposity and sex hormones may influence cancer risk through mechanisms that are independent of repetitive element DNA methylation (e.g., estrogen metabolites can have genotoxic effects). (64) Therefore, the impact of these exposures on LINE-1 and Alu methylation may correspond to a reduction in cancer incidence but not to an elimination of their carcinogenic effects entirely. (87,89)

The laboratory methods used to quantify a biomarker must be accurate and precise if the biomarker is to be used within an epidemiologic setting. (90) A pyrosequencing assay was used

in the current study to quantify repetitive element DNA methylation which has been shown to be both valid and reliable. (91,92) By repeating the assay on eight samples, it was determined that there was little intra-lab variability for the LINE-1 and Alu measures with coefficients of variation equal to 1.8% and 2.2% respectively. Thus, the measures of repetitive element DNA methylation used in the current study are accurate and reliable in terms of both inter- and intra-lab variability.

It is desirable to use a biomarker that is relatively stable over time so that only a single measure is needed to measure the construct in question precisely. A single measure of a biomarker that has a great deal of fluctuation over short periods of time (e.g. daily or monthly) will suffer from measurement error if additional adjustments are not taken (e.g., taking multiple measures of the same participant). (90) Levels of LINE-1 and Alu methylation have been found not to change over short periods of time (e.g., over the course of four days (93) or the span of a natural menstrual cycle (94)). In general, studies that have found an association between age and LINE-1 and Alu methylation report small declines in absolute percent methylation (~0.1%) from year to year. (95) The temporal stability of repetitive element DNA methylation is a characteristic of a biomarker well suited for incorporation into epidemiologic studies.

Akin to temporal stability, a biomarker would ideally have little intra-individual variability within or between different tissues and cell types so that measures taken from readily available sites (e.g. blood or saliva) can be used as proxy measures of levels in the target tissues of interest (e.g. breast or colorectal tissue). (85) Epidemiologic investigations are often forced to rely upon measures taken from specimens that are readily available for ethical and practical reasons. In the current study, we assessed DNA methylation within white blood cells. This biological source may be problematic because patterns of DNA methylation have been found to

vary across different tissues (96,97) and cell types. (98,99) However, levels of DNA methylation within certain genomic regions has been shown to be conserved across different tissues and one study found that a large degree of inter-tissue variability was attributed to DNA methylation within genes that have a tissue-specific function. (97) Until the extent to which levels of LINE-1 and Alu methylation within white blood cells correspond to levels in other tissue sites has been established, caution should be taken when interpreting the findings of this investigation.

A biomarker of an early carcinogenic effect should be able to assess the probability of developing the clinical endpoint of interest with a high degree of validity and consistency. (87) While some studies have associated lower levels of LINE-1 and Alu methylation measured in peripheral blood with the risk of several different types of cancer, the relation has yet to be firmly established and the evidence is inconsistent. (23,100) While repetitive element DNA methylation in peripheral blood cells may be indicative of general cancer risk, future prospective studies relating baseline levels of DNA methylation to the incidence of specific types and sub-types of cancer are needed to understand and appreciate the clinical relevance of the results from this study and how specific they are to a certain type of cancer.

In general, repetitive element DNA methylation is a promising biomarker of an early carcinogenic effect because it is believed to causally contribute to cancer formation and it is a temporally stable mark that can be measured using valid and reliable laboratory methods. To better assess its validity as an intermediate biomarker, however, the relation between repetitive element DNA methylation and disease and the extent to which relations identified in surrogate tissues persist in target tissues of interest needs to be better understood.

**Study Rationale**

Affecting approximately 1 in 9 women, breast cancer is the most common type of cancer and second most common cause of cancer deaths amongst Canadian females. (101) The discordance between the declining breast cancer mortality rate and relatively stagnant incidence rate demonstrates that preventive efforts have failed to mirror the advances made in the early detection and treatment of this disease. (101) Adiposity and sex hormones are two important modifiable risk factors for postmenopausal breast cancer. (102) Epigenetics, particularly DNA methylation, may play an important role in mediating the health effects of these exposures. Etiologic research targeted at furthering our understanding of the carcinogenic effects of these factors will ultimately help to further preventive efforts targeted at reducing the incidence of breast cancer and other chronic diseases related to aberrant DNA methylation.

The extent to which modifiable factors influence DNA methylation has not yet been established. (55) The large majority of the literature has focused on establishing the link between DNA methylation and cancer while a comparatively lesser body of research has examined the determinants of this epigenetic mark. (103) There exists a considerable gap in the literature in that the determinants of DNA methylation are poorly understood. Results from this investigation will help to address this gap and can provide information to help guide future etiologic research. Adiposity and sex hormones are complex exposures that have heterogeneous relationships with disease. For example, epidemiologic research has found that obesity is protective against premenopausal breast cancer but is harmful with respect to postmenopausal breast cancer. (65) Suggestive of an epigenetic mechanism, studies have also shown that childhood obesity can protect against adulthood breast cancer. (76) On the other hand, higher levels of estrogen exposure increase the risk for breast and other types of cancer but confer a protective effect

against cardiovascular and neurodegenerative diseases. (62,67,104) A better understanding of the underlying biological mechanisms will help to explain such heterogeneity which is information that may ultimately help to inform medical practice and public health guidelines.

Few studies have investigated the impact of sex hormones or adiposity on repetitive element DNA methylation and the results of those that have are inconsistent (see discussion section of Chapters 4 and 5). The current study used a comprehensive set of highly accurate and reliable hormonal and adiposity metrics. In contrast to many previous studies, the current investigation a full assessment of potential confounding factors. Results from this research provide evidence pertaining to the impact of obesity and estrogens on aberrant DNA methylation patterns which will ultimately help broaden our understanding of the health implications associated with these two established cancer risk factors.

## References


1. Deans C, Maggert K a. What Do You Mean, "Epigenetic"? *Genetics* 2015; 199(4): 887–96.
2. Carey N. The epigenetics revolution : how modern biology is rewriting our understanding of genetics, disease and inheritance. *Icon Books Ltd.* Columbia University Press; 2012.
3. Hotchkiss R. The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography. *J Biol Chem* 1948; 175(1): 315–32.
4. Jones MJ, Goodman SJ, Kobor MS. DNA methylation and healthy human aging. *Aging Cell* 2015; 14(6): 924-932.
5. Bernstein BE, Meissner A, Lander ES. The Mammalian Epigenome. *Cell* 2007; 128(4): 669–81.
6. Jones P a. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* 2012; 13(7): 484–92.
7. Tost J. DNA methylation: An introduction to the biology and the disease-associated changes of a promising biomarker. *Mol Biotechnol* 2010; 44(1): 71–81.
8. Moore LD, Le T, Fan G. DNA Methylation and Its Basic Function. *Neuropsychopharmacology* 2012; 38(1): 23–38.
9. Jin B, Li Y, Robertson KD. DNA Methylation: Superior or Subordinate in the Epigenetic Hierarchy? *Genes Cancer* 2011; 2(6): 607–17.
10. King WD, Ho V, Dodds L, Perkins SL, Casson RI, Massey TE. Relationships among biomarkers of one-carbon metabolism. *Mol Biol Rep* 2012; 39(7): 7805–12.

11. Anderson OS, Sant KE, Dolinoy DC. Nutrition and epigenetics: An interplay of dietary methyl donors, one-carbon metabolism and DNA methylation. *J Nutr Biochem* 2012; 3(8): 853–9.

12. Wu H, Zhang Y. Reversing DNA methylation: Mechanisms, genomics, and biological functions. *Cell* 2014; 156(1-2): 45–68.

13. Foley DL, Craig JM, Morley R, Olsson CJ, Dwyer T, Smith K, et al. Prospects for epigenetic epidemiology. *Am J Epidemiol* 2009; 169(4): 389–400.

14. Portela A, Esteller M. Epigenetic modifications and human disease. *Nat Biotechnol* 2010; 28(10): 1057–68.

15. Handel AE, Ebers GC, Ramagopalan S V. Epigenetics: molecular mechanisms and implications for disease. *Trends Mol Med* 2010; 16(1): 7–16.

16. Feinberg AP, Vogelstein B. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* 1983; 301(5895): 89–92.

17. Rodríguez-Paredes M, Esteller M. Cancer epigenetics reaches mainstream oncology. *Nat Med* 2011; 17(3): 330–9.

18. Liang K-C, Tseng JT, Tsai S-J, Sun HS. Characterization and distribution of repetitive elements in association with genes in the human genome. *Comput Biol Chem* 2015; 57: 1–10.

19. Singer MF. SINEs and LINEs: highly repeated short and long interspersed sequences in mammalian genomes. *Cell* 1982; 28(3): 433–4.

20. Miousse IR, Chalbot M-CG, Lumen A, Ferguson A, Kavouras IG, Koturbash I. Response of Transposable Elements to Environmental Stressors. *Mutat Res Mutat Res* 2015; 765: 19–39.

21. Rollins R a, Haghighi F, Edwards JR, Das R, Zhang MQ, Ju J, et al. Large-scale structure of genomic methylation patterns Large-scale structure of genomic methylation patterns. *Genome Res* 2005; 16(2): 157–63.

22. Zhong J, Agha G, Baccarelli AA. The Role of DNA Methylation in Cardiovascular Risk and Disease: Methodological Aspects, Study Design, and Data Analysis for Epidemiological Studies. *Circ Res* 2016; 118(1): 119–31.

23. Woo HD, Kim J. Global DNA hypomethylation in peripheral blood leukocytes as a biomarker for cancer risk: A meta-analysis. *PLoS One* 2012; 7(4): e34615.

24. Hur K, Cejas P, Feliu J, Moreno-Rubio J, Burgos E, Boland CR, et al. Hypomethylation of long interspersed nuclear element-1 (LINE-1) leads to activation of proto-oncogenes in human colorectal cancer metastasis. *Gut* 2014; 63(4): 635–46.

25. Ade C, Roy-Engel AM, Deininger PL. Alu elements: An intrinsic source of human genome instability. *Curr Opin Virol* 2013; 3(6): 639–45.

26. Daskalos A, Nikolaidis G, Xinarianos G, Savvari P, Cassidy A, Zakopoulou R, et al. Hypomethylation of retrotransposable elements correlates with genomic instability in non-small cell lung cancer. *Int J Cancer* 2009; 124(1): 81–7.

27. Padeken J, Zeller P, Gasser SM. Repeat DNA in genome organization and stability. *Curr Opin Genet Dev* 2015; 31: 12–9.

28. Miousse IR, Koturbash I. The Fine LINE: Methylation Drawing the Cancer Landscape. *Biomed Res Int* 2015; 2015: 1–8.

29. Gaudet F, Hodgson JG, Eden A, Jackson-Grusby L, Dausman J, Gray JW, et al. Induction of tumors in mice by genomic hypomethylation. *Science* 2003; 300(5618): 489–92.

30. Kawano H, Saeki H, Kitao H, Tsuda Y, Otsu H, Ando K, et al. Chromosomal Instability Associated with Global DNA Hypomethylation is Associated with the Initiation and Progression of Esophageal Squamous Cell Carcinoma. *Ann Surg Oncol* 2014; 21(4): 696–702.

31. Zhu ZZ, Sparrow D, Hou L, Tarantini L, Bollati V, Litonjua A a., et al. Repetitive element hypomethylation in blood leukocyte DNA and cancer incidence, prevalence, and mortality in elderly individuals: The Normative Aging Study. *Cancer Causes Control* 2011; 22(3): 437–47.

32. Deroo LA, Bolick SCE, Xu Z, Umbach DM, Shore D, Weinberg CR, et al. Global DNA methylation and one-carbon metabolism gene polymorphisms and the risk of breast cancer in the Sister Study. *Carcinogenesis* 2013; 35(2): 333–8.

33. Muka T, Koromani F, Portilla E, O'Connor A, Bramer WM, Troup J, et al. The role of epigenetic modifications in cardiovascular disease: A systematic review. *Int J Cardiol* 2016; 212: 174–83.

34. Lardenoije R, Iatrou A, Kenis G, Kompotis K, Steinbusch HWM, Mastroeni D, et al. The epigenetics of aging and neurodegeneration. *Prog Neurobiol* 2015; 131: 21–64.

35. Kim M, Long TI, Arakawa K, Wang R, Yu MC, Laird PW. DNA methylation as a biomarker for cardiovascular disease risk. *PLoS One* 2010; 5(3): 1–8.

36. Andreotti G, Karami S, Pfeiffer RM, Hurwitz L, Liao LM, Weinstein SJ, et al. LINE1 methylation levels associated with increased bladder cancer risk in pre-diagnostic blood DNA among US (PLCO) and European (ATBC) cohort study participants. *Epigenetics* 2013; 9(3): 404–15.

37. Liao LM, Brennan P, van Bemmel DM, Zaridze D, Matveev V, Janout V, et al. LINE-1 methylation levels in leukocyte DNA and risk of renal cell cancer. *PLoS One* 2011; 6(11): e27361.

38. Ogino S, Lochhead P, Chan AT, Nishihara R, Cho E, Wolpin BM, et al. Molecular pathological epidemiology of epigenetics: emerging integrative science to analyze environment, host, and disease. *Mod Pathol* 2013; 26(4): 465–84.

39. 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; 6(7): 828–37.

40. DePinho RA. The age of cancer. *Nature* 2000; 408(6809): 248–54.

41. Zampieri M, Ciccarone F, Calabrese R, Franceschi C, Burkle A, Caiafa P. Reconfiguration of DNA methylation in aging. *Mechanisms of Ageing and Development* 2015. 151: 60–70.

42. Huidobro C, Fernandez AF, Fraga MF. Aging epigenetics: causes and consequences. *Mol Aspects Med* 2013; 34(4): 765–81.

43. 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; 6(7): 828–37.

44. 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* 2014; 6(5): 623–9.

45. El-Maarri O, Walier M, Behne F, Van Üüm J, Singer H, Diaz-Lacava A, et al. Methylation at global LINE-1 repeats in human blood are affected by gender but not by age or natural hormone cycles. *PLoS One* 2011; 6(1): e16252.

46. Adkins RM, Krushkal J, Tylavsky F a., Thomas F. Racial differences in gene-specific DNA methylation levels are present at birth. *Birth Defects Res Part A - Clin Mol Teratol* 2011; 91(8): 728–36.

47. Lam LL, Emberly E, Fraser HB, Neumann SM, Chen E, Miller GE, et al. Factors underlying variable DNA methylation in a human community cohort. *Proc Nat Acad Sci USA* 2012; 109(2): 17253–60.

48. Axume J, Smith SS, Pogribny IP, Moriarty DJ, Caudill M a. Global leukocyte DNA methylation is similar in African American and Caucasian under conditions of controlled folate intake. *Epigenetics* 2007; 2(1): 66–8.

49. Tajuddin SM, Amaral AFS, Fernández AF, Rodríguez-Rodero S, Rodríguez RM, Moore LE, et al. Genetic and non-genetic predictors of LINE-1 methylation in leukocyte DNA. *Environ Health Perspect* 2013; 121(6): 650–6.

50. Huidobro C, Fernandez AF, Fraga MF. The role of genetics in the establishment and maintenance of the epigenome. *Cellular and Molecular Life Sciences* 2013. 70(9): 1543–73.

51. Zakhari S. Alcohol Metabolism and Epigenetic Changes. *Alcohol Res Curr Rev* 2013; 35(1): 6–16.

52. Varela-Rey M, Woodhoo A, Martinez-Chantar M-L, Mato JM, Lu SC. Alcohol, DNA methylation, and cancer. *Alcohol Res* 2013; 35(1): 25–35.

53. Anderson OS, Sant KE, Dolinoy DC. Nutrition and epigenetics: an interplay of dietary methyl donors, one-carbon metabolism and DNA methylation. *J Nutr* 2012; 23(8): 853–9.

54. Jiménez-Chillarón JC, Díaz R, Martínez D, Pentinat T, Ramón-Krauel M, Ribó S, et al. The role of nutrition on epigenetic modifications and their implications on health. *Biochimie* 2012; 94(11): 2242–63.

55. 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; 6(7): 828–37.

56. Ruiz-Hernandez A, Kuo C-C, Rentero-Garrido P, Tang W-Y, Redon J, Ordovas JM, et al. Environmental chemicals and DNA methylation in adults: a systematic review of the epidemiologic evidence. *Clin Epigenetics* 2015; 7(1): 55–79.

57. Feil R, Fraga MF. Epigenetics and the environment: emerging patterns and implications. *Nat Rev Genet* 2012; 13(2): 97–109.

58. Collotta M, Bertazzi PA., Bollati V. Epigenetics and pesticides. *Toxicology* 2013; 307: 35–41.

59. Christensen BC, Marsit CJ. Epigenomics in environmental health. *Front Genet* 2011;2:1–10.

60. Hou L, Zhang X, Wang D, Baccarelli A. Environmental chemical exposures and human epigenetics. *Int J Epidemiol* 2012; 41(1): 79–105.

61. Horsburgh S, Robson-Ansley P, Adams R, Smith C. Exercise and inflammation-related epigenetic modifications: focus on DNA methylation. *Exerc Immunol Rev* 2015; 21: 26–41.

62. Brown SB, Hankinson SE. Endogenous estrogens and the risk of breast, endometrial, and ovarian cancers. *Steroids* 2014; 99:8–10.

63. Dixon JB. The effect of obesity on health outcomes. *Mol Cell Endocrinol* 2010; 316(2): 104–8.

64. Germain D. Estrogen Carcinogenesis in Breast Cancer. *Endocrinology and Metabolism Clinics of North America* 2011; 40(3): 473–84.

65. Renehan AG, Zwahlen M, Egger M. Adiposity and cancer risk: new mechanistic insights from epidemiology. *Nat Rev Cancer* 2015; 15(8): 484–98.

66. Anderson KN, Schwab RB, Martinez ME. Reproductive risk factors and breast cancer subtypes: A review of the literature. *Breast Cancer Res Treat* 2014; 144(1): 1–10.

67. Xing D, Nozell S, Chen YF, Hage F, Oparil S. Estrogen and mechanisms of vascular protection. *Arterioscler Thromb Vasc Biol* 2009; 29(3): 289–95.

68. Brann DW, Dhandapani K, Wakade C, Mahesh VB, Khan MM. Neurotrophic and neuroprotective actions of estrogen: Basic mechanisms and clinical implications. *Steroids* 2007; 72(5): 381–405.

69. Williams KT, Schalinske KL. New insights into the regulation of methyl group and homocysteine metabolism. *J Nutr* 2007; 137(2): 311–4.

70. Friso S, Lamon-Fava S, Jang H, Schaefer EJ, Corrocher R, Choi SW. Oestrogen replacement therapy reduces total plasma homocysteine and enhances genomic DNA methylation in postmenopausal women. *Br J Nutr* 2007; 97(4): 617–21.

71. Zhang X, Ho SM. Epigenetics meets endocrinology. *J Mol Endocrinol* 2011; 46(1): 11-32.

72. Qin W, Zhu W, Shi H, Hewett JE, Ruhlen RL, MacDonald RS, et al. Soy isoflavones have an antiestrogenic effect and alter mammary promoter hypermethylation in healthy premenopausal women. *Nutr Cancer* 2009; 61(2): 238–44.

