Psychosocial Factors and Obesity: Examining the Impact of Genetic Predisposition and Epigenetic Regulation.

by

Emon Elboudwarej

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Epidemiology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Mahasin S. Mujahid, Chair
Professor Darlene D. Francis
Professor Arthur L. Reingold

Fall 2016

Psychosocial Factors and Obesity: Examining the Impact of Genetic Predisposition and Epigenetic Regulation.

Abstract

Psychosocial Factors and Obesity: Examining the Impact of Genetic Predisposition and Epigenetic Regulation.

by

Emon Elboudwarej

Doctor of Philosophy in Epidemiology

University of California, Berkeley

Professor Mahasin S. Mujahid, Chair

Background

Obesity is a metabolic condition characterized by high levels of abdominal adiposity and currently affects approximately 35% of all US adults. Obesity is commonly measured using the body mass index (BMI), measured in kg/m^2, with a level ≥ 30 indicating obesity. There are significant sex and racial/ethnic disparities in obesity; the highest levels of age-adjusted obesity are seen in non-Hispanic Black women (~58%) compared to the lowest levels seen, in non-Hispanic White women (~32%). Obesity is associated with many negative health outcomes, such as atherosclerotic cardiovascular disease (CVD) and various cancers (e.g. breast, colon, and endometrial), and has also been shown to be an independent risk factor for all-cause mortality.

Given the multifactorial etiology of obesity, examining multiple pathways that incorporate biological, behavioral, and environmental effects on weight gain may provide insights that lead to the prevention of obesity. Previous studies have demonstrated modest effects of psychosocial factors (e.g. job-related demands, relationship demands, and personal health problems) on the risk of obesity. Additionally, genome-wide association studies (GWAS) on obesity have identified up to 97 genetic risk variants that are associated with a high BMI. The effects of BMI-associated genetic risk variants detected thus far have been weak, with the strongest predictor, the fat mass and obesity gene (*FTO*), contributing a 0.39 point increase in BMI for each copy of the risk allele, explaining only ~0.34% of total genetic variance.

Two types of research studies that might help expand on the limited findings of previous research and help incorporate the effect of psychosocial factors and genetic factors together are gene by environment interaction studies and epigenetic studies (i.e. DNA methylation (DNAm) studies). Animal models have shown that acute stressors (e.g. maltreatment or maternal neglect) alter DNAm in mice and rats. Less is known about the interaction of psychosocial factors, genetic risk factors for obesity, and epigenetic regulation of those genetic risk factors in humans. However, genetic susceptibility to obesity, together with high levels of external stressors, may increase the risk of obesity and account for a previously unexplained proportion of the variance in obesity.

Methods

This dissertation uses two analytical approaches to investigate the relationship between psychosocial risk factors and genetic risk factors and obesity: a systematic review and secondary data analyses. Chapter 2 presents a detailed systematic review of available studies of gene-environment interaction studies and epigenetic studies that focus on the effects of psychosocial and genetic risk factors for obesity. Chapters 3 and 4 use secondary data from a large population-based longitudinal cohort study, known as the Multi-Ethnic Study of Atherosclerosis (MESA).

Chapter 3 explores the interaction of three psychosocial factors (i.e. chronic burden of stress, everyday hassles, and depression) and an obesity genetic risk score on obesity. Obesity genetic risk is derived from the most recent meta-analysis, which established 97 independent genetic variants associated with body mass index (BMI). This analysis was conducted using interval-censored survival modeling using a Weibull distribution. Both multiplicative and additive effects were determined, so as to give a comprehensive assessment of genetic and psychosocial interactions.

Chapter 4 investigates the effect of three psychosocial factors (i.e. chronic burden of stress, everyday hassles, and Cohen's perceived stress) on DNAm levels of obesity risk genes. I employed a two-level model for each unique gene (87 genes from 97 obesity GWAS SNPs), treating the CpG as the level-1 unit and the individual as the level-2 unit. For genes associated with psychosocial variables, we assess cross-sectional associations between DNAm and genetic expression levels. An association between DNAm and expression demonstrates the functional importance of DNAm as a gene regulator.

Significance

This dissertation examines the interactive relationship between genetic and non-genetic factors as they relate to obesity. Previous findings for BMI-associated genetic risk factors have been relatively weak. The effects of psychosocial factors on obesity have been primarily examined through mechanisms that involve behavior change, such as altered diet patterns or altered physical activity. Assessing the epigenetic effects and gene-environment interactions of psychosocial factors and obesity genetic risk may reveal pathways through which people develop a greater risk for obesity.

My findings provide evidence of an interaction between psychosocial factors and genetic risk for obesity in multiple subpopulations of the MESA cohort study. I identify several BMI-associated genes that are differentially methylated by levels of chronic stress. DNAm is significantly associated with genetic expression, revealing a functional mechanism by which exogenous factors affect genetic expression, not directly attributed to inherited genomic sequences. These findings suggest potential underlying biological mechanisms whereby psychosocial factors and genetic risk factors interact to cause obesity in a manner that is not mediated by altered behavioral patterns of energy intake and expenditure. As inherited genomic sequences are not easily modifiable to prevent negative health conditions, it is important to establish multiple systems where we can prevent additional cases of obesity by targeting a modifiable risk factor (e.g. psychosocial stress) that interacts with genes or affects their expression. Because obesity remains a major concern in the United States, investigators should continue to search for mechanisms through which it can occur, in order to help reduce the burden of obesity.

Dedication

To my mother, Mama, for her unwavering love and support that have motivated me to be an empathetic individual who cares about the health and well-being of all people. To my father, Baba, for his love and commitment to higher education, which inspired me to think critically about the problems we can solve in the world. To my brother, Omes, who showed me how to work hard for what matters. I am forever grateful to my family for supporting me.

I love you with all my heart.

Table of Contents

Acknowledgements

I would like to first and foremost thank my advisor and mentor, Dr. Mahasin Mujahid, for her support and guidance throughout my doctoral program. Her commitment to public health research with thoughtful insight helped navigate me towards issues that mattered to me personally and that would benefit the current body of knowledge on health disparities. Her critical thinking and unwavering commitment to mentorship at multiple levels of educational training have inspired my desire not only to help the greater public, but to those closest to me as well. Dr. Mujahid's commitment of time and dedication to my research, as well as to my academic, professional, and personal development over the past five years have been invaluable.

Thank you to my doctoral committee members, Dr. Darlene Francis for her support and engagement on the topic of biological embedding of social factors, and Dr. Arthur Reingold for his impact on my development and understanding of fundamental epidemiology. I also would like to thank the members of my qualifying exam, Drs. Barbara Abrams, Ralph Catalano, and Maya Petersen for their contributions to developing the dissertation topic.

Thank you to all collaborators from the Multi-Ethnic Study of Atherosclerosis (MESA) that shaped the analytical plan of multiple studies. Thank you to Dr. Lisa Barcellos, whose mentorship motivated my interest in genetic epidemiology and who supported me throughout each level of my public health training at UC Berkeley. Thank you to my colleague and friend, Dr. Farren Briggs, for his support and mentorship throughout every level of my training and beyond.

I would also like to thank my fellow epidemiology doctoral students from UC Berkeley and all of the other faculty members who have shaped my training. I will be forever grateful for your help and insight.

To my family, thank you for blessing me with the opportunities that most are not fortunate enough to have and for ensuring that I was supported in every possible way, so as to guarantee success in my life and career.

Chapter 1

Introduction

1.1 Background information

Obesity is a medical condition that generally arises from an imbalance of energy intake and energy expenditure. Most studies categorize weight classes defined by the World Health Organization, using the body mass index (BMI) measure. Adolphe Quetelet first created BMI in the mid 19th century for studying what he referred to as social physics.[1] An astronomer and mathematician from Belgium, Quetelet discovered that adult body weight is proportional to the square of one's height and was interested in measuring adiposity as a way of characterizing the overall health of a population.[2] Different levels of BMI can be reported categorically (in kg/m^2), with overweight classified as having a BMI between 25.0 and 29.9, and obesity classified as any value 30.0 and above; In the year 2000, obesity was subcategorized into three classes: class I (moderately obese, BMI 30.0 to 34.9); class II (severely obese, BMI 35.0 to 39.9); class III (very severely obese, BMI 40.0 and above).

Obesity is a major public health concern affecting approximately 35% of all US adults. Examining the current trends in obesity in the US, Flegal et al. established that while the incidence of obesity has been plateauing in recent years, compared to the period from 2003-2008, the prevalence of obesity remains at an all time high.[3] Data from the 2012 National Health and Nutrition Examination Survey (NHANES) showed a 33.6% prevalence of overweight among US adults aged 20 and older, and a prevalence of obesity of 34.9%.[3] That is, almost 70% of the US adults are either obese or overweight. There are also statistically significant racial and ethnic differences in obesity. Studies have shown that non-Hispanic Black (47.8%) and Hispanic individuals (42.5%) have a significantly higher prevalence of obesity compared to non-Hispanic Whites (32.6%).[4]

According to the National Heart Lung and Blood Institute (NHLBI), obesity is associated with many adverse health conditions, including atherosclerotic cardiovascular disease (CVD), cancer (e.g. breast, colon, and endometrial), infertility in women, and metabolic syndrome (www.nhlbi.nih.gov). Obesity has also been established as an independent risk factor for all-cause mortality.[5] Because CVD is the leading cause of death in the United States,[6] more attention has been paid to obesity as a contributing factor in recent years. Further, the American Medical Association officially recognized obesity as a "disease" in 2013, in order to increase the urgency with which researchers and public health professionals address the epidemic.[7]

In addition to the morbidity and mortality associated with obesity, the economic burden placed on the US health-care system is enormous. Using data from the U.S. Medical Expenditure Panel Survey (MEPS), Cawley and colleagues determined that in 2005, the direct cost of obesity-related health problems in the U.S. (e.g. treatment for asthma, prescription drugs for weight loss, bariatric surgery, and nutrition counseling) was an estimated $190 billion.[8] By 2006, obesity was estimated to be responsible nearly $86 billion in health-care associated spending each year, or 10%

of all public medical costs (8.5% of Medicare and 11.8% of Medicaid spending).[9] With the prevalence of obesity increasing each year, costs are projected to increase as well.

Risk factors for obesity

Although the exact causes of obesity remain unknown, there has been a great deal of research establishing myriad risk factors for obesity, commonly divided into biological, behavioral, and social factors. According to the Mayo Clinic and NIH, individual level determinants of obesity include, but are not limited to, genetic heritability, excessive eating, and an inactive or sedentary lifestyle.[10, 11] Social and economic factors have also been associated with obesity (e.g. cultural norms concerning body size, diet and food affordability), as well as neighborhood factors, such as safety and walkability. (www.mayoclinic.org)

Obesity is best conceptualized in the context of a framework in which multiple networks (biological, behavioral, and social) interact to determine weight. In a recent review of public health paradigms and multilevel frameworks for disease onset, Glass et al. (2006) pooled the results of 22 studies that have established multiple etiological pathways for obesity onset, encompassing biological, behavioral, and environmental effects on weight gain.[12] A multi-level framework for contributors to obesity is presented in Figure 1 (section 1.4). The figure highlights the health behaviors related to eating patterns and physical activity that have been extensively examined in relation to obesity. Also shown are biological factors such as Hypothalamic-Pituitary-Adrenal (HPA) activity, a hormone regulatory system for stress reactivity and metabolism, which affects weight through health behaviors regarding diet and physical activity, which are affected by mood or appetite.[12] Societal determinants, such as increased urbanization, economic growth, globalization of food markets, and food availability have been previously associated with obesity as well.[13]

Psychosocial stress vs. stressors

Stress can be operationalized in four general ways, as described by Kasl et al. [14]:

1) By an environmental condition (objective measure of a stressor).
2) By an appraisal of an environmental condition (perceived stress)
3) By a response to an environmental condition or appraisal (biological stress response)
4) As an interactive term between environmental demands and a person's capacity to meet those demands (adaptive ability).

This dissertation focuses on two categories of stress: objective measures of an environmental condition (i.e. the long-term accumulation of stressors such as difficulties with your job, health, or relationships), and the appraisal of an environmental condition (i.e. the perception of stress or unfair treatment). Psychology research in the early 1980's examined objective measures of stressors that were classified into six distinct domains: work life, health, personal life, emotions, actions or behaviors, and life as a whole.[15] This can be advantageous when attempting to quantify the amount of stress experienced in a way that is comparable across individuals, as

perceived stress can be highly variable from person to person. The distinction between the condition of stress and measurable events that can be termed stressors is important to consider, as the literature is not consistent with how the impact of stress on health is best conceptualized. Kasl et al. also noted this lack of consensus across multiple studies, where there was no clear indication that objective events or subjective psychosocial perception were more meaningful measures for determining the impact stress on one's health.[14] For this reason, the current dissertation focuses on both measures separately (objective and subjective) in order to account for potential discrepancies due to measurement, and not the effect of stress itself.