73. Fernandez S V., Huang Y, Snider KE, Zhou Y, Pogash TJ, Russo J. Expression and DNA methylation changes in human breast epithelial cells after bisphenol A exposure. *Int J Oncol* 2012; 41(1): 369–77.

74. Friedenreich CM. Review of anthropometric factors and breast cancer risk. *Eur J Cancer Prev* 2001; 10(1): 15–32.

75. James FR, Wootton S, Jackson A, Wiseman M, Copson ER, Cutress RI. Obesity in breast cancer - What is the risk factor? *Eur J Cancer* 2015; 51(6): 705–20.

76. Moley KH, Colditz GA. Effects of obesity on hormonally driven cancer in women. *Sci Transl Med* 2016; 8(323): 1–6.

77. Key TJ, Appleby PN, Reeves GK, Roddam a, Dorgan JF, Longcope C, et al. Body mass index, serum sex hormones, and breast cancer risk in postmenopausal women. *J Natl Cancer Inst* 2003; 95(16): 1218–26.

78. Kaye SA, Folsom AR, Soler JT, Prineas RJ, Potter JD. Associations of body mass and fat distribution with sex hormone concentrations in postmenopausal women. *Int J Epidemiol* 1991; 20(1): 151–6.

79. Abbenhardt C, Miller JW, Song X, Brown EC, Cheng T-YD, Wener MH, et al. Biomarkers of one-carbon metabolism are associated with biomarkers of inflammation in women. *J Nutr* 2014; 144(5): 714–21.

80. Crujeiras a. B, Casanueva FF. Obesity and the reproductive system disorders: epigenetics as a potential bridge. *Hum Reprod Update* 2014; 21(2): 249–61.

81. Fernández-Sánchez A, Madrigal-Santillán E, Bautista M, Esquivel-Soto J, Morales-González Á, Esquivel-Chirino C, et al. Inflammation, oxidative stress, and obesity. *Int J Mol Sci* 2011; 12(5): 3117–32.

82. Niu Y, Desmarais TL, Tong Z, Yao Y, Costa M. Oxidative stress alters global histone modification and DNA methylation. *Free Radic Biol Med* 2015; 82: 22–8.

83. Xin Y-J, Yuan B, Yu B, Wang Y-Q, Wu J-J, Zhou W-H, et al. Tet1-mediated DNA demethylation regulates neuronal cell death induced by oxidative stress. *Sci Rep* 2015; 5: 7645.

84. Kowluru R a., Kowluru A, Mishra M, Kumar B. Oxidative stress and epigenetic modifications in the pathogenesis of diabetic retinopathy. *Prog Retin Eye Res* 2015; 48: 40-61.

85. Einspahr JG, Alberts DS, Gapstur SM, Bostick RM, Emerson SS, Gerner EW. Surrogate end-point biomarkers as measures of colon cancer risk and their use in cancer chemoprevention trials. *Cancer Epidemiology Biomarkers and Prevention* 1997; 6(1): 37–48.

86. Zhao F. Surrogate End Points and Their Validation in Oncology Clinical Trials. *J Clin Oncol* 2016; 34(13): 1–3.

87. Weintraub WS, Lüscher TF, Pocock S. The perils of surrogate endpoints. *Eur Heart J* 2015; 36(33): 2212–8.

88. Baba Y, Ishimoto T, Kurashige J, Iwatsuki M, Sakamoto Y, Yoshida N, et al. Epigenetic field cancerization in gastrointestinal cancers. *Cancer Lett* 2016; 375: 360–6.

89. Fleming TR, Powers JH. Biomarkers and surrogate endpoints in clinical trials. *Stat Med* 2012; 31(25): 2973–84.

90. Tworoger SS, Hankinson SE. Use of biomarkers in epidemiologic studies: Minimizing the influence of measurement error in the study design and analysis. *Cancer Causes Control* 2006; 17(7): 889–99.

91. Irahara N, Nosho K, Baba Y, Shima K, Lindeman NI, Hazra A, et al. Precision of pyrosequencing assay to measure LINE-1 methylation in colon cancer, normal colonic mucosa, and peripheral blood cells. *J Mol Diagn* 2010; 12(2): 177–83.

92. Hoad G, Harrison K. The Design and Optimization of DNA Methylation Pyrosequencing Assays Targeting Region-Specific Repeat Elements. *Methods Mol Biol* 2015.

93. Byun HM, Nordio F, Coull BA, Tarantini L, Hou L, Bonzini M, et al. Temporal stability of epigenetic markers: Sequence characteristics and predictors of short-term DNA methylation variations. *PLoS One* 2012; 7(6): 1–9.

94. El-Maarri O, Walier M, Behne F, van Üüm J, Singer H, Diaz-Lacava A, et al. Methylation at global LINE-1 repeats in human blood are affected by gender but not by age or natural hormone cycles. *PLoS One* 2011; 6(1): e162525.

95. 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; 130(4): 234–9.

96. 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; 23(54): 8841–6.

97. Lokk K, Modhukur V, Rajashekar B, Märtens K, Mägi R, Kolde R, et al. DNA methylome profiling of human tissues identifies global and tissue-specific methylation patterns. *Genome Biol* 2014; 15(4): R54.

98. Houseman EA, Kim S, Kelsey KT, Wiencke JK. DNA Methylation in Whole Blood: Uses and Challenges. *Curr Environ Heal Reports* 2015; 2(2): 145–54.

99. Wu HC, Delgado-Cruzata L, Flom JD, Kappil M, Ferris JS, Liao Y, et al. Global methylation profiles in DNA from different blood cell types. *Epigenetics* 2011; 6(1): 76–85.

100. Brennan K, Flanagan JM. Is there a link between genome-wide hypomethylation in blood and cancer risk? *Cancer Prev Res* 2012; 5(12): 1345–57.

101. Canadian Society's Advisory Committee on Cancer Statistics. *Canadian Cancer Statistics 2015*. Toronto, ON: Canadian Cancer Society; 2015.

102. Stuckey A. Breast cancer: epidemiology and risk factors. *Clin Obstet Gynecol* 2011; 54(1): 96–102.

103. Burris HH, Baccarelli AA. Environmental epigenetics: From novelty to scientific discipline. *J Appl Toxicol* 2014; 34(2): 113–6.

104. Engler-Chiurazzi EB, Brown CM, Povroznik JM, Simpkins JW. Estrogens as Neuroprotectants: Estrogenic Actions in the Context of Cognitive Aging and Brain Injury. *Prog Neurobiol* 2016; In press.

# Chapter 3

# Study Design and Methods

## Introduction

The following chapter provides an overview of the specific objectives addressed in the two manuscripts. Specifically, objectives 1 to 3 were addressed in manuscript 1 (Chapter 4) and objectives 4 to 6 were addressed in manuscript 2 (Chapter 5). This section also includes aadditional information pertaining to the study population, design, and methods that were not fully described in the two manuscripts for the sake of brevity.

## Study Objectives and Hypotheses

**Objective 1:** Describe the association between measures of adiposity and DNA methylation within LINE-1 and Alu repeats in a population of healthy, inactive, postmenopausal women. **Hypothesis:** All measures of adiposity will be negatively associated with both LINE-1 and Alu methylation. **Objective 2:** Assess the association between weight changes over time and LINE-1 and Alu methylation. **Hypothesis:** Weight gain and weight cycling will be negatively related to the methylation outcomes. Weight loss will be positively associated with LINE-1 and Alu methylation. **Objective 3**: Determine if the association between adiposity and repetitive element DNA methylation differs according to the timing (lifetime vs. current), the type (intraabdominal vs. subcutaneous fat), the distribution (central vs. overall body fat), or the conceptualization (absolute vs. relative measure) of the exposure. **Hypothesis:** Historical, intraabdominal, central, and relative measures of adiposity will have a larger association DNA methylation than current, subcutaneous, overall, and absolute measures of adiposity.

**Objective 4:** Examine the association between endogenous sex hormones and repetitive element DNA methylation in a population of healthy, inactive, postmenopausal women. **Hypothesis:** Exposure to endogenous sex hormones will be negatively associated with levels of LINE-1 and Alu methylation. **Objective 5:** Compare the association between different types (androgens vs. estrogens) and durations (current vs. cumulative estrogen exposure) with endogenous sex hormone exposure on LINE-1 and Alu methylation. **Hypothesis:** Estrogens will be more strongly associated with LINE-1 and Alu methylation than androgens and cumulative estrogen exposure will be more strongly related to these outcomes than current estrogen measures. **Objective 6:** Determine if the total lifetime number of menstrual cycles adequately summarizes the independent epigenetic effects of age at menarche, age at menopause, and parity. **Hypothesis:** The shape and direction of the association between total lifetime number of menstrual cycles and the DNA methylation outcomes will be similar to that of age at menarche, age at menopause, and parity albeit with a slightly larger magnitude of effect.

**Study Design**

A cross-sectional study was conducted within an ancillary study of the Alberta Physical Activity and Breast Cancer Prevention (ALPHA) Trial (ClinicalTrials.gov identifier: NCT00522262). (ref. 1) The ALPHA Trial was a two-centered, parallel arm, randomized controlled trial that assessed the impact of a one-year aerobic exercise intervention relative to a regular inactive lifestyle on various biological factors thought to mediate the relation between physical activity and breast cancer. Participants were randomized to either the intervention or control arm. The intervention consisted of five 45 minute sessions per week of moderate to vigorous aerobic activity for one year. The control group was asked to maintain their regular

activity levels over the same length of time. Assessments were made at baseline, 6 months, and 12 months. In the ALPHA Trial, the outcomes of interest were sex hormone concentrations, adiposity, breast density, inflammation, and insulin resistance. In the ancillary study, blood samples collected from the participants were reanalyzed to obtain measures of DNA methylation, telomere length, and oxidative stress. In the current study, we relied solely upon measures taken at baseline.

**Study Participants**

Letters of invitation to participate in the ALPHA Trial were sent to women who partook in the Alberta Breast Screening Program Screen Test. Study participants were also recruited from the Alberta Family Practice Research Network and through media campaigns. Eligibility was assessed at various stages of entry into the study (before contacting the participant, at an information session, after the participant expressed an interest in enrollment, and at baseline) using a variety of methods (breast cancer screening database, telephone interview, introductory questionnaire, mammogram, physician approval, blood tests, and submaximal fitness test). Participants had to satisfy the following criteria at baseline in order to be eligible for entry into the ALPHA Trial:

| Inclusion Criteria | Exclusion Criteria |
|---|---|
| • Female<br>• Aged 50 to 75<br>• Postmenopausal for at least 2 years<br>• Sedentary (i.e. less than 90 minutes of recreational activity per week)<br>• Reside in Calgary or Edmonton<br>• Body mass index between 22.0 and 40.0 $kg/m^2$<br>• Breast density greater than zero percent<br>• Normal fasting blood levels of lipids, thyroid-stimulating hormone, and alanine aminotransferase | • Unable to participate in the exercise intervention<br>• Major comorbidity<br>• Previous cancer diagnosis (with the exception of non-melanoma skin cancer)<br>• Smoker<br>• Excessive alcohol consumption (i.e. more than 14 drinks/week)<br>• Hormone replacement therapy use in the past two years<br>• Taking medication that could affect the outcomes of interest |

**Data Collection**

Data collection is described in further detail in Chapters 4 and 5. Briefly, measures of adiposity were obtained using a CT scan, DXA scan, balance-beam scale, and tape measure. Sex hormone concentrations (i.e. estradiol, estrone, testosterone, androstenedione, and sex hormone binding globulin) were assessed from fasting blood samples. We estimated each participant's total lifetime number of menstrual cycles as a proxy for cumulative exposure to ovarian sex hormones using questionnaire data about age at menarche, age at menopause, parity, and gravidity. DNA methylation was analyzed from frozen baseline fasting blood samples. Covariate information was primarily captured through various questionnaires.

**Statistical Analysis**

Linear regression was used to model the association between the exposures and outcomes of interest while controlling for important confounders. The mean change in percent methylation associated with a standard deviation increase in the exposure of interest was used as a measure of

effect. This standardized effect estimate was chosen to allow for comparisons across different exposures. We used a backwards elimination procedure with a p-value of 0.20 to select which confounders to control in the final model. More detailed information pertaining to the analytic strategies is provided in Chapters 4 and 5.

Using a one-way ANOVA model, it was determined that the plate on which DNA methylation was assessed (n = 7) had a significant association with LINE-1 methylation ($F = 102.04$; $p < 0.0001$) and Alu methylation ($F = 12.27$; $p < 0.0001$). To assess the degree of clustering by plate we estimated the intraclass correlation coefficient by partitioning the variance into individual- and plate-level effects. It was determined that within plate clustering accounted for 67.72% and 29.97% of the variation in the LINE-1 and Alu methylation measures respectively. To account for clustering by plate, we modeled the seven plates using a random effect parameter with an unstructured covariance matrix in all analyses.

**Power Calculation**

PS – Power and Sample Size Calculation software Version 3.0 by Dupont and colleagues was used to compute the following *a priori* power estimate which is based on a two sample t-test comparing the mean difference in standardized DNA methylation levels between the first and fourth quartiles of the exposure of interest with an assumed sample size equal to 300. (2) At a 5% significance level, the smallest effect size that this study has an 80% probability of detecting is a 0.46 standard deviation change in LINE-1 or Alu methylation.

## Additional Considerations

*Conceptualization of DNA Methylation Outcomes*

For both LINE-1 and Alu elements, we assessed DNA methylation within three repeated CpG sites. Within the genome of a single cell, there are approximately 500,000 and 1,000,000 copies of LINE-1 and Alu elements respectively, not all of which are full length. (3) For each participant, the percent methylation for a repeated CpG site was estimated by dividing the number of methylated CpG sites by the total number of CpG sites within their DNA sample. We then estimated the overall level of LINE-1 and Alu methylation by averaging across the three CpG sites for each participant.

The three CpG sites within LINE-1 that were assessed for DNA methylation were highly correlated ($0.58 < r < 0.89$) which supports the appropriateness of averaging across these three CpG sites to derive an overall measure of LINE-1 methylation for our primary outcome. While two of the CpG sites assessed for Alu methylation were highly correlated ($r = 0.71$), one of the CpG sites was not correlated with the other two ($0.002 < |r| < 0.04$). To be consistent with the literature, we averaged across the three CpG sites to compute the overall Alu methylation percent as our primary outcome of interest. However, we also conducted a sensitivity analysis in which we ran two separate analyses assessing the effect of the exposures of interest on: 1) DNA methylation in the uncorrelated CpG site; and 2) the average DNA methylation within the two correlated CpG sites. The results from these two sensitivity analyses were not considerably different from the results of the main analyses that averaged across the three CpG sites. Therefore, we used the average of the three CpG sites for our final outcome.

*Estimation of Total Lifetime Number of Menstrual Cycles*

One of the primary exposures used in the current study is the total lifetime number of menstrual cycles which is thought to reflect cumulative exposure to ovarian sex hormones. (4) Although we had information on age at menarche, age at menopause, parity, and gravidity, we were missing several variables that would have allowed for more precise estimates of the total lifetime number of menstrual cycles such as menstrual cycle regularity and length, age at which a regular menstrual cycle pattern was established, oral contraceptive use, duration of gestation for a full term or terminated pregnancy, and the length of time it took for menstruation to return following a full term birth. To estimate the lifetime number of menstrual cycles, the following assumptions were made:

- A full term pregnancy resulted in a 49 week absence of cycles (36 weeks of absence due to gestation + 13 weeks of absence due to breastfeeding)
- A terminated pregnancy (i.e. miscarriage, abortion, or stillbirth) resulted in a 12 week absence of cycles
- All women had regular 28 day cycles throughout their menstrual span

A 36 week absence of menstrual cycles due to a full term pregnancy was chosen as it has been used in past research and is a number that coincides with the typical length of a full term pregnancy (i.e. most pregnancies last around 40 weeks and a regular menstrual cycle is around 4 weeks which would result in a 40 – 4 = 36 week absence in menstrual cycles). (4) We also assumed that each participant who gave birth would have an additional absence of menstrual cycles for 13 weeks after the pregnancy due to breastfeeding. We derived this estimate using data on the duration of exclusive breast feeding from the 2003-2004 Canadian Community Health

Survey (Cycle 2). (5) We used data on the duration of exclusive breast feeding instead of the duration of any breastfeeding because exclusive breastfeeding is more likely to prolong the absence of menstrual cycles. (7) Since the 2001-2002 Canadian Community Healthy Survey (Cycle 1) only reported the duration of breast feeding (not exclusive breast feeding), we used the 2003-2004 Canadian Community Healthy Survey which reported the proportion of mothers who exclusively breastfed by different categories of breastfeeding duration. The mean duration of breastfeeding was calculated for each category (e.g. everyone who was in the 1 week to <5 week category was assumed to have breastfed for 3 weeks) and was combined with the estimated proportion of mothers in that category to estimate the expected duration of exclusive breastfeeding as follows:

| **Duration of breastfeeding** | **Mean duration of breastfeeding** | **Proportion** |
|---|---|---|
| 0 weeks | 0 weeks | 16.1% |
| < 1 week | 0.5 weeks | 8.7% |
| 1 to < 5 weeks | 3 weeks | 12.0% |
| 5 to < 12 weeks | 8.5 weeks | 10.4% |
| 12 to < 20 weeks | 16 weeks | 25.0% |
| 20 to < 28 weeks | 24 weeks | 20.4% |
| 28 weeks to < 1 year | 40 weeks | 6.9% |
| 1+ year | 52 weeks | 0.6% |
| Expected duration of exclusive breast feeding = 0*0.161 + 0.5*0.087 + 3*0.120 + 8.5*0.104 + 16*0.25 + 24*0.204 + 40*0.069 + 52*0.006 = **13.26 weeks** | | |

We did not have information on the specific reason for the termination of a pregnancy. We assumed that stillbirths would comprise the majority of terminated pregnancies (as opposed to abortions or miscarriages) and assumed a 12 week absence in menstrual cycles for all terminated pregnancies as previously done. (4) A regular 4 week menstrual cycle was chosen since it is the most usual pattern found in the population and it eases the calculation of the number of menstrual cycles. Under these assumptions the total lifetime number of menstrual cycles was calculated as follows:

**Total lifetime number of menstrual cycles** = [(age at menopause – age at menarche) × 52.1775 weeks – (number of live births) × 49 weeks – (number of terminated pregnancies) × 12 weeks] / 4 weeks.

*Derived Weight Change Variables*

In addition to the adiposity measures that were captured at baseline, we wanted to explore changes in adiposity over time. In a baseline health questionnaire, participants reported their weight since age 20 in 10 year intervals. We considered the following variables which have been used previously in the literature: (7,8)

- Weight gain since age 20: baseline weight – weight at age 20
- Adult weight gain: baseline weight – smallest adult weight
- Maximum adult weight difference: largest adult weight – smallest adult weight
- Weight loss: largest adult weight - baseline weight
- Weight cycling: absolute sum of weight difference between decades

Weight gain, adult weight gain, and maximum adult weight difference were highly correlated ($0.93 < r < 0.98$). We used weight gain since age 20 as our measure of weight gain as it is used more frequently in the literature. The remaining three weight change variables (weight gain, weight loss, and weight cycling) had weak to moderate correlations with one another ($0.12 < |r| < 0.51$).

*Adjusting Dietary Folate Intake for Total Energy Intake*

In the current investigation, we considered dietary folate intake as a potential confounder and effect modifier. A more accurate and precise estimate of the relation between a dietary micronutrient and an exposure or outcome can be obtained by adjusting for total energy intake (i.e. total energy intake can confound the relation between a dietary macro- or micro-nutrient and another variable of interest). (9)

There are four primary approaches used to adjust for total caloric intake when assessing a dietary nutrient. (10) One method is to include nutrient intake and total energy intake as separate variables in a multivariable regression model. Alternatively, a variable representing the amount of energy from nutrients other than the one under scrutiny in place of total energy intake can be included in the multivariable regression model. Another option is to express nutrient intake as a percentage of the total energy intake in the multivariable model. Finally, nutrient intake can be regressed on total caloric intake and the resulting residuals can be used in place of the original nutrient variable. Since there will be no association between these nutrient residuals and total energy intake, it is not necessary to adjust for total energy intake in a multivariable model.

We used the "nutrient residual" approach to adjust for total energy intake in the current study for the following reasons: 1) it requires the inclusion of one covariate rather than two which is statistically more efficient and reduces the risk of multicollinearity; 2) it expresses the amount of energy-adjusted folate intake in a single variable which helps in the assessment of effect modification by making it easier to divide the participants into strata by folate consumption; 3) adiposity is one of the primary exposures of this investigation and it is closely linked with total energy intake such that total energy intake would exist largely on the causal

pathway and its inclusion into a multivariable regression model would likely result in a high degree of multicollinearity with the adiposity exposures.

Using a simple linear regression model, we estimated the relation between total energy intake and intake of dietary folate equivalents (1 microgram (mcg) dietary folate equivalent = 1 mcg of food folate or 0.6 mcg of folic acid from fortified foods) amongst the study participants as follows: dietary folate equivalents (mcg) = 62.98437 + 0.23179 × total energy intake (kcal). Using this model, we estimated that an individual with a total energy intake of 2000 kcal per day (a statistically arbitrary constant) would have an average intake of dietary folate equivalents of 526.56 mcg per day. Each participant's energy-adjusted folate intake was estimated as their nutrient residual + 526.56. This final variable can be interpreted as the daily intake of dietary folate equivalents in micrograms standardized to a caloric intake of 2000 kcal per day. The unadjusted folate intake variable was strongly correlated (r = 0.73) with total energy intake whereas the nutrient residual variable, as expected, had no correlation with total energy intake (r = 0.00).

*Assessing Linearity*

Within a bivariate analysis, we assessed the shape of the relation between all of our explanatory variables (i.e. all exposures and covariates) and outcome variables. We assessed the shape of the relation qualitatively by examining a plot of the residuals versus the explanatory variable in which a systematic "U-shape" would be indicative of a departure from linearity. We also detected non-linearity quantitatively by assessing the significance of a polynomial term in a model that already contains the explanatory variable of interest (i.e. the significance of β2 in the following model: $y = \beta_0 + \beta_1 * x + \beta_2 * x^2$).

## References


1. Friedenreich CM, Woolcott CG, McTiernan A, Ballard-Barbash R, Brant RF, Stanczyk FZ, Terry T, Boyd NF, Yaffe MJ, Irwin ML, et al. Alberta physical activity and breast cancer prevention trial: Sex hormone changes in a year-long exercise intervention among postmenopausal women. *J Clin Oncol* 2010; 28: 1458–66.
2. Dupont WD and Plummer WD. Power and Sample Size Calculations: A Review and Computer Program. *Control Clin Trials* 1990; 11: 116 – 128.
3. Miousse IR, Chalbot M-CG, Lumen A, Ferguson A, Kavouras IG, Koturbash I. Response of Transposable Elements to Environmental Stressors. *Mutat Res* 2015; 765: 19–39.
4. Chavez-MacGregor M, Elias SG, Onland-Moret NC, van der Schouw YT, Van Gils CH, Monninkhof E, Grobbee DE, Peeters PHM. Postmenopausal Breast Cancer Risk and Cumulative Number of Menstrual Cycles. *Cancer Epidemiol Biomarkers Prev* 2005; 14: 799–804.
5. Statistics Canada. Canadian Community Health Survey, 2003: Cycle 2.1: General File [public-use microdata file]; 2004.
6. Gray RH, Campbell OM, Apelo R, Eslami SS, Zacur H, Ramos RM, Gehret JC, Labbok MH. Risk of ovulation during lactation. *Lancet* 1990; 335: 25–9.
7. Friedenreich CM, Courneya KS, Bryant HE. Case-control study of anthropometric measures and breast cancer risk. *Int J Cancer* 2002; 99: 445–52.
8. Trentham-Dietz A, Newcomb PA, Egan KM, Titus-Ernstoff L, Baron JA, Storer BE, Stampfer M, Willett WC. Weight change and risk of postmenopausal breast cancer (United States). *Cancer Causes Control* 2000; 11: 533–42.

9. Willett WC, Howe GR, Kushi L. Adjustment for total energy intake in epidemiologic studies. *Am J Clin Nutr* 1997; 65(4): 1220S – 1228S.

10. Mackerras D. Energy adjustment: The concepts underlying the debate. *J Clin Epidemiol* 1996; 49: 957–62.

# Chapter 4

# Associations between Adiposity and Repetitive Element DNA Methylation in Healthy Postmenopausal Women

## INTRODUCTION

More than half of the human genome is made up of recurring copies of non-coding base sequences known as repetitive elements. (1) Transposable elements – the most common type of repetitive element – are now recognized as having an important role in the maintenance of genomic integrity and the regulation of gene expression. (2) Two major types of transposable elements active in human DNA are long interspersed nuclear element-1 (LINE-1) and Alu elements. By means of retrotransposition, LINE-1 and Alu repeats can become inserted into new genomic locations which can have harmful biological effects such as chromosomal instability or the aberrant expression of tumour suppressor and oncogenes. (3,4)

Epigenetic processes are the primary means through which the expression of these transposable elements are regulated. (5) In particular, DNA methylation serves as a mechanism to supress these potentially harmful retrotransposons. (5). In the normal cellular state, LINE-1 and Alu repeats typically have high overall levels of methylation relative to other genomic regions. (6) Lower levels of DNA methylation within these repetitive element regions may serve as a mechanism whereby healthy individuals become epigenetically predisposed to developing cancer and other illnesses due to the aberrant expression of LINE-1 and Alu elements. (7–10)

A growing body of epidemiologic evidence has linked hypomethylation of LINE-1 and Alu to an increased risk of developing several types of cancer including breast, (11) colorectal, (9,12) lung, (9) cervical, (13) ovarian, (14) bladder, (15) head and neck, (16) liver, (17) and

stomach cancer. (18) Emerging evidence has also implicated the role of repetitive element methylation in the etiology of other illnesses such as cardiovascular disease. (19)

Unlike genetic predispositions, epigenetic aberrations such as low levels of LINE-1 and Alu methylation are potentially modifiable as evidenced by their association with various demographic, environmental, and behavioural risk factors. (20) Despite their suspected role in the etiology of cancer and other illnesses, however, the determinants of these epigenetic factors and the timing during which they are most susceptible to change remains largely unexplained. (21)

Obesity continues to be an important public health priority because of its association with a wide variety of chronic diseases and its continued increase in prevalence worldwide. (22) Although adiposity is known to play a role in the etiology of various types of cancer and other illnesses, the underlying biological mechanisms have yet to be fully understood. (23,24) The aim of the current study is to determine the extent to which historical and current levels of adiposity and measures of weight change over time are associated with repetitive element DNA methylation amongst healthy postmenopausal women.

## SUBJECTS AND METHODS

### *Study Design and Participants*

The current investigation is a cross-sectional study nested within an ancillary study of the Alberta Physical Activity and Breast Cancer Prevention (ALPHA) Trial (ClinicalTrials.gov identifier: NCT00522262) which has been previously described. (25) Briefly, participants were 320 postmenopausal women aged 50 to 74 years who were residents of Calgary or Edmonton, were not current smokers or excessive drinkers (i.e. no more than 2 drinks per day), had no major

co-morbidities or previous cancer diagnoses, were inactive, and were not currently taking hormone replacement therapy or medications that could affect estrogen metabolism. The study subjects were enrolled in a year-long randomized exercise intervention that examined the effects of aerobic activity on biomarkers believed to be involved in the etiology of breast cancer. The current study employs various measurements that were taken at baseline.

### *Exposure and Covariate Data Collection*

Adiposity was captured through a variety of methods as previously described. (26) Briefly, body fat percent and fat mass were quantified using a dual energy x-ray absorptiometry (DXA) scan. Total fat and lean body mass were measured using the Hologic QDR 4500W scanner (Hologic Inc., Bedford, MA, USA) at the Calgary study site and the Lunar Prodigy scanner (Lunar General Electric Medical Systems, Madison, WI, USA) at the Edmonton site. Various measures of central adiposity including intraabdominal and subcutaneous fat were measured at the umbilicus using a computed tomography (CT) scan. A PQ5000 VisionMaster CT scanner (Marconi, Cleveland, OH, USA) and MX8000 multi-slice CT Scanner (Phillips Medical Systems, Cleveland, OH, USA) was used at the Calgary and Edmonton study sites respectively.

Each participant's waist and hip circumference as well as weight and height were measured by trained specialists using a tape measure, balance beam scale, and stadiometer. These measures were subsequently used to calculate each participants' waist-to-hip ratio and body mass index (BMI, $kg/m^2$). Weight at different ages was recorded in a self-administered baseline health questionnaire (BHQ) in ten-year intervals starting at age 20. These variables were used to derive the following weight change variables: weight gain (current weight – weight at

age 20), weight loss (largest adult weight – current weight), and weight cycling (sum of absolute weight change between decades).

Demographic information, medical and reproductive factors, menstrual history, smoking history, and hormone replacement therapy use were also captured in the BHQ. Total past year physical activity in metabolic equivalent (MET) hours per week was assessed using the validated Past Year Total Physical Activity Questionnaire. (27) Past year diet and alcohol intake were measured using a Canadian adaptation of the U.S. National Cancer Institute's Diet History Questionnaire Version 1. (ref. 28) Energy-adjusted folate intake (i.e. the amount of folate intake standardized to a total-caloric intake of 2000 kcal) was estimated using a nutrient residual model. (29)

We hypothesized that increased number of ovulatory cycles and resulting exposure to high estrogen levels is the primary mechanism through which age at menarche, age at menopause, and parity influences repetitive element DNA methylation. To create a summary index of the influence of these factors on cumulative estrogen exposure, these variables were combined into a single variable – total lifetime number of menstrual cycles which is believed to represent cumulative estrogen exposure and has been previously linked with postmenopausal breast cancer risk. (30)

Assuming a regular four week menstrual cycle, a 49 week absence of menstrual cycles per full term pregnancy, and a 12 week absence in menstrual cycles per miscarriage, abortion, or stillbirth, the total lifetime number of menstrual cycles was estimated as follows: [(age at menopause – age at menarche) × 52.1775 weeks – (number of live births) × 49 weeks – (number of terminated pregnancies) × 12 weeks] / 4 weeks.

### *Assessment of Repetitive Element Methylation*

A bisulfite PCR pyrosequencing assay was used to analyze repetitive element DNA methylation. Forty mL of fasting blood was taken from each participant at baseline and stored at -86oC as previously described. (25) Purification of nucleic acid was automated on the Hamilton STARlet liquid handling instrument (Hamilton Robotics Inc., Reno, USA). The Macherey-Nagel NucleoMag Blood 200 µl kit (Macherey-Nagel GmbH & Co. KG, Düren, Germany) was used to extract DNA from each participant's buffy coat solution and DNA samples were quantified using the Qubit dsDNA HS Assay Kit (ThermoFisher Scientific Inc., Waltham, USA). Of the 320 baseline buffy coat samples extracted, 298 yielded sufficient DNA quantity for the methylation assay. DNA methylation analysis was conducted at the McGill University and Génome Québec Innovation Centre (Montréal, Canada). Bisulfite conversion of the DNA samples was accomplished using the EZ-96 DNA Methylation-Gold Kit (Zymo Research, Irvine, USA, Catalog No. D5007) as per the manufacturer's protocol. Regions covering LINE-1 and Alu repeats were amplified using the HotStarTaq DNA Polymerase kit (Qiagen, Hilden, Germay). We used the same oligonucleotide primers reported by Kile et al. (31). Pyrosequencing of these PCR products was done using the PyroMark Q24 (Qiagen, Hilden, Germay) which quantified the percentage of methylation at three distinct CpG sites for both LINE-1 and Alu elements. The overall level of methylation for both outcomes was defined as the average percentage of methylation across the three CpG locations. Levels of methylation were measured on eight duplicate samples to assess the internal reliability of this method. Inter-batch coefficients of variation for the LINE-1 and Alu methylation measures were 1.83% and 2.19% respectively. Out of the 298 samples analyzed, the DNA methylation assay failed for two samples and three

samples did not pass internal quality control checks for both analyses. These observations were removed from the data set, resulting in a sample size of 293.

### *Statistical Analyses*

The relation between adiposity and repetitive element DNA methylation was assessed using linear regression. Continuous measures were standardized so that coefficients could be interpreted as the change in methylation percent per one standard deviation increase in the adiposity measure. Ethnicity, smoking status, age at first birth, and the presence of comorbidities were modeled as categorical variables. The remaining covariates and measures of adiposity were modeled as continuous variables with the exception of lifetime number of menstrual cycles which was modeled as a categorical variable with quintile cut-points since bivariate analyses suggested a curvilinear relationship with LINE-1 methylation. For each methylation outcome, a backwards elimination approach using a liberal p-value of 0.20 was used to identify a parsimonious model with which to control for confounding. A random effects parameter with an unstructured covariance structure representing the seven plates on which the measurement of DNA methylation took place was included in all analyses to account for intra-assay clustering. Based on findings from previous studies and their established link with one-carbon metabolism, potential interactions with dietary folate and alcohol intake were explored. (32) Four participants were missing information on important covariates and were therefore excluded from the analysis resulting in a final sample size of 289. All statistical tests were two-sided and were carried out at a 5% significance level using SAS version 9.4.

## RESULTS

The 289 women included in the final analysis had a mean age of 61 years (range = 50 to 75 years). The vast majority (91%) of these participants were Caucasian. Over the course of the year prior to their enrollment in the ALPHA Trial, participants had a median total physical activity level of 108 MET-hours per week, a median alcohol consumption of 1.8 grams per day, and a median energy-adjusted dietary folate intake of 510 micrograms (mcg) of dietary folate equivalents (DFE) per day (1 mcg DFE is equal to 1 mcg of food folate or 0.6 mcg of folic acid from fortified foods) (Table 1). Approximately one third (33%) of the participants were former smokers and the rest were never smokers. Around 38% of the subjects had at least one minor comorbidity. On average, participants gave birth to two children and most had their first child before the age of 25. Only 8.65% of the participants were nulliparous. The mean age of menarche and menopause was 13 and 48 years respectively. The median lifetime number of menstrual cycles was 456 cycles for this group of women.

In general, study participants were overweight with a mean body fat percent of 42.5% and BMI of 29.3 $kg/m^2$ (Table 2). Since the age of 20, the subjects gained an average of 21 kg. Amongst all of the adiposity variables, the strongest correlation was between total abdominal fat area and subcutaneous fat area (Pearson's $r = 0.96$) and the weakest correlation was between weight loss and hip circumference ($r = -0.001$). In general, weight loss ($0.001 < |r| < 0.35$), weight change ($0.11 < |r| < 0.56$), weight at age 20 ($0.08 < |r| < 0.47$), and waist-to-hip ratio ($0.02 < |r| < 0.64$) were not strongly correlated with the other adiposity measures. With the exception of these four variables, all of the other adiposity measures were moderately to strongly correlated ($0.54 < |r| < 0.96$). Observed levels of LINE-1 and Alu methylation were normally distributed with a mean of 72.73% (Standard Error (SE): 0.90) and 19.09% (SE: 0.13) respectively after

adjusting for intra-plate variability. Our measures of LINE-1 and Alu methylation were significantly weakly correlated ($r = 0.14$; p-value = 0.02).

As per the backwards elimination procedure, LINE-1 methylation analyses adjusted for ethnicity and lifetime number of menstrual cycles and Alu methylation analyses controlled for the presence of comorbidities. Adjusting for age at menarche, age at menopause, and parity as separate variables rather than as a single combined variable as represented by the lifetime number of menstrual cycles did not meaningfully alter the effect estimates (data not shown).

All of the adiposity variables were significantly positively associated with LINE-1 methylation after controlling for confounding ($p < 0.05$) with the exception of subcutaneous fat area ($p = 0.08$), waist-to-hip ratio ($p = 0.53$), weight loss ($p = 0.93$), and weight change ($p = 0.11$) (Table 3). Hip circumference appeared to have the largest association with LINE-1 methylation in which a one standard deviation (9.90 cm) increase in hip circumference was associated with a mean increase of 0.33% in LINE-1 percent methylation after controlling for confounding. None of the adiposity variables were associated with Alu methylation in the multivariable model (Table 3). Neither dietary folate nor alcohol consumption significantly modified the effect of any of the adiposity metrics on either DNA methylation outcome (data not shown).

## DISCUSSION

In the current study, various measures of adiposity were found to be positively related to LINE-1 methylation. No associations with Alu methylation were observed. While these differences may point towards underlying biological differences between LINE-1 and Alu

methylation, (2) the null findings with respect to the latter outcome may be attributable to the lack of variability in our measure of Alu methylation.

The strength of the relation between adiposity and LINE-1 methylation was generally consistent across the different adiposity measures. While we did not assess the statistical significance of the differences in effect sizes, there exists small qualitative differences between the exposures that are worthy of mention. In particular, total body fat volume (e.g., total fat mass, overall weight, or waist circumference) appeared to have a slightly larger effect on LINE-1 methylation compared to the relative adiposity measures (e.g. body fat percent, BMI, or waist-to-hip ratio).

In the current study, the effect of adiposity during both early and late adulthood was assessed. Compared to weight at age 20, we found that current weight had a slightly larger effect on LINE-1 methylation. While this finding may indicate that more recent adiposity levels have a greater influence on LINE-1 methylation than adiposity at earlier ages, the different magnitudes of effect may be a result of greater misclassification in the historical self-reported weight variable. Nonetheless, our results suggest that DNA methylation remains susceptible to influence from lifestyle factors later in life.

Despite statistically significant relations with waist circumference and hip circumference, LINE-1 methylation was not associated with waist-to-hip ratio. Although used as a measure of adiposity, waist-to-hip ratio is more indicative of body type (e.g. “apple-shaped” or “pear-shaped”). Given the presence of a significant association found with measures of central body fat captured by CT scan, these findings imply that LINE-1 methylation is affected by body fat and not body shape.