Psychosocial stressors and obesity

Psychosocial stressors, defined as social or environmental exposures or demands that place a burden on the adaptive capabilities of an individual,[16] may be risk factors for obesity. The link between psychosocial stressors and CVD has been examined extensively in previous research.[17, 18] A 2005 review described multiple cases studies where acute psychosocial stressors (e.g. earthquakes or terrorist attacks) led to significant increases in sudden cardiac arrest.[17] Caregiver strain and occupational stress (e.g. high job demands, low job control) were also consistently associated with cardiovascular disease outcomes such as increased progression of atherosclerosis, or incident coronary heart disease from multiple longitudinal studies using validated measures of stress and clinically measured outcomes.[17] However, the link between psychosocial stressors and obesity is less well-established.

A 2004 review by Overgaard et al. examined ten cross-sectional studies that found only a weak or no association between psychosocial stressors, in the job and work-related domain, and obesity.[19] The review focused on one job-related stressor (job demands) and two job-related factors: job latitudes (skill discretion and decision authority), and job strain (high demands, with low influence). Only one out of eight studies examining job demands showed a statistically significant relationship between the stressor and obesity. Additionally, only two of ten studies found a statistically significant relationship between at least one of the job-related factors and obesity.

Longitudinal data can provide stronger evidence of a cause effect relationship between psychosocial stressors and obesity (i.e. that the stressor preceded the onset of obesity). Two of the largest nationally representative, population-based, longitudinal studies, the Coronary Artery Risk Development in Young Adults (CARDIA) study and the Midlife in the United States (MIDUS) study, have examined whether social and psychosocial factors such as job-related demands, job immobility, and poor social relationships lead to weight gain and an increased likelihood of obesity.[20, 21] In 2011, Wardle et al. conducted a meta-analysis of 14 longitudinal cohort studies examining the stressor-obesity relationship in adults. They limited the analysis to prospective cohort studies that examined stressors such as caregiver strain (i.e. providing basic care for children or the elderly), and work-related stressors. While the meta-analysis found a statistically significant effect of stressors on objective measures of adiposity (e.g. waist circumference), the relationship was observed only in males, not females, and the effect sizes were small (e.g. males from 15 studies had a combined effect size r=0.024, 95% CI (0.006,0.042)). [22]

There have been four large population-based longitudinal cohort studies, outside the US that have examined the relationship between psychosocial stressors and obesity. These studies were not included in the meta-analysis by Wardle et al. Much like the findings from cross-sectional studies, results from longitudinal studies were inconsistent. As early as 1986, Van Strien et al. established that negative life events (e.g. a death in the family, spousal separation, or financial problems), were associated with weight gain over a six-month follow-up period.[23] However, again the observed effect was significant in only men, and focused primarily on the effect of stressors and subsequent emotional eating behaviors that would contribute to obesity (i.e. the diet-mediated effect of stress and obesity).[23]

A study by Harding et al. of 5,118 participants from the Australian Diabetes Obesity and Lifestyle study (AusDiab) showed that increased psychosocial stress s associated with weight gain over a five-year follow up period.[24] Stress in this study was measured two ways: first, by a "perceived stress questionnaire" comprised of 30 items (e.g. feelings of tension over the previous 12-month period) and second, a 13-item stressful life events questionnaire, summed over a 12-month period. Harding et al. found a string association between both "perceived stress" and life event stressors and weight gain over five years in both men and women. Two studies examining the Whitehall Cohort II population also found that work-stress was positively associated with subsequent change in BMI.[25, 26] However, results varied by baseline BMI categories, where lean men with higher stress exhibited weight *loss* and obese men with higher stress exhibited weight *gain* over the follow up period, with no interaction observed in women.[26]

One of the most well known longitudinal studies of obesity in the United States, conducted by Block et al., suggested a role of psychosocial stressors for the recently observed increases in obesity.[20] Block et al. found that psychosocial stressors such as "job-related demands", and "difficulty paying bills" were significantly associated with subsequent weight gain in men.[20] Additionally, measures such as "perceived constraints in life," and "strains in relations with family," were significantly associated with subsequent weight gain in women.

The conflicting evidence across multiple longitudinal cohort studies and multiple cross-sectional studies indicates that other potential mechanisms linking stressors and obesity should be investigated. As mentioned previously, obesity is conceptualized in a multi-level framework, with many contributing factors (section 1.3; Figure 1). Evidence provided by Block and colleagues highlights the need for expanding obesity research to include psychosocial factors that contribute to obesity, above and beyond the more proximal factors traditionally associated with weight gain, such as diet and activity. Despite conflicting results of observational studies in humans, we know from extensive research using experimental models in rats and mice that there is biological plausibility for how the biological response to stressors can lead to obesity.[27] These biological mechanisms by which stressors can induce obesity will be explored in the following section.

Potential mechanism linking psychosocial stressors to obesity

One mechanism by which stress is believed to influence obesity is through a relationship between stress and eating behavior. A review in 2014 summarized the current arguments concerning the

relationship between stress and eating and found that increased stress alters eating habits, but not necessarily in a consistent fashion. Some individuals increase food intake when they experience stress, while others decrease food intake.[28] This review however, did not differentiate between acute and chronic stress. Acute stress is defined as a physiological stress response to an unexpected, single event brought on by a stressor (e.g. being attacked by a predator)[29] Chronic stress is the persistence of acute stressors over time, which elicits a unique psychophysiological response different from that resulting from acute stress.[29] The distinction between acute and chronic stress is important to keep in mind, as conflicting results from studies researching the stress-obesity relationship may be partly attributed to the type of stress being examined.

A systematic review conducted in 2012 by Moore et al. showed overwhelming support for the hypothesis that high levels stress are associated with less healthy eating habits;[30] it should be noted that this was based on cross-sectional studies only and included studies that examined objective stressors (job demands) *and* subjective stress (Cohen's perceived stress scale[31] and Karasak's demand/control model[32]). However, another review in 2014 showed that eating habits are affected differently by acute and chronic stress.[33] Acute stress is thought to suppress appetite through the release of corticotropin-releasing hormone (CRH), which initiates an immediate response to fight against danger, which is evolutionarily advantageous.[33] Alternatively, when a stressor is activating the stress response chronically, glucocorticoids levels will be elevated for prolonged periods of time and activate lipoprotein lipase in adipose tissue, which leads to increased fat storage.[33] The evolutionary and biological reasoning for fat storage would lead one to assume that there should be unequivocal evidence that those experiencing chronic stressors over the long term are more likely to store fat and thus, become obese. This is not necessarily the case, as evidenced by the previous studies summarized thus far. Therefore, additional mechanisms must be examined.