Intra-abdominal fat had a slightly larger effect on LINE-1 methylation compared to subcutaneous abdominal fat, the results of which did not reach statistical significance. These findings indicate that fat deposits in the former region are more responsible for driving the association between abdominal adiposity and DNA methylation which is consistent with the literature showing a larger health effect of visceral fat compared to subcutaneous fat. (33)

As with the findings from the baseline adiposity measures, weight gain since age 20 was positively associated with LINE-1 methylation. Neither weight cycling nor weight loss was related to LINE-1 methylation. These null findings may be a result of the lack of statistical power caused by the small sample size and narrow range of values observed in this study population.

It is important to recognize that the foregoing associations are not entirely independent as evidenced by the moderate to strong correlations between most of our adiposity measures. As such, it is difficult to separate the effect of any one adiposity measure from that of another.

We were able to evaluate the effects of several potential confounding factors in this study but did not find evidence for major confounding. Specifically, adjusting for the effects of age, physical activity, smoking, hormone replacement therapy use, and alcohol intake in the regression analyses did not meaningfully alter the effect estimates. This finding may be partially attributable to the exclusion criteria used in the underlying study which reduced variability in these potential confounders. Although folate is a suspected determinant of DNA methylation, (34) it may have not been an important predictor of our outcomes because the participants had adequate intake of dietary folate.

Epidemiologic evidence to date regarding the relation between obesity and repetitive element methylation is inconsistent. (35) The current findings are supported by two separate

weight loss intervention studies that found a positive relation between body fat and LINE-1 methylation. (36,37) A positive association between BMI and LINE-1 methylation has also been observed in cross-sectional analyses conducted in different study settings. (38,39) However, no association between adiposity and LINE-1 methylation were reported by one study using similar adiposity metrics, (40) by one weight loss intervention on a similar study population of a comparable size (i.e. three hundred overweight/obese postmenopausal women), (41) and in one large pooled-analysis. (42) In contrast, a negative relation between adiposity and LINE-1 methylation has been reported in younger, premenopausal populations. (32,43) While fewer studies have looked at the relation between body fat and Alu methylation, similar inconsistencies exist in the literature. (35) For example, investigations into the relation between BMI and Alu methylation have found negative, (44) curvilinear (45), and null associations. (42) Differences in terms of the gender distribution and menopausal status of the study populations as well as the different methods of DNA methylation assessment and adjustment for confounders may account for discrepancies in the literature with respect to the direction and significance of the observed effect. (35)

Given the current state of knowledge regarding the relation between LINE-1 methylation and disease risk, the clinical significance of these findings remains unclear. Our findings suggest that adiposity confers a protective epigenetic effect in postmenopausal women which is not consistent with previous research that found associations between increased levels of adiposity and cancer risk. (24) The epigenetic benefits reported in the current study may be outweighed by the harmful impact that adiposity has on cancer risk through other biological mechanisms which include an effect on hormonal, insulin, or inflammatory pathways. (24) Hence, the relation between adiposity and cancer might be predominantly mediated by mechanisms other than

repetitive element DNA methylation in postmenopausal women. Another possible explanation is that, in addition to low levels of repetitive element methylation, aberrantly high levels of repetitive element methylation may confer an elevated risk of cancer. This speculation is substantiated by studies that have found statistically significant positive (46,47) and curve-linear associations (48) between LINE-1 methylation and cancer risk.

The effect sizes we observed ranged from increases in mean LINE-1 percent methylation of 0.20% to 0.33% per standard deviation in the adiposity exposure. Although some epidemiologic studies suggest that small changes in percent LINE-1 methylation (~1%) may confer meaningful increases in cancer risk (11), a better understanding of the relation between LINE-1 methylation and disease risk is needed to assess the clinical relevance of these results.

The potential limitations of this study need to be recognized. We assessed DNA methylation in peripheral leukocytes. Changes in adiposity have been shown to impact the distribution of white blood cell types (49) which, in turn, has been shown to affect measures of LINE-1 methylation. (42) We were unable to adjust for the proportion of white blood cells in the current study and we cannot rule out the possibility that our results are confounded by cell type. However, one prior study found that the distribution of white blood cells did not confound the relation between BMI and LINE-1 methylation. (42) Therefore, we expect the effects of unmeasured confounding to be minimal. In addition to cell-specificity, patterns of DNA methylation are tissue specific. (50,51) The extent to which blood measures of repetitive element methylation correspond to measures done in target tissues has yet to be established. (52) Therefore, the current findings may not be generalizable to other tissues of interest. Given the cross-sectional setting of the current investigation, there may be concern regarding the directionality of the observed relations. Since significant relations of a similar magnitude and in

a similar direction were detected with the historical measures of body fat, it can be inferred that more recent adiposity measures preceded the changes in LINE-1 methylation. Finally, the study sample consisted primarily of healthy postmenopausal Caucasian women who volunteered for a year-long randomized exercise trial. As such, these findings may not be generalizable to men, premenopausal women, unhealthy populations, or to individuals of different racial backgrounds.

Despite these limitations, this study has several noteworthy strengths. Within the context of repetitive element DNA methylation studies, this investigation has used the most comprehensive list of body fat metrics to date. We examined both lifetime and current as well as overall and relative measures of obesity in addition to measures of body fat distribution and weight changes over time. Furthermore, the methods used to measure adiposity are among the highest quality used in health research. Nearly every repetitive element methylation study to date has focused on BMI (35) which is an imperfect measure of obesity. (53) Only two previous studies have considered the impact of body fat percent on LINE-1 methylation, both of which were quantified using bioelectrical impedance analysis (BIA). (32,40) The current study measured body fat percent using a DXA scan which is superior to BIA methods. (54) Thus, in addition to its breadth, this study used measures of body fat that are more accurate and precise to those used in the existing literature. Many prior studies examining this relation failed to adequately control for confounding. (35) In contrast, information was collected on several suspected determinants of DNA methylation so that a complete assessment of confounding was possible in the current investigation.

In conclusion, increased levels of adiposity were positively associated with higher levels of LINE-1 methylation amongst healthy postmenopausal women. Alu methylation does not appear to be related to adiposity in this population. While not consistent with the literature

linking excess body fat to poor health outcomes, these results add to the growing body of literature seeking to understand the extent to which DNA methylation can be modified through lifestyle choices and help to broaden our understanding of the health consequences associated with obesity. These findings also suggest that DNA methylation is susceptible to behavioural influence in older populations. To better understand the relation between obesity and repetitive element DNA methylation, future research should consider menopausal status and dietary factors as potential effect modifiers, include adiposity measures other than BMI or waist-to-hip ratio such as weight or waist circumference, and explore the foregoing relations across multiple time points.

## ACKNOWLEDGEMENTS

The ALPHA Trial was originally funded by a research grant from the Canadian Breast Cancer Research Alliance and the ancillary study is funded by grants from the Canadian Institutes of Health Research (CIHR) and the Canadian Cancer Society. Funding for the DNA extraction was provided by an Alberta Cancer Foundation grant for the Biospecimen Processing Unit within the Translational Laboratories (TL) at the Tom Baker Cancer Center (TBCC). Devon Boyne is supported by funding from the Queen's University Terry Fox Foundation Training Program in Transdisciplinary Cancer Research in partnership with CIHR and an Ontario Graduate Scholarship. Christine Friedenreich was supported by an Alberta Innovates Health Solutions Health Senior Scholar Award and by the Alberta Cancer Foundation Weekend to End Women's Cancers Breast Cancer Chair. Kerry Cournyea is supported by the Canada Research Chairs Program. The ALPHA Trial study set-up was performed by Kim van der Hoek and Marla Orenstein. The Study Coordinators were Rosemary Crosby and Ame-Lia Tamburrini. Data preparation for this ancillary study was done by Qinggang Wang. DNA extraction and quantification was performed by Angela Chan and Michelle Dean in the TL. Nigel Brockton provided advice on the DNA extraction protocols developed for this study.

## CONFLICTS OF INTEREST

No conflicts of interest are declared by the authors.

## References


1. De Koning AJ, Gu W, Castoe TA, Batzer MA, Pollock DD. Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet* 2011; 7(12): e1002384.
2. Miousse IR, Chalbot MCG, Lumen A, Ferguson A, Kavouras IG, Koturbash I. Response of Transposable Elements to Environmental Stressors. *Mutat Res Mutat Res* 2015; 765: 19–39.
3. Kemp JR, Longworth MS. Crossing the LINE Toward Genomic Instability: LINE-1 Retrotransposition in Cancer. *Front Chem* 2015; 3: 68.
4. Ade C, Roy-Engel AM, Deininger PL. Alu elements: An intrinsic source of human genome instability. *Curr Opin Virol* 2013; 3(6): 639–45.
5. Bernstein BE, Meissner A, Lander ES. The Mammalian Epigenome. *Cell* 2007; 128(4): 669–81.
6. Rollins RA, Haghighi F, Edwards JR, Das R, Zhang MQ, Ju J, et al. Large-scale structure of genomic methylation patterns Large-scale structure of genomic methylation patterns. *Genome Res* 2005; 16(2): 157–63.
7. Matsuda Y, Yamashita S, Lee YC, Niwa T, Yoshida T, Gyobu K, et al. Hypomethylation of Alu repetitive elements in esophageal mucosa, and its potential contribution to the epigenetic field for cancerization. *Cancer Causes Control* 2012; 23(6): 865–73.
8. Miousse IR, Koturbash I. The Fine LINE: Methylation Drawing the Cancer Landscape. *Biomed Res Int* 2015; 2015:1–8.
9. Zhu ZZ, Sparrow D, Hou L, Tarantini L, Bollati V, Litonjua AA, et al. Repetitive element hypomethylation in blood leukocyte DNA and cancer incidence, prevalence, and

mortality in elderly individuals: The Normative Aging Study. *Cancer Causes Control* 2011; 22(3): 437–47.

10. Rodríguez-Paredes M, Esteller M. Cancer epigenetics reaches mainstream oncology. *Nat Med 2011*; 17(3): 330–9.

11. Deroo L a, Bolick SCE, Xu Z, Umbach DM, Shore D, Weinberg CR, et al. Global DNA methylation and one-carbon metabolism gene polymorphisms and the risk of breast cancer in the Sister Study. *Carcinogenesis* 2014; 35(2): 333–8.

12. King WD, Ashbury JE, Taylor SA, Tse MY, Pang SC, Louw JA, et al. A cross-sectional study of global DNA methylation and risk of colorectal adenoma. *BMC Cancer* 2014; 14(1): 488.

13. Piyathilake CJ, Macaluso M, Alvarez RD, Chen M, Badiga S, Siddiqui NR, et al. A higher degree of LINE-1 methylation in peripheral blood mononuclear cells, a one-carbon nutrient related epigenetic alteration, is associated with a lower risk of developing cervical intraepithelial neoplasia. *Nutrition* 2011; 27(5): 513–9.

14. Senthong A, Kitkumthorn N, Rattanatanyong P, Khemapech N, Triratanachart S, Mutirangura A. Differences in LINE-1 methylation between endometriotic ovarian cyst and endometriosis-associated ovarian cancer. *Int J Gynecol Cancer* 2014; 24(1): 36–42.

15. Cash HL, Tao L, Yuan J-M, Marsit CJ, Houseman EA, Xiang Y-B, et al. LINE-1 hypomethylation is associated with bladder cancer risk among nonsmoking Chinese. *Int J Cancer* 2012; 130(5): 1151–9.

16. Kawano H, Saeki H, Kitao H, Tsuda Y, Otsu H, Ando K, et al. Chromosomal Instability Associated with Global DNA Hypomethylation is Associated with the Initiation and

Progression of Esophageal Squamous Cell Carcinoma. *Ann Surg Oncol* 2014; 21(4): 696–702.

17. Di J, Han X, Gu W, Wang Y, Zheng Q, Zhang P, et al. Association of hypomethylation of LINE-1 repetitive element in blood leukocyte DNA with an increased risk of hepatocellular carcinoma. *J Zhejiang Univ Sci B* 2011; 12(10): 805–11.
18. Gao Y, Baccarelli A, Shu XO, Ji B-T, Yu K, Tarantini L, et al. Blood leukocyte Alu and LINE-1 methylation and gastric cancer risk in the Shanghai Women's Health Study. *Br J Cancer* 2012; 106(3): 585–91.
19. Muka T, Koromani F, Portilla E, O'Connor A, Bramer WM, Troup J, et al. The role of epigenetic modifications in cardiovascular disease: A systematic review. *Int J Cardiol* 2016; 212: 174-83.
20. 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* 201; 6(7): 828–37.
21. Jones MJ, Goodman SJ, Kobor MS. DNA methylation and healthy human aging. *Aging Cell* 2015; 14(6): 924–32.
22. Dixon JB. The effect of obesity on health outcomes. *Mol Cell Endocrinol* 2010; 316(2): 104–8.
23. Vucenik I, Stains JP. Obesity and cancer risk: evidence, mechanisms, and recommendations. *Ann N Y Acad Sci* 2012; 1271(1): 37–43.
24. Renehan AG, Zwahlen M, Egger M. Adiposity and cancer risk: new mechanistic insights from epidemiology. *Nat Rev Cancer* 2015; 15(8): 484–98.

25. Friedenreich CM, Woolcott CG, McTiernan A, Ballard-Barbash R, Brant RF, Stanczyk FZ, et al. Alberta physical activity and breast cancer prevention trial: Sex hormone changes in a year-long exercise intervention among postmenopausal women. *J Clin Oncol* 2010; 28(9): 1458–66.

26. Friedenreich CM, Woolcott CG, McTiernan A, Terry T, Brant R, Ballard-Barbash R, et al. Adiposity changes after a 1-year aerobic exercise intervention among postmenopausal women: a randomized controlled trial. *Int J Obes* 2011; 35(3): 427–35.

27. Friedenreich CM, Courneya KS, Neilson HK, Matthews CE, Willis G, Irwin M, et al. Reliability and validity of the past year total physical activity questionnaire. *Am J Epidemiol* 2006; 163(10): 959–70.

28. Csizmadi I, Kahle L, Ullman R, Dawe U, Zimmerman TP, Friedenreich CM, et al. Adaptation and evaluation of the National Cancer Institute's Diet History Questionnaire and nutrient database for Canadian populations. *Public Health Nutr* 2007; 10(1): 88–96.

29. Willet W, Howe G, Kushi L. Adjustment for total energy intake in epidemiologic studies. *Am J Clin Nutr* 1997; 65(4): 1220S – 8S.

30. Chavez-MacGregor M, Elias SG, Onland-Moret NC, van der Schouw YT, Van Gils CH, Monninkhof E, et al. Postmenopausal breast cancer risk and cumulative number of menstrual cycles. *Cancer Epidemiol Biomarkers Prev* 2005; 14(4): 799–804.

31. Kile ML, Baccarelli A, Tarantini L, Hoffman E, Wright RO, Christiani DC. Correlation of global and gene-specific DNA methylation in maternal-infant pairs. *PLoS One* 2010; 5(10): e13730.

32. Piyathilake CJ, Badiga S, Alvarez RD, Partridge EE, Johanning GL. A lower degree of PBMC L1 methylation is associated with excess body weight and higher HOMA-IR in the presence of lower concentrations of plasma folate. *PLoS One* 2013; 8(1): e54544.

33. Kihara S, Matsuzawa Y. Fat Distribution and Cardiovascular Disease Risk. *Curr Cardiovasc Risk Rep* 2015; 9(3): 4–9.

34. Anderson OS, Sant KE, Dolinoy DC. Nutrition and epigenetics: an interplay of dietary methyl donors, one-carbon metabolism and DNA methylation. *J Nutr Biochem* 2012; 23(8): 853–9.

35. van Dijk SJ, Molloy PL, Varinli H, Morrison JL, Muhlhausler BS. Epigenetics and human obesity. *Int J Obes* 2014; 39:1–13.

36. Delgado-Cruzata L, Zhang W, McDonald JA, Tsai WY, Valdovinos C, Falci L, et al. Dietary Modifications, Weight Loss, and Changes in Metabolic Markers Affect Global DNA Methylation in Hispanic, African American, and Afro-Caribbean Breast Cancer Survivors. *J Nutr* 2015; 145(4): 783–90.

37. Martín-Núñez GM, Cabrera-Mulero R, Rubio-Martín E, Rojo-Martínez G, Olveira G, Valdés S, et al. Methylation levels of the SCD1 gene promoter and LINE-1 repeat region are associated with weight change: An intervention study. *Mol Nutr Food Res* 2014; 58(7): 1528–36.

38. Cash HL, McGarvey ST, Houseman EA, Marsit CJ, Hawley NL, Lambert-Messerlian GM, et al. Cardiovascular disease risk factors and DNA methylation at the LINE-1 repeat region in peripheral blood from Samoan Islanders. *Epigenetics* 2011; 6(10): 1257–64.

39. Perng W, Villamor E, Shroff MR, Nettleton JA, Pilsner JR, Liu Y, et al. Dietary intake, plasma homocysteine, and repetitive element DNA methylation in the Multi-Ethnic Study of Atherosclerosis (MESA). *Nutr Metab Cardiovasc Dis* 2014; 24(6): 614–22.

40. 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* 2014: 6(5): 623–9.

41. Duggan C, Xiao L, Terry MB, McTiernan A. No effect of weight loss on LINE-1 methylation levels in peripheral blood leukocytes from postmenopausal overweight women. *Obesity* 2014; 22(9): 2091–6.

42. 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* 2012; 41(1): 126–39.

43. Marques-Rocha JL, Milagro FI, Mansego ML, Mourão DM, Martínez JA, Bressan J. LINE-1 methylation is positively associated with healthier lifestyle but inversely related to body fat mass in healthy young individuals. *Epigenetics* 2016; 11(1): 49-60.

44. Jintaridth P, Tungtrongchitr R, Preutthipan S, Mutirangura A. Hypomethylation of Alu Elements in Postmenopausal Women with Osteoporosis. *PLoS One* 2013; 8(8): 1–8.

45. Na YK, Hong HS, Lee DH, Lee WK, Kim DS. Effect of body mass index on global DNA methylation in healthy Korean women. *Mol Cells* 2014; 37(6): 467–72.

46. Andreotti G, Karami S, Pfeiffer RM, Hurwitz L, Liao LM, Weinstein SJ, et al. LINE1 methylation levels associated with increased bladder cancer risk in pre-diagnostic blood DNA among US (PLCO) and European (ATBC) cohort study participants. *Epigenetics* 2013; 9(3): 404–15.

47. Liao LM, Brennan P, van Bemmel DM, Zaridze D, Matveev V, Janout V, et al. LINE-1 methylation levels in leukocyte DNA and risk of renal cell cancer. *PLoS One* 2011; 6(11): e27361.

48. Tajuddin SM, Amaral AFS, Fernández AF, Chanock S, Silverman DT, Tardón A, et al. LINE-1 methylation in leukocyte DNA, interaction with phosphatidylethanolamine N-methyltransferase variants and bladder cancer risk. *Br J Cancer* 2014; 110(8): 2123–30.

49. Dixon JB, O'Brien PE. Obesity and the white blood cell count: Changes with sustained weight loss. *Obes Surg* 2006; 16(3): 251–7.

50. Lowe R, Slodkowicz G, Goldman N, Rakyan VK. The human blood DNA methylome displays a highly distinctive profile compared with other somatic tissues. *Epigenetics* 2015; 10(4): 274–81.

51. 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; 23(54): 8841–6.

52. Ashbury JE, Taylor SA, Tse MY, Pang SC, Louw JA, Vanner SJ, et al. Biomarkers measured in buccal and blood leukocyte DNA as proxies for colon tissue global methylation. *Int J Mol Epidemiol Genet* 2014; 5(2): 120–4.

53. Rothman KJ. BMI-related errors in the measurement of obesity. *Int J Obes* 2008; 32: S56–9.

54. Sun G, French CR, Martin GR, Younghusband B, Green RC, Xie YG, et al. Comparison of multifrequency bioelectrical impedance analysis with dual-energy X-ray absorptiometry for assessment of percentage body fat in a large, healthy population. *Am J Clin Nutr* 2005; 81(1): 74–8.

**Table 4.1. Baseline Characteristics of a Subset of ALPHA Trial Participants (N = 289)**

| Covariate | Mean, Median, or Frequency |
|---|---|
| Age (years) | |
| Mean (SD) | 60.98 (5.51) |
| 50 to 60 years, n (%) | 152 (52.60) |
| 61 to 75 years, n (%) | 137 (47.40) |
| Ethnicity | |
| Caucasian, n (%) | 263 (91.00) |
| Other, n (%) | 26 (9.00) |
| Past Year Physical Activity (MET-hours per week), Median (IQR) | 107.56 (77.01 – 149.19) |
| Alcohol Consumption (grams per day), Median (IQR) [a] | 1.77 (0.55 – 5.24) |
| Energy-Adjusted Folate Intake (micrograms DFE / day), Median (IQR) [a, b] | 510.14 (453.92 – 567.37) |
| Smoking Status | |
| Former Smoker, n (%) | 96 (33.22) |
| Non-smoker, n (%) | 193 (66.78) |
| Previous Hormone Replacement Therapy Use | |
| Yes, n (%) | 158 (54.67) |
| No, n (%) | 131 (45.33) |
| Age at first birth | |
| Nulliparous, n (%) | 25 (8.65) |
| 16-24 years, n (%) | 154 (53.29) |
| 25-43 years, n (%) | 110 (38.06) |
| Parity, mean (SD) | 2.42 (1.29) |
| Age at menarche, mean (SD) | 12.77 (1.49) |
| Age at menopause, mean (SD) | 48.41 (6.30) |
| Lifetime Number of Menstrual Cycles (#), Median (IQR) | 455.94 (394.51 – 485.03) |
| Quintile 1: <376.08, n (%) | 56 (19.38) |
| Quintile 2: 376.08 to <432.85, n (%) | 58 (20.07) |
| Quintile 3: 432.85 to <471.19, n (%) | 58 (20.07) |
| Quintile 4: 471.19 to <497.28, n (%) | 57 (19.72) |
| Quintile 5: 497.28+, n (%) | 60 (20.76) |
| Time since menopause (years), Median (IQR) | 10.51 (6.06 – 17.64) |
| Presence of Comorbidities [c] | |
| Yes, n (%) | 110 (38.06) |
| No, n (%) | 179 (61.94) |

[a] Missing information for five participants

[b] One microgram dietary folate equivalent (DFE) = one microgram of food folate or 0.6 micrograms of folic acid from fortified foods

[c] Ever been diagnosed with at least one of the following: high cholesterol or triglycerides, myocardial infarction, angina pectoris, stroke, rheumatoid arthritis or osteoarthritis, osteoporosis, blood clots in the veins of legs or pelvis, blood clots in the lung, or thyroid problems.

**Table 4.2. Baseline Adiposity and DNA Methylation Measures from a Subset of ALPHA Trial Participants (N = 289)**

| Current Adiposity | Mean (SD) |
|---|---|
| *CT Scan* | |
| Abdominal fat area ($cm^2$) | 437.54 (160.65) |
| Intraabdominal fat area ($cm^2$) | 104.35 (56.52) |
| Subcutaneous fat area ($cm^2$) | 333.19 (121.84) |
| | |
| *DXA Scan* | |
| Body fat percent (%) | 42.46 (5.31) |
| Fat mass (kg) | 31.58 (8.70) |
| | |
| *Baseline Physical Assessment* | |
| Baseline Weight (kg) | 76.34 (12.86) |
| Waist circumference (cm) | 89.27 (10.51) |
| Hip circumference (cm) | 110.23 (9.90) |
| Waist to hip ratio (waist / hip circumference) | 0.81 (0.06) |
| BMI ($kg/m^2$) | 29.25 (4.39) |

| Weight Changes Over Time (kg) | Mean (SD) |
|---|---|
| Weight at age 20 | 55.38 (6.93) |
| Weight Gain | 20.85 (11.36) |
| Weight Loss | 1.09 (3.18) |
| Weight Change | 13.99 (8.82) |

| DNA Methylation (%) | Mean (SE)[a] |
|---|---|
| LINE-1 methylation | 72.73 (0.90) |
| Alu methylation | 19.09 (0.13) |

[a] The least squares mean and standard error of the least squares mean after adjusting for intra-plate variability

**Table 4.3. The Association between Baseline Measures of Adiposity and Repetitive Element Methylation amongst Healthy Postmenopausal Women (N = 289)**

| | LINE-1 Methylation (%) | | | Alu Methylation (%) | | |
|---|---|---|---|---|---|---|
| | β[a] | 95% C.I. | p-value | β[b] | 95% C.I. | p-value |
| **Current Adiposity** | | | | | | |
| *CT Scan* | | | | | | |
| Abdominal fat area ($cm^2$) | 0.20 | 0.02 - 0.39 | 0.03 | 0.03 | -0.04 - 0.09 | 0.40 |
| Intraabdominal fat area ($cm^2$) | 0.21 | 0.02 - 0.40 | 0.03 | 0.04 | -0.03 - 0.10 | 0.27 |
| Subcutaneous fat area ($cm^2$) | 0.17 | -0.02 - 0.35 | 0.08 | 0.02 | -0.04 - 0.08 | 0.55 |
| *DXA Scan* | | | | | | |
| Body fat percent (%) | 0.20 | 0.01 - 0.39 | 0.048 | 0.03 | -0.04 - 0.09 | 0.44 |
| Fat mass (kg) | 0.26 | 0.07 - 0.45 | 0.01 | 0.03 | -0.04 - 0.09 | 0.40 |
| *Baseline Physical Assessment* | | | | | | |
| Baseline weight (kg) | 0.30 | 0.11 - 0.49 | 0.002 | 0.03 | -0.04 - 0.09 | 0.42 |
| Waist circumference (cm) | 0.21 | 0.02 - 0.40 | 0.03 | 0.04 | -0.03 - 0.10 | 0.24 |
| Hip circumference (cm) | 0.33 | 0.14 - 0.52 | 0.001 | 0.02 | -0.05 - 0.08 | 0.60 |
| Waist to hip ratio (waist / hip circumference) | -0.06 | -0.25 - 0.13 | 0.53 | 0.03 | -0.03 - 0.09 | 0.32 |
| BMI ($kg/m^2$) | 0.21 | 0.03 - 0.40 | 0.03 | 0.03 | -0.04 - 0.09 | 0.40 |
| **Weight Changes Over Time** | | | | | | |
| Weight at age 20 (kg)[c] | 0.22 | 0.03 - 0.41 | 0.02 | -0.01 | -0.07 - 0.04 | 0.65 |
| Weight Gain[c] | 0.20 | 0.01 - 0.39 | 0.04 | 0.02 | -0.04 - 0.07 | 0.57 |
| Weight Loss (kg) | 0.01 | -0.18 - 0.19 | 0.93 | -0.01 | -0.08 - 0.05 | 0.83 |
| Weight Change (kg)[d] | 0.16 | -0.04 - 0.35 | 0.11 | 0.02 | -0.03 - 0.08 | 0.39 |

[a] Estimated increase in percent methylation associated with a standard deviation increase in the exposure after adjusting for ethnicity, cumulative number of lifetime menstrual cycles, and within plate clustering

[b] Estimated increase in percent methylation associated with a standard deviation increase in the exposure after adjusting for the presence of comorbidities and within plate clustering

[c] Two participants were excluded from the analysis due to missing data

[d] Thirteen participants were excluded from the analysis due to missing data

# Chapter 5

# Endogenous Estrogen Exposure is Associated with Repetitive Element DNA Methylation in Healthy Postmenopausal Women

## INTRODUCTION

It is becoming increasingly apparent that repetitive DNA elements play an important role in the etiology of a wide variety of diseases. (1) Two common types of repetitive elements that are mobile in the human genome are long interspersed nuclear element-1 (LINE-1) and Alu elements. (2) Epigenetic mechanisms, such as high levels of DNA methylation, help to suppress the expression of these potentially harmful transposable elements. (3) Although it has yet to be established, emerging evidence suggests that healthy individuals with lower levels of DNA methylation within LINE-1 and Alu repeats have an elevated risk of developing certain types of cancer, (4,5) cardiovascular disease, (6) and possibly neurodegenerative disorders. (7) Despite their suspected role in the initiation of disease, the determinants of LINE-1 and Alu hypomethylation have yet to be fully uncovered.

Sex hormones have a complex and heterogeneous relation to disease. For example, estrogen is an established risk factor for breast, ovarian, and endometrial cancers (8) but has been shown to be protective against colorectal cancer, (9) heart disease, (10) and neurodegeneration. (11) Even within a single disease type there exists heterogeneity with respect to the etiologic role of sex hormones as demonstrated by the varied impact that hormonal risk factors have on different breast cancer subtypes. (12) Understanding the underlying biological mechanisms may help to explain such disparity in the literature.

Emerging evidence suggests that epigenetic dysregulation is a mechanism through which certain exposures impact disease risk. (13) In particular, changes in levels of LINE-1 and Alu methylation might be mechanisms whereby sex hormones increase the risk of certain cancers or, conversely, reduce the risk of colorectal cancer, heart disease, or neurodegenerative disorders.

The aim of the current study is to describe the relation between endogenous sex hormones and repetitive element DNA methylation in healthy, postmenopausal women. To address this objective, we assessed the effects of serum endogenous sex hormone concentrations and cumulative ovarian sex hormone exposure, as measured by lifetime number of menstrual cycles (LNMC), on levels of LINE-1 and Alu methylation in a group of healthy, inactive, postmenopausal women. Sex hormones have been shown to repress the activity of DNA methyl-transferases and methyl-CpG-binding proteins – enzymes necessary for the maintenance of DNA methylation patterns. (14) As such, we hypothesized that greater exposure to sex hormones would be associated with lower levels of DNA methylation within repetitive elements. As done in past research, (15,16) we also explored if the relation between sex hormones and repetitive element DNA methylation was modified by alcohol and folate consumption – two factors thought to affect patterns of DNA methylation through their role in one carbon metabolism. (17,18)

## MATERIALS AND METHODS

### *Participants and Study Design*

We conducted a cross-sectional study using baseline information from an ancillary study of the Alberta Physical Activity and Breast Cancer Prevention (ALPHA) Trial (ClinicalTrials.gov identifier: NCT00522262). (19) The aim of this year-long randomized

exercise intervention trial was to examine the effects of aerobic activity on various biomarkers of etiologic relevance to breast cancer. Participants consisted of 320 healthy, inactive, postmenopausal women between 50 to 74 years who resided in either Calgary or Edmonton, Alberta. To be eligible for inclusion, participants had to be non-smokers, not consume excessive amounts of alcohol (i.e. no more than 2 drinks per day), have no major co-morbidities, have not been previously diagnosed with cancer, and not taking medications that could affect estrogen levels (e.g. hormone replacement therapy).

### *Data Collection*

Demographic and covariate information such as age, ethnicity, previous smoking, and medical history (e.g. co-morbidities, previous hormone replacement therapy use, and hysterectomy status) were captured in a baseline health questionnaire. Variables used to estimate LNMC (i.e. age at menarche, age at menopause, parity, and gravidity) were also self-reported in the baseline health questionnaire. We used the validated Past Year Total Physical Activity Questionnaire to assess physical activity. (20) Dietary folate intake and alcohol consumption in the past year was estimated using a validated food frequency questionnaire. (21) The nutrient residual method was used to standardize dietary folate intake to a total caloric intake of 2000 kcal. (22) Body fat percent was measured via dual energy x-ray absorptiometry scan as previously described. (23)

### *Blood Collection and Hormone Assays*

The blood collection and processing protocol and hormone assays used in the current study have been previously described. (19) Briefly, fasting blood samples were collected from all

participants at baseline and were processed to obtain serum before being frozen at -86ºC. Samples were sent to the Reproductive Endocrine Research Laboratory at the University of Southern California in Los Angeles for analysis. A radioimmunoassay with preceding organic solvent extraction and Celite column partition steps (CAS No. 68855-54-9, Mallinckrodt Baker) were used to measure total estrone, estradiol, androstenedione, and testosterone concentrations. The Immulite Analyzer (Siemens Health Corporation) was used to assess sex hormone-binding globulin (SHBG). Intra- and inter-batch coefficients of variation were between 4.0% to 7.5% and 8.0% to 13.0% respectively. All of the participants had sex hormone concentrations within the range of detection.

***DNA Methylation Analysis***

We analyzed repetitive element DNA methylation in peripheral white blood cells that were extracted from frozen baseline blood samples. DNA was isolated from sample buffy coats on the Hamilton STARlet (Hamilton Robotics Inc.) liquid handler using the Macherey-Nagel NucleoMag Blood 200 µl kit (Macherey-Nagel GmbH & Co. KG). DNA yields were measured on a Qubit flourometer using the Qubit dsDNA HS Assay Kit (ThermoFisher Scientific Inc.). It was determined that 298 out of the 320 samples extracted had sufficient DNA yields for downstream methylation analyses. The assessment of repetitive element DNA methylation was performed at McGill University and Génome Québec Innovation Centre in Montréal, Quebec. The DNA samples were treated with sodium bisulfite using the EZ-96 DNA Methylation-Gold Kit (Zymo Research, Catalog No. D5007). The HotStarTaq DNA Polymerase kit (Qiagen) was used to amplify target regions within the LINE-1 and Alu repeats. The oligonucleotide primers used in the current study have been previously described. (24) The total proportion of methylated

cytosine bases at three distinct CpG sites were quantified by pyrosequencing PCR products using the PyroMark Q24 (Qiagen) for both LINE-1 and Alu elements. The final LINE-1 and Alu methylation outcomes used in the current investigation were defined as average percent methylation across the three CpG sites that were analyzed. We assessed the internal consistency of this assay by repeating the methylation analyses on eight samples. For LINE-1 and Alu methylation, the inter-plate coefficients of variation were 1.83% and 2.19% respectively. The assay was unable to detect methylation values for two samples and three samples were excluded because they failed to meet internal quality control standards which resulted in a final sample size of 293.

***Statistical Analysis***

Linear regression was used to model the relation between sex hormone exposure and repetitive element DNA methylation while controlling for confounding. The endogenous sex hormone variables and LNMC were modeled as categorical variables using quartile cut-points. The significance of a linear trend in the data was assessed by modeling the rank of the categories as a continuous ordinal variable. We also modeled the exposures as continuous variables if the relationship was linear. To allow for comparisons across different exposures, effect sizes for the continuous variables were standardized such that the regression coefficient represented the mean percent methylation increase associated with a one standard deviation increase in the exposure of interest.

Depending on the analysis, we adjusted for categorical representations of LNMC, estradiol, testosterone, and SHBG. Estradiol, testosterone, and SHBG were controlled for in the analyses of LNMC and its component variables. In addition to LNMC, estrogens (i.e. estradiol or

estrone) were adjusted for testosterone and SHBG, testosterones (i.e. testosterone or androstenedione) were adjusted for estradiol and SHBG, and SHBG was adjusted for testosterone and estradiol. In all analyses, we included a random effects parameter with an unstructured covariance matrix to account for inter-plate variability in the DNA methylation assay. We decided which of the remaining covariates to adjust for in the final model using a backwards elimination approach with a p-value of 0.20. Age, past year total physical activity, past year alcohol consumption, past year dietary folate intake, and body fat percent were modeled as continuous variables. Categorical covariates included ethnicity, smoking history, presence of comorbidities, hormone replacement therapy use, and age at first birth

Amongst postmenopausal women, adipose tissue is the primary source of endogenous estrogen exposure such that its relation with cancer is believed to be almost entirely mediated by sex hormones. (25) Body fat percent was therefore excluded from analyses of the endogenous sex hormones because it largely exists on the causal pathway in this population. We conducted sensitivity analyses in which we subsequently adjusted for body fat percent in these analyses.

Four participants were excluded due to missing data resulting in a final sample size of 289. One observation was excluded from the Alu methylation analyses as it was determined to be an influential observation with an outlying Alu methylation value of 22.7% which resulted in a sample size of 288. Five individuals were not included in the LINE-1 analyses because they were found to have been missing important covariate information which left 284 participants in these analyses.

Following this main effect analysis, we considered potential interactions with dietary folate and alcohol consumption. Median cut-points were used to define low and high groups of alcohol and folate consumption. Due to missing alcohol and folate consumption data, four

additional participants were removed from the Alu methylation analysis. We also examined whether or not the relation between LNMC and the methylation outcomes differed according to whether or not the participant had ever used hormone replacement therapy or had a surgically induced menopause. All analyses were conducted at the 5% significance level in SAS v.9.4.

## RESULTS

The study participants were 61 years of age on average and were predominantly Caucasian with a mean body fat percent of 42.5% and a median of 108 MET-hours/week of total physical activity in the past year (Table 1). Participants consumed a median of 1.77 grams of alcohol and 510 micrograms of dietary folate equivalents per day in the year before their enrollment into the ALPHA Trial. Of the 289 participants included in this analysis, 96 (33%) were former smokers, 158 (55%) had previously used hormone replacement therapy, 216 (75%) had a natural menopause, and 110 (38%) had at least one comorbidity.

Percent LINE-1 and Alu methylation followed a normal distribution with a corresponding mean of 72.7% (SE: 0.9) and 19.1% (SE: 0.1; Table 2). All of the serum concentrations of endogenous sex hormones had right skewed distributions with a median value of 10.0 pg/mL (IQR: 7.0 – 14.0), 32.0 pg/mL (IQR: 24.0 – 45.0), 24.2 ng/dL (IQR: 17.4 – 32.4), 571 pg/mL (IQR: 421 – 766), and 40.0 nmol/L (IQR: 29.0 – 54.0) for estradiol, estrone, testosterone, androstenedione, and SHBG respectively. In general, participants underwent menarche at 13 years and reached menopause at 48 years. Participants had an average of two children and 25 (9%) were nulliparous. Lifetime number of menstrual cycles was left skewed with a median of 456 menstrual cycles (IQR: 395 – 485).

In addition to a set of hormonal variables selected *a priori*, we additionally adjusted for ethnicity and past year alcohol consumption in the LINE-1 methylation analyses and smoking history in the Alu methylation analyses as determined by the backwards elimination procedure.