Genetics factors in obesity

In recent years, a strong argument has been made for a genetic contribution to obesity in humans. The biological and evolutionary basis for appetite and fat storage due to hormone release invokes the role of genes, as we know that stress-response biomarkers (e.g. glucocorticoids, such as cortisol) are heavily regulated by genetic factors.[34] Once again, this highlights the multifactorial nature of weight gain.

Many genetic epidemiologic studies, including genome linkage, fine mapping, and candidate gene studies have identified several single nucleotide polymorphisms (SNPs) associated with BMI. Additionally, genome wide association studies (GWAS) have improved upon other approaches by allowing for the examination of a wide range of common variants associated with obesity. A recent meta-analysis of GWA studies has identified up to 97 genetic predictors of high BMI, including the strongest genetic factor discovered to date, the fat mass and obesity gene, *FTO*.[35, 36] As opposed to a single gene's contribution to disease, discovering a network of genetic factors and pathways of genetic interaction highlights the potential for polygenic mechanisms contributing to obesity.[37] Modern technologies have allowed for a comprehensive assessment of the genetic contribution to obesity, with many genes working together to lead to obesity.

GWAS discoveries have not, however, identified very strong predictors of obesity. For example, genetic variation in the strongest predictor, *FTO*, confers only a 0.39-point increase of BMI for each copy of the risk allele. This would explain only 0.34% of the total genetic variance with respect to obesity.[35] In order to address this limitation, Belsky and colleagues developed a method for examining total polygenic genetic risk using an obesity genetic risk score (oGRS) and have found encouraging results.[38] The oGRS can be used for trans-ethnic replication (i.e. replication in populations of varied ancestry), even if the risk variants were established in populations of European descent,[39] making it an attractive measure for generalizing genetic risk to broader populations. Still, the genetic role in obesity is not enough to explain even a majority of obesity cases, and it is believed that interactions between genes and environment provide more insight into etiology of disease. Therefore, it is important to establish other risk factors that work independently of (or synergistically with) genetic factors that increase the likelihood of being obese.

Gene-environment interactions and obesity

One way of examining the multifactorial causal connection between weight gain and obesity is by conducting studies that examine the interaction of endogenous and exogenous factors. Interest in the gene by environment interaction literature suggests a genetic predisposition for obesity works synergistically with non-genetic factors to increase the risk of obesity. However, this literature remains focused on the interactions of gens and dietary factors or socioeconomic status.[40] A study was conducted by Andreasen et al. exploring the interactive effect of *FTO* and physical activity on obesity, found that carriers of the A allele at the rs9939609 locus showed a significant increase in BMI with physical passivity, compared to those homozygous for the A allele.[41] Additionally, a gene-diet interaction was established for *APOA5[42]* as well as for *ADRB3*,[43] and *IL6R*,[44] each showing an interaction with higher energy intake leading to increases in the risk of obesity.

Epigenetics and obesity

Understanding psychosocial stress from a vantage point that includes genetic interactions is important and timely. Recent studies have shown that genes responsible for regulating the stress response (e.g. CRH and glucocorticoids such as cortisol) can be differentially expressed in those with and without obesity, potentially due to epigenetic regulation.[45] Examining the impact of stress on DNA methylation and genetic expression can help to elucidate alternative complex mechanisms linking psychosocial factors and obesity.

Epigenetics has recently emerged as a mechanism by which genetic expression influences can affect disease risk. Epigenetics is a term used for anything that might be an un-inherited regulation of inheritable genetic transcripts. This may occur in many ways, from chromosomal binding and telomere lengths[46, 47] to histone acetylation[48] or DNA methylation. DNA methylation, which is defined as the addition of a methyl group (CH_3) to cytosine/guanine base pairs, may help to understand how genes are being expressed in the body; this is due to the regulatory nature of DNA

methylation, as it is impeding or promoting genetic transcription and is affected by exogenous factors. The growing interest in epigenetic regulation is supported by the fact that non-genetic exposures, such as diet, smoking and environmental toxins, are directly associated with altered DNA methylation levels.[49-51] Increased DNA methylation leads to differential expression of genes by blocking transcription[52] (i.e. preventing a gene from being expressed or "turned on" in the body).

Methylation is a recurring process that is believed to be dynamic throughout the life course, with evidence showing that methylation is reversible[53, 54]. Examining factors that contribute to methylation can reveal modifiable molecular mechanisms for genetic expression that transcend inheritance of genomic sequences. The potential of psychosocial factors affecting DNA methylation can illuminate how regulation of genetic expression by stress is contributing to obesity in humans. The present dissertation focuses on DNA methylation as the regulatory agent of genetic expression.

Stress, obesity, and DNA methylation

The impact of stressors on DNA methylation has been thoroughly documented in animal models, with numerous studies showing how acute stressors (e.g. maltreatment or maternal neglect) alter DNA methylation in mice[55, 56] and rats.[57] The relationship between stress and methylation has not been as clear in humans, as observational studies do not allow for similar control of cause and effect where subjects would be experimentally stressed and then measured for subsequent changes in DNA methylation. Sasaki et al provided a thorough review of studies on stressors and DNA methylation in 2013.[58] They examined studies that assessed the relationship between DNA methylation and the following stress/stressors/social factors: abuse and neglect in early life and maternal care in early life lastly, differing DNA methylation patterns of glucocorticoid receptor genes was observed in the cord blood of children with mothers who experience high stress during pregnancy.

The majority of studies included in the review by Sasaki et al. focused on acute stressors and early life events. Thus, there is a significant gap in the literature with respect to measures of chronic stress and DNA methylation in adults. One study of DNA methylation differences in monozygotic twins, conducted by Fraga et al., showed that DNA methylation significantly changed between co-twins over their life course.[59] Their study showed that despite similar experiences in early life, given that twins are generally reared together, DNA methylation became markedly different in adulthood.[59]

An epigenome wide analysis was recently conducted in nearly 500 subjects of European ancestry, found five DNA methylation markers that were associated with obesity.[60] Further, Zhao et al. in a 2014 twin study examined methylation patterns in stress genes and found a significant association with obesity at the promoter region of serotonin transporter gene, *SLC6A4*.[61] Zhao and colleagues found associations between methylation of genes related to stress response and obesity, indicating that potentially different life stressors or management of the stress response can

be different in genetically similar people. In other words, the environment impacted obesity through DNA methylation in a way previously unexplained by genetic sequences alone.

Prior research has highlighted how early life exposures to stressful events can result in observable epigenetic modifications later in life, by broad increases in DNA methylation across the entire genome and at specific genes.[62, 63] Less is known about modifications that occur due to psychosocial stressors experienced in adulthood. While there have been previous studies investigating the role of DNA methylation on obesity, to my knowledge, none has conducted such a study from a candidate gene approach based on obesity risk genes discovered from GWA studies. Given the dynamic nature of the factors contributing to obesity, understanding the epigenetic mechanisms will be essential in bridging the divide between socio-environmental exposures and genetic predisposition to obesity.