We first examined the association between serum sex hormone concentrations (Table 3) and cumulative ovarian sex hormone exposure (Table 4) with LINE-1 and Alu methylation. With respect to LINE-1 methylation, a dose-response relationship was observed across quartiles of estradiol (p-trend = 0.01) and estrone (p-trend = 0.049). In particular, a one standard deviation increase in estradiol (12.0 pg/mL) and a one standard deviation increase in estrone (17.7 pg/mL) were significantly associated with a mean increase in LINE-1 methylation of 0.21% (95% CI: 0.03 to 0.39; p = 0.03) and 0.25% (95% CI: 0.04 to 0.46; p = 0.02) respectively. Testosterone, androstenedione, and SHBG were not related to LINE-1 methylation (p > 0.05). None of the exposures were associated with Alu methylation in this analysis (p > 0.05).

Based on an examination of the residuals and the significance of a polynomial term, it was determined that there was a non-linear "U-shaped" relation between LNMC and LINE-1 methylation. As such, we did not consider a continuous representation of this variable with respect to LINE-1 methylation. The categorical representation of LNMC was related to LINE-1 methylation (p = 0.02). Age at menopause similarly had a curve-linear relation with LINE-1 methylation after adjusting for confounding (p = 0.02). The data were suggestive of a dose-response relationship between parity and LINE-1 methylation (p-trend = 0.07). Specifically, a one standard deviation increase in the number of live births (1.3 births) was associated with a mean increase in LINE-1 methylation of 0.21% (95% CI: 0.03 to 0.38; p = 0.02).

We subsequently investigated dietary folate and alcohol consumption as effect modifiers of the relations of interest (Table 5). Dietary folate was found to modify the relationship between

SHBG and LINE-1 methylation (p-interaction = 0.01) and between LNMC and Alu methylation (p-interaction = 0.02). In particular, SHBG had a positive relation with LINE-1 methylation in the low folate consumption group and no relation with LINE-1 methylation amongst high folate consumers. Similarly, LNMC had a negative relation with Alu methylation but only in the presence of low folate consumption. Although the interaction term did not reach statistical significance, these results were consistent with those found in the analyses that examined effect modification by alcohol consumption in which the effect of the exposure became more pronounced in the presence of high alcohol intake.

In a sensitivity analysis, we found that subsequent adjustments for body fat percent did not meaningfully alter the significance or the magnitude of any of the estimated effects (data not shown). In addition, we found that the observed relations did not differ according to whether or not the participant had a surgically induced menopause or had ever taken hormone replacement therapy (data not shown).

## DISCUSSION

In the current study, exposure to endogenous estrogens was found to be related to repetitive element DNA methylation in the white blood cells of healthy postmenopausal women. Serum concentrations of estradiol and estrone were positively related to LINE-1 methylation. SHBG was also found to be positively related to LINE-1 methylation but only in the presence of low folate intake. None of the serum sex hormone concentrations were related to Alu methylation and although the data were suggestive of an inverse dose-response relation between testosterone and LINE-1 methylation (p-trend = 0.07), no significant associations were found

with any of the androgen measures. These findings suggest that estrogens and SHBG play a role in maintaining high levels of LINE-1 methylation in healthy, postmenopausal women.

To date, few studies have examined the relation between serum sex hormone concentrations and repetitive element DNA methylation and there is no clear consensus in the literature. Of particular relevance are two prior investigations done on healthy, postmenopausal women. Using measures similar to our own, Ulrich et al. (2012) found that androstenedione, estrone, estradiol, SHBG, and testosterone were inversely related to LINE-1 methylation but only in the presence of low folate consumption. (15) The direction of the relation with the estrogen measures, the presence of an association with the androgen measures, and the interaction with dietary folate reported by Ulrich and colleagues are inconsistent with our findings. In contrast, Iwasaki et al. (2012) found no associations between endogenous sex hormone concentrations and global DNA methylation nor did they find effect modification by folate or alcohol consumption. (16) It is difficult to draw any conclusions pertaining to the relation between sex hormone concentrations and repetitive element DNA methylation given the current state of the literature. Differences with respect to the adjustment of covariates, methods of assessment, and the racial distribution of the underlying study populations may account for the disparity in these results.

This is the first study to explore the relation between cumulative exposure to ovarian sex hormones, as represented by LNMC, and repetitive element DNA methylation. We found that LNMC had a “U-shaped” curvilinear relation with LINE-1 methylation and that this relation did not depend on dietary folate intake. On average, women in the first and fourth quartiles of LNMC had the highest levels of LINE-1 methylation whereas participants in the second and third quartiles had the lowest levels of LINE-1 methylation. In contrast, a negative linear relation

between LNMC and Alu methylation was observed but only in the presence of low folate consumption. Although it did not reach statistical significance, the data suggested that higher levels of alcohol consumption had a similar moderating effect. Since none of the participants included in this study were considered to be heavy drinkers, it is possible that the augmentation of the effect of LNMC may be more pronounced amongst women who consume excessive amounts of alcohol.

Our findings suggest that total lifetime number of menstrual cycles may not fully capture the independent effects of age at menarche and parity on repetitive element DNA methylation. A higher LNMC is associated with a higher age at menopause, lower age at menarche, and lower number of births. As such, the relation between LNMC and repetitive element DNA methylation is expected to be similar for age at menopause and in the opposite direction for age at menarche and parity. While the relation between age at menopause and repetitive element DNA methylation was consistent with that for LNMC, age at menarche and parity were not. For example, we found a significant positive linear association between parity and LINE-1 methylation that was suggestive of a dose-response relationship which is not consistent with the curvilinear relation observed with LNMC. These results indicate that age at menarche and parity might influence repetitive element DNA methylation through additional mechanisms other than cumulative estrogen and progesterone exposure.

The findings with the serum estrogen concentrations are inconsistent with those from the analyses on cumulative ovarian sex hormone exposure. Such disparities may be explained by the type, timing, and duration of sex hormone exposure. Unlike the serum estrogen measures, the LNMC variable additionally accounts for progesterone. With respect to the timing and duration of the exposures, LNMC reflects long-term premenopausal ovarian sex hormone exposure

whereas the serum hormone concentrations markers represent short-term postmenopausal estrogen levels. The relation between ovarian sex hormones and repetitive element DNA methylation may therefore be dependent upon the type of hormone (i.e. progesterone or estrogen), the duration of exposure, and menopausal status.

In the current investigation, the effect of endogenous estrogen exposure on repetitive element DNA methylation differed depending on the genomic region under scrutiny which suggests that LINE-1 and Alu elements are not equally susceptible to epigenetic dysregulation by endogenous estrogen exposure. Although both measures are often considered surrogate markers of global DNA methylation, (26) these results suggest that LINE-1 and Alu methylation may be better conceptualized as independent epigenetic events.

There are limitations to the current investigation. Some variables that would have allowed for a more precise estimate of the participant's LNMC were unavailable. Another potential limitation is our use of a single time point measure of serum sex hormone concentrations. However, sex hormone concentrations have been shown to be relatively stable in postmenopausal women and a single time point measure has been shown to adequately predict cancer risk for up to 20 years. (27,28) Thus, the benefits of including two or more measures of sex hormone concentrations are expected to be minimal and have little impact on the final results of this study. In the current study, we made no adjustments for multiple comparisons. The statistical significance of these results need to be interpreted with caution. DNA methylation patterns have been shown to be specific to different cell- and tissue-types. (29–33) We assessed DNA methylation in peripheral white blood cells and did not take into considerations the distribution of cell types. As such, our results may be confounded by cell type. Similarly, the relations we reported may not persist in other target tissues of interest such as the breast, brain,

or heart. Finally, the highly selective study population reduces the external validity of these results which may not be generalizable to younger premenopausal women, men, or diseased populations of non-Caucasian ethnicities.

This study also has important strengths. This study is the first to consider cumulative ovarian sex hormone exposure as represented by LNMC in addition to several plasma sex hormone concentrations. Moreover, this study is the largest to date focused on the relation between endogenous sex hormone exposure and DNA methylation. Besides this investigation, the two largest studies examining the relation between sex hormone concentrations and DNA methylation in healthy, postmenopausal women were conducted by Iwasaki et al. (16) and Ulrich et al. (15) with sample sizes of 185 and 173 participants respectively. The adjustment for several suspected confounders is another major strengths of this investigation. The exclusion criteria of the underlying study reduced or entirely eliminated variability in several potential confounders (e.g. age, sex, smoking, alcohol consumption, physical activity, etc.) and we had access to information on several covariates which we were able to assess in regression analyses.

These results suggest that repetitive element DNA methylation may be a mechanism through which endogenous estrogen exposure has an impact on the risk of certain diseases. In particular, LINE-1 methylation may be a pathway whereby higher levels of serum estrogen concentrations protect against colorectal cancer, heart disease, and neurodegeneration in postmenopausal women. Additional etiologic research on repetitive element DNA methylation and disease risk is needed in order to determine whether the effect sizes we observed are of clinical relevance and to determine the joint effects of LINE-1 and Alu hypomethylation on disease risk. For example, the positive trend with LINE-1 methylation at higher levels of LNMC might be offset by its negative relation with Alu methylation, thereby resulting in an overall

increased risk of disease attributed to DNA methylation changes occurring at higher levels of cumulative ovarian sex hormone exposure.

In summary, there is evidence that supports the relation between endogenous estrogen exposure and repetitive element DNA methylation. Few studies have examined this association to date and the results of such studies are heterogeneous. With respect to the foregoing relation, future research should consider folate intake, alcohol consumption, and other factors related to one-carbon metabolism as potential effect modifiers. Researchers should also consider targeting populations with more variable folate and alcohol intake to better assess the moderating effect of these dietary factors. In addition to controlling for suspected determinants of DNA methylation, future investigations should adjust for hormonal factors other than the one under scrutiny. The extent to which LNMC fully captures the independent effects of its component variables is questionable. When considered as an exposure of interest, future studies should additionally consider the independent effects of the reproductive and menstrual cycle variables used to estimate LNMC.

### *Conclusion*

Endogenous estrogen exposure is associated with repetitive element DNA methylation in healthy postmenopausal women. This relation may depend upon the timing, duration, and type of sex hormone, the participants' folate intake, and the specific repetitive element under scrutiny. Epigenetic mechanisms such as repetitive element DNA methylation may help to explain the heterogeneous relation that estrogens have with the risk of different diseases. Further research into the relation between repetitive element DNA methylation and disease risk is needed in order to appreciate the clinical significance of these results.

### *Competing interests*

No potential conflicts of interest are declared by the authors.


### *Acknowledgements*

Devon Boyne is supported by funding from the Queen's University Terry Fox Foundation Training Program in Transdisciplinary Cancer Research in partnership with the Canadian Institutes for Health Research (CIHR) and an Ontario Graduate Scholarship. Christine Friedenreich was supported by an Alberta Innovates Health Solutions Health Senior Scholar Award and by the Alberta Cancer Foundation Weekend to End Women's Cancers Breast Cancer Chair. Kerry S. Courneya is supported by the Canada Research Chairs Program.

The ALPHA Trial was originally funded by a research grant from the Canadian Breast Cancer Research Alliance and the ancillary study is funded by grants from CIHR and the Canadian Cancer Society. Funding for the DNA extraction was provided by an Alberta Cancer Foundation grant for the Biospecimen Processing Unit within the Translational Laboratories (TL) at the Tom Baker Cancer Center (TBCC). The ALPHA Trial study set-up was performed by Kim van der Hoek and Marla Orenstein. The Study Coordinators were Rosemary Crosby and Ame-Lia Tamburrini. Data preparation for this ancillary study was done by Qinggang Wang. DNA extraction and quantification was performed by Angela Chan and Michelle Dean in the Translational Labratory. Dr. Nigel Brockton provided advice on the DNA extraction protocols developed for this study.

## References


1. Hancks DC, Kazazian HH. Roles for retrotransposon insertions in human disease. *Mob DNA* 2016; 7: 9.
2. Padeken J, Zeller P, Gasser SM. Repeat DNA in genome organization and stability. *Curr Opin Genet Dev* 2015; 31:12–9.
3. Bernstein BE, Meissner A, Lander ES. The Mammalian Epigenome. *Cell* 2007; 128: 669–81.
4. Baba Y, Ishimoto T, Kurashige J, Iwatsuki M, Sakamoto Y, Yoshida N, Watanabe M, Baba H. Epigenetic field cancerization in gastrointestinal cancers. *Cancer Lett* 2016; 375: 360–66.
5. Barchitta M, Quattrocchi A, Maugeri A, Vinciguerra M, Agodi A. LINE-1 hypomethylation in blood and tissue samples as an epigenetic marker for cancer risk: A systematic review and meta-analysis. *PLoS One* 2014; 9: e109478.
6. Muka T, Koromani F, Portilla E, O'Connor A, Bramer WM, Troup J, Chowdhury R, Dehghan A, Franco OH. The role of epigenetic modifications in cardiovascular disease: A systematic review. *Int J Cardiol* 2016; 212: 174–83.
7. Jakovcevski M, Akbarian S. Epigenetic mechanisms in neurological disease. *Nat Med* 2012; 18: 1194–204.
8. Brown SB, Hankinson SE. Endogenous estrogens and the risk of breast, endometrial, and ovarian cancers. *Steroids* 2014; 99: 8–10.
9. Barzi A, Lenz AM, Labonte MJ, Lenz HJ. Molecular pathways: Estrogen pathway in colorectal cancer. *Clin Cancer Res* 2013; 19: 5842–8.

10. Mendelsohn ME. Protective effects of estrogen on the cardiovascular system. *Am J Cardiol* 2002; 89: 12–7.

11. Arevalo M-A, Azcoitia I, Garcia-Segura LM. The neuroprotective actions of oestradiol and oestrogen receptors. *Nat Rev Neurosci* 2015; 16: 17–29.

12. Anderson KN, Schwab RB, Martinez ME. Reproductive risk factors and breast cancer subtypes: A review of the literature. *Breast Cancer Res Treat* 2014; 144: 1–10.

13. Jirtle RL, Skinner MK. Environmental epigenomics and disease susceptibility. *Nat Rev Genet* 2007; 8:253–62.

14. Zhang X, Ho SM. Epigenetics meets endocrinology. J Mol Endocrinol 2011; 46: R11–32.

15. Ulrich CM, Toriola AT, Koepl LM, Sandifer T, Poole EM, Duggan C, McTiernan A, Issa JPJ. Metabolic, hormonal and immunological associations with global DNA methylation among postmenopausal women. *Epigenetics* 2012; 7: 1020–8.

16. Iwasaki M, Ono H, Kuchiba A, Kasuga Y, Yokoyama S, Onuma H, Nishimura H, Kusama R, Yoshida T, Tsugane S. Association of postmenopausal endogenous sex hormones with global methylation level of leukocyte DNA among Japanese women. *BMC Cancer* 2012; 12: 323.

17. Varela-Rey M, Woodhoo A, Martinez-Chantar M-L, Mato JM, Lu SC. Alcohol, DNA methylation, and cancer. *Alcohol Res* 2013; 35: 25–35.

18. Anderson OS, Sant KE, Dolinoy DC. Nutrition and epigenetics: an interplay of dietary methyl donors, one-carbon metabolism and DNA methylation. *J Nutr Biochem* 2012; 23: 853–9.

19. Friedenreich CM, Woolcott CG, McTiernan A, Ballard-Barbash R, Brant RF, Stanczyk FZ, Terry T, Boyd NF, Yaffe MJ, Irwin ML, et al. Alberta physical activity and breast

cancer prevention trial: Sex hormone changes in a year-long exercise intervention among postmenopausal women. *J Clin Oncol* 2010; 28: 1458–66.

20. Friedenreich CM, Courneya KS, Neilson HK, Matthews CE, Willis G, Irwin M, Troiano R, Ballard-Barbash R. Reliability and validity of the past year total physical activity questionnaire. *Am J Epidemiol* 2006; 163: 959–70.

21. Csizmadi I, Kahle L, Ullman R, Dawe U, Zimmerman TP, Friedenreich CM, Bryant H, Subar AF. Adaptation and evaluation of the National Cancer Institute's Diet History Questionnaire and nutrient database for Canadian populations. *Public Health Nutr* 2007; 10: 88–96.

22. Willet W, Howe G, Kushi L. Adjustment for total energy intake in epidemiologic studies. *Am J Clin Nutr* 1997; 65: 1220S – 8S.

23. Friedenreich CM, Woolcott CG, McTiernan a, Terry T, Brant R, Ballard-Barbash R, Irwin ML, Jones C a, Boyd NF, Yaffe MJ, et al. Adiposity changes after a 1-year aerobic exercise intervention among postmenopausal women: a randomized controlled trial. *Int J Obes* 2011; 35:427–35.

24. Kile ML, Baccarelli A, Tarantini L, Hoffman E, Wright RO, Christiani DC. Correlation of global and gene-specific DNA methylation in maternal-infant pairs. *PLoS One* 2010; 5: e13730.

25. Endogenous Hormones Breast Cancer Collaborative Group. Body Mass Index, Serum Sex Hormones, and Breast Cancer Risk in Postmenopausal Women. *JNCI J Natl Cancer Inst* 2003; 95: 1218–26.

26. Yang AS, Estécio MRH, Doshi K, Kondo Y, Tajara EH, Issa J-PJ. A simple method for estimating global DNA methylation using bisulfite PCR of repetitive DNA elements. *Nucleic Acids Res* 2004; 32: e38.

27. Zhang X, Tworoger SS, Eliassen AH, Hankinson SE. Postmenopausal plasma sex hormone levels and breast cancer risk over 20 years of follow-up. *Breast Cancer Res Treat* 2013; 137: 883–92.

28. Hankinson SE, Manson JE, Spiegelman D, Willett WC, Longcope C, Speizer F. Reproducibility of plasma hormone levels in postmenopausal women over a 2-3 year period. *Cancer Epidemiol, Biomark Prev* 1995; 4: 649–54.

29. Ashbury JE, Taylor SA, Tse MY, Pang SC, Louw JA, Vanner SJ, King WD. Biomarkers measured in buccal and blood leukocyte DNA as proxies for colon tissue global methylation. *Int J Mol Epidemiol Genet* 2014; 5: 120–4.

30. Chalitchagorn K, Shuangshoti S, Hourpai N, Kongruttanachok N, Tangkijvanich P, Thong-ngam D, Voravud N, Sriuranpong V, Mutirangura A. Distinctive pattern of LINE-1 methylation level in normal tissues and the association with carcinogenesis. *Oncogene* 2004; 23: 8841–6.

31. Lowe R, Slodkowicz G, Goldman N, Rakyan VK. The human blood DNA methylome displays a highly distinctive profile compared with other somatic tissues. *Epigenetics* 2015; 10: 274–81.

32. Adalsteinsson BT, Gudnason H, Aspelund T, Harris TB, Launer LJ, Eiriksdottir G, Smith A V., Gudnason V. Heterogeneity in White Blood Cells Has Potential to Confound DNA Methylation Measurements. *PLoS One* 2012; 7: 1–9.

33. Zhu Z-Z, Hou L, Bollati V, Tarantini L, Marinelli B, Cantone L, Yang AS, Vokonas P, Lissowska J, Fustinoni S, et al. Predictors of global methylation levels in blood DNA of healthy subjects: a combined analysis. *Int J Epidemiol* 2012; 41:126–39.