Limitations of existing research

Numerous studies have established a connection between psychosocial factors and obesity via coping mechanisms that influence one's diet and/or physical activity. However, few have assessed the interaction of psychosocial factors and the genetic risk of obesity. Furthermore, no studies to date have examined the potential effect of psychosocial factors on DNA methylation of all GWAS derived obesity risk genes.

A potential criticism of observational methylation studies in humans is that it can be too difficult to determine the exact cause of the DNA methylation change. DNA methylation is a dynamic and ongoing process (demethylation and remethylation occur daily and have the potential to persist), and exactly how the changes occur or persist is still being established. As mentioned previously, countless exposures can be causing methylation levels to change, and being able to isolate risk factors and quantify their impact on methylation can be challenging; however, it is necessary in studying the genetic determinants of obesity.

A majority of epigenetic and DNA methylation studies have focused on developmental stages of life and early childhood, given that a great deal of genetic programming (e.g. imprinting and x-inactivation), occurs during fetal development.[64] However, psychosocial factors that may be influencing DNA methylation can occur at any time point throughout an individual's life, including when relationship difficulties or job-related stressors are present. As most healthy individuals are susceptible to obesity at any time, it is important to characterize how psychosocial factors in adulthood might be contributing to overall risk of obesity.

1.2 Specific Aims

The overall goal of this dissertation is to improve our understanding of the link between psychosocial factors and obesity by examining the role of genetic predisposition and epigenetic regulation of genetic risk factors. My primary, secondary, and tertiary specific aims are the following:

1) Examine the current state of the literature on the relationship between genomic and psychosocial factors associated with obesity.
2) Examine associations between self-reported psychosocial factors and obesity risk, and whether associations are modified by genetic risk in a longitudinal study.
    a. Hypotheses:
        i. Individuals with higher baseline levels of psychosocial factors will have higher rates of obesity over the study follow-up time, independent of potentially confounding covariates.
        ii. Associations between psychosocial factors and obesity will be modified by genetic risk. An obesity genetic risk score will modify the effects of psychosocial factors, such that associations will be strongest among those with a high genetic risk.
3) Examine associations between psychosocial factors and DNA methylation in genetic risk variants associated with obesity.
    a. Hypotheses:
        i. There will be a statistically significant association between psychosocial factors and DNA methylation levels of obesity genes.
        ii. There will be a statistically significant association between DNA methylation levels of obesity genes and gene expression levels.

The first aim will be investigated by conducting a systematic review of publicly available studies. Aims 2 and 3 will be achieved using secondary data analyses of the longitudinal cohort study known as the Multi-Ethnic Study of Atherosclerosis (MESA). The MESA dataset is from a population-based, multi-ethnic cohort, with measures of psychosocial factors, genetic variants and DNA methylation, in addition to objective measures of obesity.

1.3 Tables and Figures

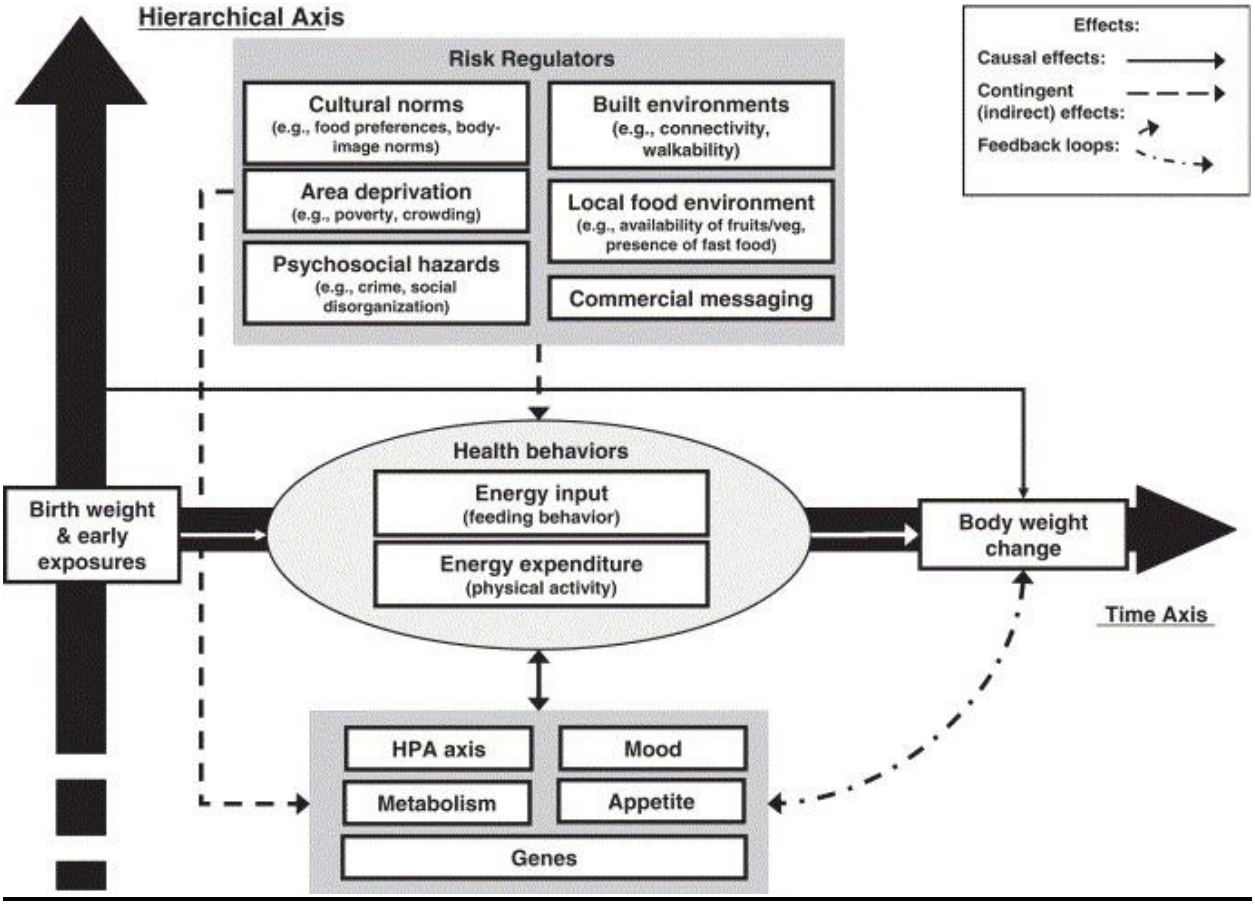


Figure 1. Multi-level framework for obesity risk

Chapter 2

The contribution of genomic and psychosocial factors to obesity

2.1 Introduction

Obesity, commonly measured by body mass index or BMI in (kg/m^2), has become a significant public health problem in modern America, affecting approximately 35% of all US adults.[3] Examining the current trends in obesity in the US, Flegal et al. established that while the prevalence of obesity is plateauing in more recent years, compared to the period from 2003-2008, the prevalence of obesity remains at all time highs.[3] Sex and racial/ethnic disparities in obesity remain a significant concern, with nearly a two-fold increase in the age-adjusted prevalence among non-Hispanic Black women (~58%) compared to non-Hispanic White women (~32%).[3] The impact of obesity on important health outcomes is significant. Obesity is associated with an increased risk of many negative health conditions, such as atherosclerotic cardiovascular disease (CVD), various types of cancer (e.g. breast, colon, and endometrial), reproductive infertility in women, and metabolic syndrome (www.nhlbi.nih.gov), among others. Obesity has also been established as an independent risk factor for all-cause mortality.[5]

Obesity is a key example of a health condition with a multifactorial etiology. Multiple networks (biological, behavioral, and social) may act independently or together to affect weight. In a recent review of public health paradigms and multilevel frameworks for disease onset by Glass et al. (2006), multiple pathways for obesity etiology that incorporated biological, behavioral, and environmental effects on weight gain were described.[12] Glass and colleagues highlighted the health behaviors related to eating patterns and physical activity that have been extensively examined in relation to obesity.

Indeed, the two most well-established, proximal risk factors of obesity are energy intake and energy expenditure.[65] Nevertheless, significant disparities persist in the population by factors not directly related to diet or physical activity.[66, 67] A seminal paper by Rand and Kuldau showed significant differences in obesity by race and sex, age group, and socioeconomic status.[68] This paper established that myriad social factors, and not merely heritable traits, could have a lasting impact on the risk of obesity. However, only a few studies have previously examined the effect of psychosocial factors (e.g. depression, relationship strain, difficulty paying bills, and job related demands), on obesity.[20, 21, 69, 70]

Notwithstanding limited or conflicting evidence in humans,[19] previous animal studies have shown that chronic exposure to stressors increases levels of glucocorticoids in the blood, which in turn leads to excess fat storage. We also know from extensive experimental studies in rats and mice where the biological response to stressors over an extended period of time can lead to adiposity, independent of dietary intake or level of physical activity.[27] Thus, there is a scientific basis for investigating the role of more distal factors that contribute to obesity in humans.

Prior studies have shown that neither psychosocial nor behavioral factors alone can explain the total variance in BMI. Genetic contributions to obesity research have expanded as obesity is a highly heritable trait and can be polygenic or a rare, familial, monogenic condition.[71] Genome-wide association studies (GWAS), which expanded dramatically around the late 2000's, led to multiple discoveries of single nucleotide polymorphisms (SNPs) that were associated with a wide array of diseases and conditions, including autoimmune disorders such as multiple sclerosis and systemic lupus erythematosus and metabolic conditions, such as type 2 diabetes mellitus.[72]

Numerous obesity GWAS have been conducted in the past decade, with the most recent meta-analysis conducted by Locke et al. in 2014. This meta-analysis identified 97 genetic predictors of high BMI (56 novel loci) from 339,224 individuals.[36] Genetic findings will continue to expand as risk variants are constantly being discovered, due to the greater power of larger studies. The utility of this constant expansion, however, is not yet clear. Despite the many genetic variants discovered for obesity, genetic variation in even the strongest predictor, the fat mass and obesity gene, or *FTO*, contributed to only a 0.39 point increase in BMI for each copy of the risk allele, which can explain only ~0.34% of the total genetic variance.[35] While genetic findings have been limited, it is not to say that the genetic contribution to obesity is negligible. Given the highly expensive procedures required to discover these genes (approximately of $125,000 per discovered locus)[72] and the growing desire to treat obesity in a genetically personalized way,[37] different methods that can address a larger proportion of the obesity variance must be prioritized.

A distinct area of research that unites genetic and non-genetic determinants of disease is Gene by Environment (GxE) interactions. GxE studies look at the joint effect of genetic variants and exogenous factors, often referred to as the environment, but which can be any non-genetic factor. With respect to obesity research in the GxE framework, Choquet et al. found that the environmental components of many studies have focused on physical activity, dietary habits, age, sex, and race/ethnicity[73] However, studies to date have not extensively examined interactions between genetic and psychosocial social factors, which is an important limitation in the field. There are plausible biological mechanisms by which stress and gene interactions may lead to obesity. Animal models have shown that chronic exposure to stress can cause the prolonged release of stress-response hormones such as glucocorticoids.[74] Glucocorticoids are regulated by genetic factors,[34] and persistent high levels in the blood can lead to fat storage.[74] Therefore, the genetic influence on stress-response, coupled with an environment of high levels of external stressors, may increase the risk of obesity and characterize a previously unexplained proportion of the obesity variance.

Epigenetics has also emerged as a mechanism by which genetic risk factors for obesity might be over-expressed or under-expressed in people with high levels of psychosocial stress. Although obesity is a highly heritable trait, as evidenced by concordance rates derived from studies utilizing monozygotic twins,[75] heritability is not simply genetic inheritance.[76, 77] The lack of strong effects in single variants mentioned previously has contributed to the widespread desire to discover the "missing heritability" of complex genetic diseases, whereby another distinct area of research, epigenetic inheritance, has been discussed as a possible avenue of exploration.[78] Epigenetics is a term used for the inherited or un-inherited regulation of genetic expression that is not due to changes in inherited genetic material. This may occur in many ways, from chromosomal binding and telomere length[46, 47] to histone acetylation[48] to the most commonly studied mechanism,

DNA methylation.[79] DNA methylation has become the most prominent aspect of epigenetics to emerge in recent studies due to the its relative economical advantages over more costly approaches, expansive genome-wide coverage thanks to modern microarray technologies,[80] and its direct and lasting impact on gene expression.[81] DNA methylation is a dynamic process that fluctuates constantly and occurs at every stage of life,[82] making it an attractive candidate for public health interventions even well into adulthood.