**Table 5.1. Baseline Characteristics of Subset of ALPHA Trial Participants (N=289)**

| Variable | Mean, Median, or Frequency |
|---|---|
| Age (years) | |
| Mean (SD) | 61.0 (5.5) |
| 50 to 60 years, n (%) | 152 (52.6) |
| 61 to 75 years, n (%) | 137 (47.4) |
| Ethnicity | |
| Caucasian, n (%) | 263 (91.0) |
| Other, n (%) | 26 (9.0) |
| Body mass index, mean (SD) | 29.3 (4.4) |
| Body fat percent, mean (SD) | 42.5 (5.3) |
| Past Year Physical Activity (MET-hours per week), median (IQR) | 108 (77 – 149) |
| Alcohol Consumption (grams per day), median (IQR) [1] | 1.77 (0.55 – 5.24) |
| Energy-Adjusted Folate Intake (micrograms DFE / day), median (IQR) [1,2] | 510 (454 – 567) |
| Smoking Status | |
| Former Smoker, n (%) | 96 (33.2) |
| Non-smoker, n (%) | 193 (66.8) |
| Previous Hormone Replacement Therapy Use | |
| Yes, n (%) | 158 (54.7) |
| No, n (%) | 131 (45.3) |
| Hysterectomy | |
| Yes, n (%) | 73 (25.3) |
| No, n (%) | 216 (74.7) |
| Age at first birth | |
| Nulliparous, n (%) | 25 (8.7) |
| 16-24 years, n (%) | 154 (53.3) |
| 25-43 years, n (%) | 110 (38.1) |
| Presence of Comorbidities[3] | |
| Yes, n (%) | 110 (38.1) |
| No, n (%) | 179 (61.9) |

[1] Missing information for five participants

[2]One microgram dietary folate equivalent (DFE) = one microgram of food folate or 0.6 micrograms of folic acid from fortified foods

[3]Ever been diagnosed with at least one of the following: high cholesterol or triglycerides, myocardial infarction, angina pectoris, stroke, rheumatoid arthritis or osteoarthritis, osteoporosis, blood clots in the veins of legs or pelvis, blood clots in the lung, or thyroid problems.

**Table 5.2. Distribution of Sex Hormone Exposure and DNA Methylation Baseline Measures from ALPHA Trial Participants (N = 289)**

| Variable | Mean or Median |
|---|---|
| **Cumulative Sex Hormone Exposure** | |
| Cumulative Number of Menstrual Cycles (#), median (IQR) | 456 (395 - 485) |
| Age at menarche (years), mean (SD) | 12.8 (1.5) |
| Age at menopause (years), mean (SD) | 48.4 (6.3) |
| Parity (#), mean (SD) | 2.42 (1.29) |
| | |
| **Current Sex Hormone Exposure** | |
| Estradiol (pg/mL), median (IQR) | 10.0 (7.0 - 14.0) |
| Estrone (pg/mL), median (IQR) | 32.0 (24.0 - 45.0) |
| Testosterone (ng/dL), median (IQR) | 24.2 (17.4 - 32.4) |
| Androstenedione (pg/mL), median (IQR) | 571 (421 - 766) |
| Sex-Hormone Binding Globulin (nmol/L), median (IQR) | 40.0 (29.0 - 54.0) |
| | |
| **DNA Methyaltion** | |
| LINE-1 Methylation (%), mean (SD)[1] | 72.9 (0.9) |
| Alu Methylation (%), mean (SD)[1] | 19.1 (0.1) |

[1] The least squares mean and standard error of the least squares mean after adjusting for intra-plate variability

**Table 5.3. The Relation Between Current Sex Hormone Concentrations and Repetitive Element Methylation**

| | | LINE-1 Methylation (N = 284) | | | Alu Methylation (N = 288) | |
|---|---|---|---|---|---|---|
| | **Range or SD[1]** | **N** | **Adjusted Mean (%) or β[2,3] (95% CI)** | **P-value[5]** | **N** | **Adjusted Mean (%) or β[2,4] (95% CI)** | **P-value[5]** |
| **Estradiol (pg/mL)[6]** | | | | | | | |
| Categorical | < 8.0 | 74 | 72.54 (70.90 to 74.17) | 0.02 | 74 | 19.13 (18.84 to 19.42) | 0.13 |
| | 8.0 to < 11.0 | 85 | 73.01 (71.37 to 74.64) | | 85 | 19.15 (18.86 to 19.43) | |
| | 11.0 to <15.0 | 63 | 73.11 (71.46 to 74.77) | | 65 | 18.98 (18.69 to 19.27) | |
| | 15.0+ | 62 | 73.38 (71.72 to 75.05) | | 64 | 19.03 (18.74 to 19.33) | |
| | | | P-trend: 0.01 | | | P-trend: 0.10 | |
| Continuous | 12.0 | 284 | 0.21 (0.03 to 0.39) | 0.03 | 288 | -0.04 (-0.09 to 0.02) | 0.19 |
| **Estrone (pg/mL)[6]** | | | | | | | |
| Categorical | <25.0 | 78 | 72.75 (71.11 to 74.40) | 0.17 | 78 | 19.11 (18.82 to 19.40) | 0.58 |
| | 25.0 to < 33.0 | 76 | 72.75 (71.10 to 74.40) | | 76 | 19.11 (18.82 to 19.40) | |
| | 33.0 to < 46.0 | 62 | 72.93 (71.28 to 74.58) | | 63 | 19.01 (18.72 to 19.30) | |
| | 46.0+ | 68 | 73.32 (71.66 to 74.98) | | 71 | 19.07 (18.78 to 19.36) | |
| | | | P-trend: 0.049 | | | P-trend: 0.41 | |
| Continuous | 17.7 | 284 | 0.25 (0.04 to 0.46) | 0.02 | 288 | -0.01 (-0.08 to 0.05) | 0.65 |
| **Testosterone (ng/dL)[7]** | | | | | | | |
| Categorical | <17.4 | 71 | 73.44 (71.79 to 75.09) | 0.07 | 72 | 19.07 (18.78 to 19.36) | 0.36 |
| | 17.4 to <24.1 | 70 | 72.80 (71.16 to 74.45) | | 71 | 18.99 (18.70 to 19.28) | |
| | 24.1 to <32.5 | 74 | 72.95 (71.31 to 74.59) | | 74 | 19.08 (18.80 to 19.37) | |
| | 32.5+ | 69 | 72.85 (71.21 to 74.50) | | 71 | 19.14 (18.85 to 19.43) | |
| | | | P-trend: 0.07 | | | P-trend: 0.33 | |
| Continuous | 14.0 | 284 | -0.17 (-0.37 to 0.04) | 0.11 | 288 | 0.04 (-0.02 to 0.10) | 0.19 |
| **Androstenedione (pg/mL)[7]** | | | | | | | |
| Categorical | <421 | 71 | 73.12 (71.45 to 74.79) | 0.74 | 72 | 19.09 (18.80 to 19.39) | 0.51 |
| | 421 to <568 | 72 | 73.04 (71.37 to 74.70) | | 72 | 19.07 (18.78 to 19.37) | |
| | 568 to <763 | 70 | 72.86 (71.20 to 74.51) | | 72 | 19.00 (18.71 to 19.30) | |
| | 763+ | 71 | 73.09 (71.44 to 74.75) | | 72 | 19.11 (18.82 to 19.41) | |
| | | | P-trend: 0.80 | | | P-trend: 0.99 | |
| Continuous | 317.1 | 284 | 0.07 (-0.12 to 0.27) | 0.46 | 288 | 0.05 (-0.01 to 0.11) | 0.09 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **SHBG (nmol/L)[8]** | | | | | | | |
| Categorical | <30.0 | 73 | 72.77 (71.13 to 74.40) | 0.47 | 75 | 19.03 (18.74 to 19.32) | 0.15 |
| | 30.0 to <41.0 | 71 | 73.08 (71.44 to 74.72) | | 72 | 19.16 (18.87 to 19.45) | |
| | 41.0 to <55.0 | 74 | 73.02 (71.38 to 74.67) | | 74 | 19.11 (18.81 to 19.40) | |
| | 55+ | 66 | 73.17 (71.52 to 74.82) | | 67 | 18.99 (18.70 to 19.28) | |
| | | | P-trend: 0.18 | | | P-trend: 0.53 | |
| Continuous | 19.4 | 284 | 0.04 (-0.14 to 0.23) | 0.66 | 288 | -0.02 (-0.08 to 0.04) | 0.46 |

[1]The range is presented for categorical exposures and the SD is presented for continuous variables
[2]The adjusted least squares mean estimate is presented for categorical variables and the expected unit increase in percent methylation per one SD increase in the exposure after adjusting for confounders is presented for continuous variables
[3]All analyses adjusted for LNMC, past year alcohol consumption, and ethnicity
[4]All analyses adjusted for LNMC and smoking status
[5]P-value corresponding to the significance of the categorical or continuous variable
[6]Additionally adjusted for testosterone and SHBG
[7]Additionally adjusted for estradiol and SHBG
[8]Additionally adjusted for estradiol and testosterone

**Table 5.4. The Relation between Cumulative Ovarian Sex Hormone Exposure and Repetitive Element Methylation**

| | | LINE-1 Methylation (N = 284) | | | Alu Methylation (N = 288) | | |
|---|---|---|---|---|---|---|---|
| | **Range or SD[1]** | **N** | **Adjusted Mean (%) or β[2,3] (95% CI)** | **P-value[5]** | **N** | **Adjusted Mean (%) or β[2,4] (95% CI)** | **P-value[5]** |
| **LNMC (#)** | | | | | | | |
| Categorical | <395 | 70 | 73.2 (71.6 to 74.8) | 0.02 | 73 | 19.1 (18.8 to 19.4) | 0.07 |
| | 395 to <456 | 70 | 72.6 (70.9 to 74.2) | | 70 | 19.1 (18.8 to 19.3) | |
| | 456 to <485 | 73 | 73.0 (71.4 to 74.6) | | 73 | 19.0 (18.7 to 19.3) | |
| | 485+ | 71 | 73.3 (71.7 to 74.9) | | 72 | 19.1 (18.9 to 19.4) | |
| | | | P-trend: 0.32 | | | P-trend: 0.78 | |
| Continuous | 84.9 | 284 | - | - | 288 | -0.04 (-0.09 to 0.02) | 0.16 |
| **Age at menarche (years)** | | | | | | | |
| Categorical | <12 | 48 | 72.8 (71.1 to 74.5) | 0.052 | 49 | 19.0 (18.7 to 19.3) | 0.21 |
| | 12 | 80 | 73.3 (71.7 to 75.0) | | 80 | 19.2 (18.9 to 19.4) | |
| | 13 | 81 | 73.1 (71.4 to 74.7) | | 83 | 19.1 (18.8 to 19.4) | |
| | 14+ | 75 | 72.7 (71.0 to 74.4) | | 76 | 19.0 (18.7 to 19.3) | |
| | | | P-trend: 0.25 | | | P-trend: 0.32 | |
| Continuous | 1.5 | 284 | -0.07 (-0.24 to 0.11) | 0.44 | 288 | -0.04 (-0.09 to 0.02) | 0.17 |
| **Age at menopause (years)** | | | | | | | |
| Categorical | <45 | 56 | 73.3 (71.7 to 75.0) | 0.02 | 58 | 19.2 (18.9 to 19.5) | 0.21 |
| | 45 to <50 | 63 | 72.5 (70.9 to 74.2) | | 64 | 19.1 (18.8 to 19.4) | |
| | 50 to <55 | 131 | 73.1 (71.5 to 74.7) | | 132 | 19.1 (18.8 to 19.3) | |
| | 55+ | 34 | 73.2 (71.5 to 74.9) | | 34 | 19.0 (18.6 to 19.4) | |
| | | | P-trend: 0.71 | | | P-trend: 0.04 | |
| Continuous | 6.4 | 284 | - | - | 288 | -0.05 (-0.11 to 0.01) | 0.08 |
| **Parity (#)** | | | | | | | |
| Categorical | 0 | 25 | 72.8 (71.1 to 74.5) | 0.15 | 25 | 19.1 (18.9 to 19.5) | 0.56 |
| | 1 | 29 | 73.0 (71.3 to 74.7) | | 29 | 19.0 (18.7 to 19.3) | |
| | 2 | 105 | 72.9 (71.2 to 74.5) | | 107 | 19.1 (18.8 to 19.4) | |
| | 3+ | 125 | 73.3 (71.6 to 74.9) | | 127 | 19.1 (18.8 to 19.4) | |
| | | | P-trend: 0.07 | | | P-trend: 0.68 | |
| Continuous | 1.3 | 284 | 0.21 (0.03 to 0.38) | 0.02 | 288 | -0.01 (-0.06 to 0.05) | 0.86 |

[1]The range is presented for categorical exposures and the SD is presented for continuous variables

[2]The adjusted least squares mean estimate is presented for categorical variables and the expected unit increase in percent methylation per one SD increase in the exposure after adjusting for confounders is presented for continuous variables
[3]Adjusted for past year alcohol consumption, ethnicity, estradiol, testosterone, and SHBG
[4]Adjusted for smoking status, estradiol, testosterone, and SHBG
[5]P-value corresponding to the significance of the categorical or continuous variable

**Table 5.5. Assessment of Effect Modification by Folate and Alcohol Consumption (N = 284)**

| | LINE-1 Methylation | | | | | |
|---|---|---|---|---|---|---|
| | **Low Folate[1]** | **High Folate[1]** | | **Low Alcohol[1]** | **High Alcohol[1]** | |
| | **Mean or β (95% CI)[2]** | **Mean or β (95% CI)[2]** | **P-value[3]** | **Mean or β (95% CI)[2]** | **Mean or β (95% CI)[2]** | **P-value[3]** |
| **Estradiol** | 0.20 (-0.004 to 0.41) | 0.24 (-0.13 to 0.61) | 0.85 | 0.17 (-0.23 to 0.56) | 0.21 (0.004 to 0.41) | 0.84 |
| **Estrone** | 0.30 (0.05 to 0.56) | 0.18 (-0.12 to 0.47) | 0.49 | 0.23 (-0.06 to 0.52) | 0.24 (-0.01 to 0.50) | 0.95 |
| **Testosterone** | -0.05 (-0.32 to 0.22) | -0.28 (-0.56 to -0.01) | 0.20 | -0.29 (-0.56 to -0.02) | -0.03 (-0.31 to 0.24) | 0.17 |
| **Androstenedione** | 0.13 (-0.10 to 0.36) | -0.03 (-0.34 to 0.28) | 0.41 | 0.06 (-0.17 to 0.30) | 0.08 (-0.22 to 0.38) | 0.92 |
| **SHBG** | 0.29 (0.03 to 0.55) | -0.17 (-0.41 to 0.08) | 0.01 | -0.08 (-0.32 to 0.17) | 0.21 (-0.06 to 0.48) | 0.12 |
| **LNMC** | | | | | | |
| <395 | 73.0 (71.3 to 74.7) | 73.4 (71.7 to 75.1) | 0.70 | 73.1 (71.4 to 74.7) | 73.35 (71.7 to 75.1) | 0.45 |
| 395 to <456 | 72.7 (71.0 to 74.4) | 72.5 (70.8 to 74.1) | | 72.2 (70.6 to 73.9) | 72.9 (71.2 to 74.6) | |
| 456 to <485 | 73.0 (71.3 to 74.6) | 73.0 (71.4 to 74.7) | | 72.8 (71.2 to 74.5) | 73.16 (71.5 to 74.8) | |
| 485+ | 73.3 (71.6 to 75.0) | 73.3 (71.6 to 75.0) | | 73.4 (71.7 to 75.1) | 73.24 (71.6 to 74.9) | |
| | **Alu Methylation** | | | | | |
| | **Low Folate[1]** | **High Folate[1]** | | **Low Alcohol[1]** | **High Alcohol[1]** | |
| | **Mean or β (95% CI)[2]** | **Mean or β (95% CI)[2]** | **P-value[3]** | **Mean or β (95% CI)[2]** | **Mean or β (95% CI)[2]** | **P-value[3]** |
| **Estradiol** | -0.04 (-0.10 to 0.03) | -0.06 (-0.17 to 0.05) | 0.72 | -0.06 (-0.18 to 0.06) | -0.03 (-0.10 to 0.03) | 0.65 |
| **Estrone** | -0.04 (-0.12 to 0.03) | 0.01 (-0.08 to 0.10) | 0.34 | -0.03 (-0.12 to 0.06) | -0.01 (-0.09 to 0.07) | 0.75 |
| **Testosterone** | 0.02 (-0.07 to 0.10) | 0.07 (-0.01 to 0.16) | 0.31 | 0.05 (-0.03 to 0.14) | 0.04 (-0.05 to 0.12) | 0.80 |
| **Androstenedione** | 0.05 (-0.02 to 0.12) | 0.06 (-0.04 to 0.15) | 0.90 | 0.06 (-0.01 to 0.13) | 0.04 (-0.05 to 0.13) | 0.66 |
| **SHBG** | 0.03 (-0.05 to 0.11) | -0.06 (-0.14 to 0.01) | 0.09 | -0.03 (-0.10 to 0.05) | -0.01 (-0.10 to 0.07) | 0.82 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **LNMC** | | | | | | |
| <395 | 19.2 (18.9 to 19.5) | 19.0 (18.7 to 19.3) | 0.03 | 19.1 (18.8 to 19.4) | 19.2 (18.9 to 19.5) | 0.12 |
| 395 to <456 | 19.1 (18.8 to 19.4) | 19.1 (18.8 to 19.4) | | 19.1 (18.8 to 19.4) | 19.1 (18.8 to 19.4) | |
| 456 to <485 | 19.1 (18.8 to 19.4) | 18.9 (18.6 to 19.2) | | 18.9 (18.6 to 19.2) | 19.0 (18.7 to 19.3) | |
| 485+ | 19.0 (18.7 to 19.3) | 19.3 (19.0 to 19.6) | | 19.3 (19.0 to 19.6) | 19.0 (18.7 to 19.3) | |
| Continuous | -0.10 (-0.18 to -0.03) | 0.03 (-0.05 to 0.10) | 0.02 | 0.01 (-0.06 to 0.09) | -0.09 (-0.18 to -0.01) | 0.16 |

[1]Median folate cut-point = 510 micrograms DFE/day; Median alcohol consumption cut-point = 1.77 g/day
[2]The adjusted least squares mean estimate is presented for categorical variables and the expected unit increase in percent methylation per one SD increase in the exposure after adjusting for confounders is presented for continuous variables. Refer to Table 3 and Table 4 for information regarding the variables adjusted for in each analysis.
[3]Significance of the interaction term in the multivariable model.

# Chapter 6

# Conclusion

## Summary of Findings

Both adiposity and endogenous sex hormones were related to repetitive element DNA methylation in the white blood cells of healthy, inactive, postmenopausal women. Body fat and serum estrogen concentrations had a similar impact on LINE-1 methylation with respect to the size, shape, and direction of effect (i.e. a standard deviation increase in the exposure resulted in a linear increase of approximately 0.20 to 0.30 in mean percent LINE-1 methylation). Lifetime number of menstrual cycles had a non-linear "U-shaped" relation with LINE-1 methylation. In contrast, cumulative exposure to ovarian sex hormones had a negative linear relation with Alu methylation but only amongst low folate consumers. Androgens did not have an impact on repetitive element DNA methylation in this study population.

## Limitations

A number of limitations with this cross-sectional analysis merit discussion. These include issues with respect to the assessment and interpretation of DNA methylation data as well as classic epidemiologic considerations.

DNA methylation was measured in blood leukocytes and an implied assumption is that this measurement is meaningful with respect to chronic diseases that arise in other tissue sites. Tissue specificity is a major unresolved issue within the field of epigenetic epidemiology. Although blood measures of DNA methylation have been associated with disease risk, (1) DNA methylation levels differ between tissues and the extent to which they are homogenous within a

tissue has yet to be established. Although DNA methylation within certain genomic regions have been shown to be conserved across different tissues, there is some evidence that repetitive element DNA methylation is tissue- and cell-type specific. (2–5) As such, these results must be interpreted with caution. Associations between adiposity and DNA methylation have been observed within several tissues (e.g. blood, fat, muscle, placenta, umbilical cord, and saliva) which suggests that this exposure may induce systemic epigenetic effects. (6) However, target tissues of interest such as breast or brain tissue may be more or less susceptible to epigenetic dysregulation induced by adiposity and sex hormones in comparison to blood and other surrogate tissues. As such, one cannot be certain whether the relations we observed persist in other tissues and to what extent. Future studies should assess the impact of adiposity, sex hormones, and other exposures of interest on levels of DNA methylation within multiple tissues to determine if such exposures have systemic epigenetic effects and the extent to which the effects detected in surrogate tissues mirror the effects in target tissues.

Another important limitation of this investigation is the lack of knowledge pertaining to the relation between DNA methylation and disease risk. Repetitive element hypomethylation is believed to be both a cause and a consequence of carcinogenesis. (7) As such, prospective cohort studies are needed to determine the relation between DNA methylation in healthy individuals and subsequent disease risk. (8,9) Epigenetic epidemiology is an area of research still in its infancy and there have been relatively few prospective studies completed to date. (10) This lack of etiologic research makes it difficult to assess the clinical relevance of these results and the relative impact of the exposures in relation to one another. For example, hip circumference had a larger impact on LINE-1 methylation in comparison to waist circumference but it cannot be determined if such a difference is meaningful. Standardized methods of assessment, reference

epigenomes, and prospective cohort studies are required in order to better understand the results of this study and other investigations into the determinants of DNA methylation.

Another limitation of this study is the current state of technology used to quantify DNA methylation. The assay that was used in this study provided a pooled summary measure for the degree of methylation within three repeated CpG sites for both LINE-1 and Alu elements. Although a method feasible for use in an epidemiologic context does not yet exist, higher resolution data would have allowed for more precision in describing the epigenetic effects of adiposity and sex hormones (e.g., the change in repetitive element DNA methylation within certain genomic regions, chromosomes, or alleles). Such an understanding may have had implications in terms of the interpretation of the results (e.g. the methylation of repetitive elements within certain genomics regions may be more or less susceptible to dysregulation and may be more or less relevant in terms of its impact on disease risk). Scientists should continue to develop assays that would allow for high-throughput analysis of DNA methylation within repetitive elements at a base-pair resolution such that the epigenetic effects of an exposure of interest can be better characterized.

Although reverse causality, recall bias, and selection bias can arise in cross-sectional settings, these classic epidemiologic biases are not an issue with respect to the internal validity of the current investigation. It is unlikely that the methylation outcomes under study would have impacted the hormonal or adiposity measures. The use of current hormonal and adiposity measures as proxies for the levels of these exposures that preceded the changes in DNA methylation would have resulted in non-differential misclassification which would have only made it more difficult to detect an effect if one was present. Therefore, the directionality of the relationship can be established despite the cross-sectional setting of the current study.

Recall and selection bias are also not a concern in the current study. The underlying study in which this thesis project is nested was completed in 2007. Quantification of the DNA methylation outcomes was done in 2015. At the time of recruitment, the assessment of the exposures, and the collection of the blood samples, none of the participants or research personal would have had knowledge about any of the women's DNA methylation levels. Therefore, there is no reason to suspect that any of the measures susceptible to recall error (e.g. weight at age 20) were biased because there is no reason to suspect that a women with, for example, high levels of LINE-1 methylation would tend to underestimate their weight to a greater extent than women with low levels of LINE-1 methylation. Selection bias arises when participants in a certain exposure/outcome stratum (e.g. people with high exposure levels and high levels of the outcome) have a higher probability of being included in the study in relation to individuals in other exposure/outcome categories. Since the probability of inclusion into the study cannot be related to the outcome under investigation, selection bias is not an issue.

**Strengths**

Despite the limitations of the current study, there are several strengths which will be briefly mentioned (more detailed descriptions can be found in Chapters 4 and 5). Relative to other studies in this field, the current investigation had a large sample size, a comprehensive list of exposures, and high quality measures. Although we were able to assess several potential confounders that were not considered in past research, uncontrolled confounding may be an issue as we do not yet know what to adjust for given the lack of research in this area.

## Generalizability

The current study was conducted using a highly unique study population (i.e. postmenopausal, healthy, women). Although it is not an issue with respect to the internal validity of this study, such selection bias reduces the external validity of these results. Sex hormone concentrations differ between men and women and between premenopausal and postmenopausal women. In diseased populations, the relation of interest may depend upon the type, the severity of the illness, and the treatment administered to the patient. Therefore, the results of this study should not be generalized to men, premenopausal women, or diseased populations.

## Implications and Future Directions

These results add to the growing body of evidence showing how lifestyle factors impact the epigenome. The determinants of DNA methylation and the biological windows during which the epigenome is most modifiable are poorly understood. (8,9) Some research suggests that environmental factors experienced in utero and during early childhood can induce long lasting epigenetic aberrations that affect the risk of disease later in life. (11) The results from this investigation are promising in that they support the notion that behavioural interventions in older populations have the potential to offset epigenetic predispositions acquired early in life. Future research should explore the effects of sex hormones and adiposity on DNA methylation within other genomic regions and at different time points. Researchers interested in studying the effects of other exposures on the epigenome should consider adiposity and sex hormones as potential confounders, particularly when dealing with a postmenopausal population.

The findings of this study were inconsistent for LINE-1 and Alu methylation which suggests that these measures are two distinct epigenetic phenomena. In the literature, LINE-1

and Alu methylation are typically viewed as surrogate markers of global DNA methylation. (12) Our results suggest that the determinants of DNA methylation differ depending on the genomic region under investigation. In particular, LINE-1 methylation was more susceptible to the effects of body fat and sex hormones than Alu methylation which was only associated with cumulative exposure to ovarian sex hormones. This observation is supported by other research reporting different epigenetic effects of an exposure across different repetitive elements. (13) Similarly, the impact of DNA methylation on disease risk often differ with respect to the repetitive element under investigation in that LINE-1 methylation is more strongly related to the risk of certain diseases than Alu methylation and vice versa. (14,15) As such, LINE-1 and Alu methylation are best thought of as independent biological events. This notion is supported by the fact that these retrotransposons occupy distinct genomic regions and have several distinguishing characteristics such as their method of transposition, copy number, genomic coverage, sequence length, structure, and degree of methylation. (16–18)

There is evidence that the duration of estrogen exposure is particularly relevant when investigating repetitive element DNA methylation. We found that cumulative rather than current estrogen exposure was related to Alu methylation. In a premenopausal population, El-Maarri et al. (2011) found that blood levels of LINE-1 and Alu methylation were not affected by changes in sex hormones that occurred over the course of one full menstrual cycle. (19) Taken together, these results suggest that repetitive element methylation may be impervious to short term changes in estrogen exposure which is consistent with the literature demonstrating the relative stability of these epigenetic marks over time. (20,21) Long-term rather than short-term estrogen exposure may therefore be primarily responsible for inducing changes in repetitive element DNA methylation.

Our results suggest that LINE-1 methylation may be a mechanism through which estrogen exposure confers a protective effect with respect to colorectal cancer, cardiovascular disease, and possibly neurodegeneration amongst postmenopausal women. The positive relation observed between adiposity and LINE-1 methylation may be explained by its effect on sex hormone concentrations. After adjusting for estrogens, we found that the effects of body fat were attenuated which suggests that estrogens may mediate the relation between adiposity and LINE-1 methylation in postmenopausal women (data not shown). Future research should also explore the extent to which the relation between adiposity and DNA methylation is mediated by sex hormones and other mechanisms such as chronic inflammation and oxidative stress.

**Conclusion**

Lower levels of repetitive element DNA methylation may predispose individuals to developing cancer and other illnesses. Adiposity and sex hormones appear to have a role in maintaining high levels of DNA methylation within LINE-1 repeats in postmenopausal women. This protective effect may be a mechanism through which estrogens reduce the risk of certain illnesses. Our findings suggest that high folate consumption may attenuate the potentially harmful effects of long term premenopausal estrogen exposure associated with lower levels of Alu methylation. The results from this investigation are preliminary and should be interpreted with caution until future research has clarified the role of repetitive element DNA methylation in the initiation of cancer, cardiovascular disease, and neurodegenerative disorders as well as the extent to which the relations we observed can be generalized to other tissues of interest.

## References


1. Woo HD, Kim J. Global DNA hypomethylation in peripheral blood leukocytes as a biomarker for cancer risk: A meta-analysis. *PLoS One* 2012; 7(4): e34615.
2. 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* 2012; 41(1): 126–39.
3. 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; 23(54): 8841–6.
4. Delgado-Cruzata L, Vin-Raviv N, Tehranifar P, Flom J, Reynolds D, Gonzalez K, et al. Correlations in global DNA methylation measures in peripheral blood mononuclear cells and granulocytes. *Epigenetics* 2014; 9(11): 1504–10.
5. Ashbury JE, Taylor SA, Tse MY, Pang SC, Louw JA, Vanner SJ, et al. Biomarkers measured in buccal and blood leukocyte DNA as proxies for colon tissue global methylation. *Int J Mol Epidemiol Genet* 2014; 5(2): 120–4.
6. Van Dijk SJ, Molloy PL, Varinli H, Morrison JL, Muhlhausler BS. Epigenetics and human obesity. *Int J Obes* 2014; 39:1–13.
7. Miousse IR, Koturbash I. The Fine LINE: Methylation Drawing the Cancer Landscape. *Biomed Res Int* 2015; 2015: 1–8.
8. Ng JW, Barrett LM, Wong A, Kuh D, Smith G, Relton CL. The role of longitudinal cohort studies in epigenetic epidemiology: challenges and opportunities. *Genome Biol* 2012; 13(6): 246.

9. 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; 6(7): 828–37.

10. Verma M, Rogers S, Divi RL, Schully SD, Nelson S, Su LJ, et al. Epigenetic Research in Cancer Epidemiology: Trends, Opportunities, and Challenges. *Cancer Epidemiol Biomarkers Prev* 2014; 23(2): 223–33.

11. Wadhwa PD, Buss C, Entringer S, Swanson JM. Developmental origins of health and disease: Brief history of the approach and current focus on epigenetic mechanisms. *Seminars in Reproductive Medicine* 2009; 27(5): 358–68.

12. Yang AS, Estécio MRH, Doshi K, Kondo Y, Tajara EH, Issa J-PJ. A simple method for estimating global DNA methylation using bisulfite PCR of repetitive DNA elements. *Nucleic Acids Res*; 2004; 32(3): e38.

13. 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; 6(7): 828–37.

14. Gao Y, Baccarelli A, Shu XO, Ji B-T, Yu K, Tarantini L, et al. Blood leukocyte Alu and LINE-1 methylation and gastric cancer risk in the Shanghai Women's Health Study. *Br J Cancer* 2012; 106(3): 585–91.

15. Barry KH, Moore LE, Liao LM, Huang WY, Andreotti G, Poulin M, et al. Prospective study of DNA methylation at LINE-1 and Alu in peripheral blood and the risk of prostate cancer. *Prostate* 2015; 1725(15): 1718–25.

16. Miousse IR, Chalbot M-CG, Lumen A, Ferguson A, Kavouras IG, Koturbash I. Response of Transposable Elements to Environmental Stressors. *Mutat Res* 2015; 765: 19–39.

17. Padeken J, Zeller P, Gasser SM. Repeat DNA in genome organization and stability. *Curr Opin Genet Dev* 2015; 31: 12–9.

18. Rollins R a, Haghighi F, Edwards JR, Das R, Zhang MQ, Ju J, et al. Large-scale structure of genomic methylation patterns Large-scale structure of genomic methylation patterns. *Genome Res* 2006; 16(2): 157–63.

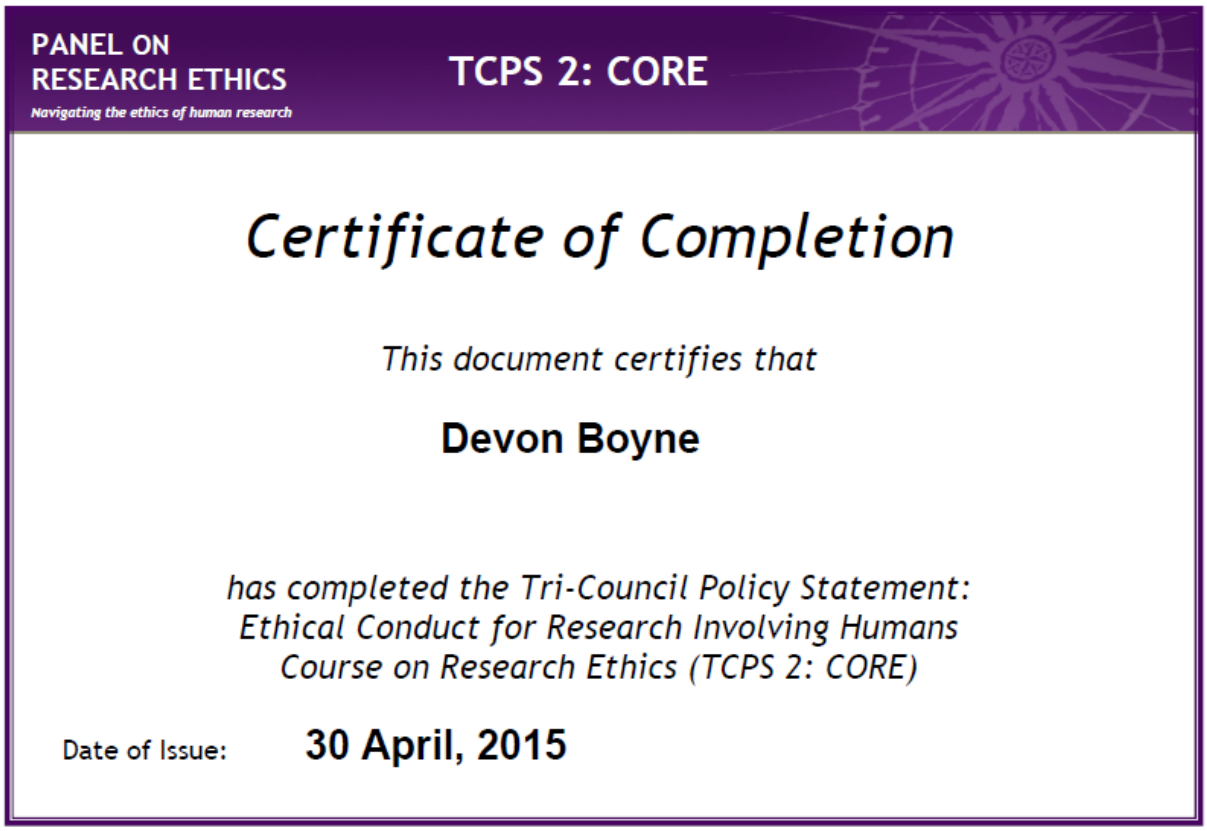
19. El-Maarri O, Walier M, Behne F, Van Üüm J, Singer H, Diaz-Lacava A, et al. Methylation at global LINE-1 repeats in human blood are affected by gender but not by age or natural hormone cycles. *PLoS One* 2011; 6(1): e16252.

20. Byun HM, Nordio F, Coull BA, Tarantini L, Hou L, Bonzini M, et al. Temporal stability of epigenetic markers: Sequence characteristics and predictors of short-term DNA methylation variations. *PLoS One* 2012; 7(6): 1–9.

21. 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; 130(4): 234–9.

# Appendix A

## Ethics Certificate and Letter of Approval



PANEL ON RESEARCH ETHICS
*Navigating the ethics of human research*

TCPS 2: CORE

*Certificate of Completion*

*This document certifies that*

**Devon Boyne**

*has completed the Tri-Council Policy Statement:*
*Ethical Conduct for Research Involving Humans*
*Course on Research Ethics (TCPS 2: CORE)*

Date of Issue: **30 April, 2015**

**QUEEN'S UNIVERSITY HEALTH SCIENCES & AFFILIATED TEACHING HOSPITALS RESEARCH ETHICS BOARD (HSREB)**

**HSREB Initial Ethics Clearance**

August 04, 2015

Mr. Devon Boyne
Department of Public Health Sciences
Queen's University

**ROMEO/TRAQ: #6015972**
**Department Code: EPID-518-15**
**Study Title: Examining the Relationship between Breast Cancer Risk Factors and DNA Methylation**
**Co-Investigators: Dr. W. King**
**Review Type: Delegated**
**Date Ethics Clearance Issued: August 04, 2015**
**Ethics Clearance Expiry Date: August 04, 2016**

Dear Mr. Boyne,

The Queen's University Health Sciences & Affiliated Teaching Hospitals Research Ethics Board (HSREB) has reviewed the application and granted ethics clearance for the documents listed below. Ethics clearance is granted until the expiration date noted above.

- Protocol
- Peer Review

**Documents Acknowledged:**

- CORE Certificate – D. Boyne
- ALPHA Trial Manual of Operations – July 7, 2015

**Amendments:** No deviations from, or changes to the protocol should be initiated without prior written clearance of an appropriate amendment from the HSREB, except when necessary to eliminate immediate hazard(s) to study participants or when the change(s) involves only administrative or logistical aspects of the trial.

**Renewals:** Prior to the expiration of your ethics clearance you will be reminded to submit your renewal report through ROMEO. Any lapses in ethical clearance will be documented on the renewal form.

**Completion/Termination:** The HSREB must be notified of the completion or termination of this study through the completion of a renewal report in ROMEO.

**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.

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.

Yours sincerely,

Albert J. Clark.

Chair, Health Sciences Research Ethics Board

*The HSREB operates in compliance with, and is constituted in accordance with, the requirements of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2); the International Conference on Harmonisation Good Clinical Practice Consolidated Guideline (ICH GCP); Part C, Division 5 of the Food and Drug Regulations; Part 4 of the Natural Health Products Regulations; Part 3 of the Medical Devices Regulations, Canadian General Standards Board, and the provisions of the Ontario Personal Health Information Protection Act (PHIPA 2004) and its applicable regulations. The HSREB is qualified through the CTO REB Qualification Program and is registered with the U.S. Department of Health and Human Services (DHHS) Office for Human Research Protection (OHRP). Federalwide Assurance Number: FWA#:00004184, IRB#:00001173*

*HSREB members involved in the research project do not participate in the review, discussion or decision.*