The effect of stressors on DNA methylation has been thoroughly documented in animal models, with numerous studies showing how acute stressors (e.g. maltreatment or maternal neglect) alter DNA methylation in mice[55, 56] and rats.[57] The relationship between stress and methylation has not been as clear in humans, as observational studies do not allow for similar control of cause and effect where subjects would be experimentally stressed and then measured for subsequent changes in DNA methylation. However several studies have shown that adverse childhood events (e.g. abuse and neglect) can alter phenotypic trajectories[83] and can carry the impact on DNA methylation well into adulthood[84] through broad increases in DNA methylation across the entire genome and at specific genes.[62, 63] Less is known about modifications that occur due to psychosocial stressors experienced in adulthood.

Incorporating genetic and non-genetic approaches into research to discover complex, multifactorial biological pathways is essential for understanding obesity in a comprehensive way. Intervening on environmental (non-genetic) factors that interact with (or block the function of) genetic markers is possible immediately, whereas genetic findings may not translate into a direct public health intervention in the near term. The growth of obesity as a major public health concern supports the need for determining all potential avenues by which obesity may arise and be prevented. The overall goal of this study was to conduct a systematic review of epidemiologic research examining the intersection of genetic and psychosocial factors in relation to obesity. Specifically, I reviewed studies that examined: 1) if the interaction of genetic and psychosocial factors is synergistically associated with obesity and 2) if epigenetic pathways (via regulation by DNA methylation) could identify a new mechanism by which psychosocial factors lead to obesity. We assessed the strengths and limitations of the current body of literature on this topic and make recommendations for future studies.

2.2 Methods

This review was done according to Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) Guidelines.

Search strategy

An electronic search of the National Center for Biotechnology Information (NCBI) PubMed database and Google scholar was conducted separately for GXE studies and epigenetic studies published between January 1, 2007 and June 1, 2016. These dates were selected based on a review by Visccher et al. that showed that the first GWAS was reported in 2005, but the expansion of

GWAS publications was initiated in 2007.[72] However, the expansion of GWAS publications was initiated in 2007 by the Welcome Trust Case Control Consortium (WTCCC), comprising a large dataset that was designed specifically for array-based SNP discovery with significant coverage of the human genome.[72, 85]

The searches included a combination of Medical Subject Headings, or MeSH terms, and terms open to "All Fields" in order not to omit potential studies that did not contain exact MeSH terminology. Obesity can be measured in a multitude of ways; therefore, we allowed for our search to include terms such as body-mass index (BMI), waist-hip ratio (WHR) and waist circumference (WC). Terms for psychosocial factors included but were not limited to "psychological stress", "stressor", or subcomponents of commonly used stress scores, such as "discrimination", "hassles" or "job strain."[86, 87] We decided to also include terms such as "SES", "education", "depression", and "income" in order to include studies that might have emphasized these measure but also looked at psychosocial factors in a sub-analysis.

Essential to our main goal for this review is synthetizing information on gene-environment interactions, as they incorporate the joint effects of each exposure, and not merely studies that address genetic and non-genetic components separately. Therefore, we specifically searched for the MeSH term "Gene-Environment Interaction," which is a well-known category of genetic studies. For the epigenetics search, we used identical terms that were common between the two study types, but switched out "Gene-Environment Interaction" with either "Epigenetics" or "DNA Methylation." Only human studies published in English were searched, as there are many animal models in genetic research that do not provide relevant information about self-reported psychosocial factors and obesity. The complete PubMed search criteria are available in Supplementary table S1.

Selection of studies and inclusion/exclusion criteria

For the GxE articles, inclusion criteria were based on studies of human adults (aged 19 years or older) published in English between January 1, 2007 and June 1, 2016. This resulted in 111 articles for title and abstract review (Figure 1). We additionally excluded studies in the following prioritized order: 1) Studies with the wrong outcome, meaning any study that did not consider obesity or some measurement of obesity as it's main or secondary outcome. 2) Studies that were focused on any population other than human adults (i.e. animals or children). This reduced the number of studies to 31 GxE articles that were retained for full review. We next excluded articles in the following prioritized order: 1) any studies that did not include psychosocial factors as their environmental exposure. 2) Any studies that were the wrong study type, meaning mediation analyses, gene-gene interaction studies, methods papers, Narrative reviews, commentary, abstracts only, or studies without primary data. We allowed for the non-systematic inclusion of cited references from our relevant articles, which added one study, yielding a final set of four GxE studies for review and summarization.

For Epigenetic studies, inclusion criteria were based on studies of human adults (aged 19 years or older) published in English between January 1, 2007 and June 1, 2016. This produced 106 articles

for title and abstract review (Figure 1). The search strategy for Epigenetic studies was designed to retrieve articles that assessed either the impact of DNA methylation on obesity, the impact of psychosocial factors on DNA methylation of obesity risk variants, or the combination of all three components. Exclusion criteria were prioritized in the following order: 1) Studies with the wrong outcome, meaning any study that did not consider obesity or some measurement of obesity as the main or secondary outcome. 2) Studies that were focused on any population other than human adults. This resulted in 35 epigenetic articles to be considered for full review. We then excluded articles in the following prioritized order: 1) any study that did not have some aspect of objectively or subjectively measured psychosocial factors incorporated into it, 2) any studies that were the wrong study type, meaning mediation analyses, methods papers, narrative reviews, commentaries, abstracts only, or studies without primary data. We allowed for less stringent criteria on the final set of epigenetic articles in order to incorporate studies that looked at genes involved in pathways that are associated with psychosocial factors and biological response mechanisms to stressors,[88-90] but did not measure psychosocial factors directly. An additional study was non-systematically included from cited references, resulting in a final set of five epigenetic studies for detailed review and summarization. A flowchart of the article filtering process was created using the Google drive web tool, draw.io diagrams (section 2.5; Figure 1).

2.3 Results

GxE review

Study characteristics

Our GxE search resulted in two longitudinal cohort studies,[91, 92] and two case-control studies.[93] Each study was comprised of a varied set of demographic characteristics: Kring 2010 (N=126 stressed caregivers/122 non-stressed, white men and women),[94] Marmorstein et al. (N=903 men, 521 women of various race/ethnicities (white, black, Native-American, Asian-Pacific Islander, Hispanic),[91] Singh et al. (N=5,805 men and women of various race/ethnicities (white, black, Native-American, Hispanic),[92] and Kring 2011 (N=475 obese cases/709 non-obese, all White men). Each study assessed sex and race/ethnicity by self-report. With the exception of the Marmorstein study, which had an overall mean age of 21.9 years at the end of follow-up,[91] the rest of the studies were comprised of a generally older population of adults with mean ages of ~49,[93] ~62,[92] and ~63 years of age.[94]

Study outcomes and exposure measurement

Three of the four studies examined psychosocial factors measuring validated measures of chronic stress: the chronic burden of stress scale (Singh, 2014),[92] chronically stressed caregivers (Kring, 2010),[94] and another seven-item chronic stress score (Kring, 2011).[93] The fourth study conducted by Marmorstein and Hart considered receipt of public assistance in childhood as evidence of a stressor.[91] Each study looked at a combination of obesity measures, including hip circumference,[92] waist circumference,[93, 94] and body-mass index (BMI).[91] A particular

strength of the studies examined in the current review is that each used measures of chronic psychosocial factors that preceded obesity measurements, establishing a clear temporal sequence of events.

Genetic factors

None of the four GxE papers incorporated information from obesity risk genes that have been derived from genome-wide association studies (GWAS) or within the strongest candidate gene to date, *FTO*.[35] Singh et al. employed a genome-wide approach as a method of discovering novel loci associated with obesity (as measured by hip circumference).[92] Kring et al. 2010 selected Apolipoprotein E (*APOE*) as an attractive candidate gene because of its relationship to Type 2 diabetes mellitus and cardiovascular conditions, as well as its role in lipid metabolism.[94] Kring et al. 2011 selected *APOE* in an effort to replicate prior findings. Last, Marmorstein et al. selected the *MAOA* gene, given its established role in moderating responses to stressors and for its potential role in obesity in males.[91]

Main findings

Each of the four GxE studies found a statistically significant interaction between a psychosocial factor and genetic determinant of obesity. Quantitative results revealed a total of 8 statistically significant positive interactions, from three unique genes (*EBF1*, *APOE*, *MAOA*), multiple measures of psychosocial factors, and multiple measures of obesity. While Singh et al. did not report specific effect sizes, the investigators showed five SNPs within the Early B-cell Factor 1 gene (*EBF1*) that interacted with a chronic burden of stress score, and reached genome-wide significance for hip circumference (p-value <9.46E-08).[92] The study was conducted in 2,460 White men and women. While data were available on other racial or ethnic groups, none of the gene-stress interactions was significant in non-whites. For the lead SNP (*EBF1* rs4704963:T>C) revealed that carriers of the minor allele (CC/CT) had a linear increase in waist girth, hip girth, and BMI at higher chronic stress levels. Each of three SNPs (rs17056278, rs17056298, and 17056398) that were also available in an independent sample from the Framingham Heart Study[95] were replicated for hip circumference (p-value <0.02), and validated by significant associations with BMI and waist circumference (p-value <0.02).

Two studies that employed a targeted candidate gene approach found significant GxE associations with waist circumference.[93, 94] Kring et al. found a statistically significant interaction between one of three genetic variants (*APOE*:rs439401) with caregiver status, in both additive (p-value=0.047) and dominant (p-value=0.026) genetic models.[94] Chronically stressed individuals who were homozygous for the minor T-allele at rs43940 had a statistically significantly greater waist circumference compared to individuals without stress. The study was conducted in 341 White women, and the findings were replicated in a subsequent study of 1,184 White men, using a self-report seven-item chronic stress score[93]. The replication study showed that a 5cm increase in waist circumference was significantly associated with a 9% increase in the odds of being homozygous for the minor T-allele among stressed men (OR=1.09 [1.01;1.17]). Similarly, a two-

unit increase in BMI level was significantly associated with rs43940 and chronic stress (OR=1.09 [1.02;1.17]).

Last, Marmorstein et al. found a relationship between the short allele for a variant in the X-linked gene, Monoamine oxidase A (*MAOA*) and a psychosocial stressor (receipt of public assistance) with BMI.[91] This study utilized 903 men of multiple racial/ethnic backgrounds. The interaction of *MAOA* with receiving public assistance alone was not statistically significant. However, the study found that age-dependent BMI trajectories (from adolescence into adulthood) were statistically different in those with the *MAOA* short allele that also received public assistance, compared to other groups.[91] These results expand on the previous GxE study findings by providing evidence of a functional polymorphism that interacts with a psychosocial factor to contribute to a more rapid increase in BMI level.

Limitations

Each study acknowledged limitations of their findings. The studies were primarily limited by the absence of replication, and at times referred to their studies as exploratory analyses, while stressing the need for continued work in the areas of research that integrate genetic and non-genetic factors. Only one investigator had the ability to replicate GxE interactions in an independent dataset[92, 93]

Each of the GxE studies was done primarily on White/European subjects due to the availability of samples,[61, 94, 96] maximizing statistical power or lack of main effects in non-white populations,[92] or the necessity to replicate prior findings in White populations.[93] Studies only conducted in White populations limit the ability to generalize the findings to other populations of different ancestries. Therefore, future studies should focus on replicating these findings in independent datasets that include racial/ethnic diverse groups. A primary concern in genome-wide analyses is the issue of making false discoveries from conducting many statistical tests (i.e. multiple testing concerns). Bonferonni corrected p-values were used in GxE studies in order to address this concern.[92] Additionally, the total number of tests performed were reduced by avoiding genome-wide scans, in favor of a targeted approach based on specific loci from candidate genes. [91, 93, 94]

As with most case-control studies, the potential for recall bias is certainly possible. The studies in the current review used a combination of self-reported measures of psychosocial factors,[91-94] self-reported height and weight,[91, 92] or clinical assessment of obesity measures.[93, 94] Self-reporting psychosocial factors is common practice, however, the potential for recalling stressful events differentially by those who have obesity versus those who do not, could lead to inflated effect sizes between stress and obesity. Furthermore, different scales for psychosocial stress makes it difficult to synthesize information across various studies, as investigators define surrogate measures for stress in myriad ways. Ultimately, the intention is to identify biological mechanisms underlying the relationship between stress and obesity through inflammatory pathways, serotonin response, or HPA activity/cortisol release. Measuring these biomarkers directly can be costly and

thereby limit the number of participants in a study, which effectively leads to decreased power for studies to identify statistically significant associations.

Epigenetic Review

Study characteristics

Our search resulted in five cross-sectional studies, (one of which also had longitudinal measures). Each study was comprised of a relatively small sample of various racial/ethnic groups: Na et al. (N=284, Korean women),[97] Perez-Cornago et al. (N=41, Spanish (White) of unspecified sex),[96] Zhang et al. (N=165, men and women of various race/ethnicities (non-Hispanic White, non-Hispanic Black, Hispanic)),[98] Gomes et al. (N=126, Brazilian men and women),[99] and Zhao et al. (N=168, White men).[61] Each study included individuals with a wide range of ages. While one study made no explicit reference to the mean age of participants,[96] the studies by Na et al. and Zhang et al. reported similar mean ages of 31.9 and 29.4 years respectively.[97, 98] The remaining two studies had older participants of 70.8 years[99] and 55 years of age, respectively.[61]

Study outcomes and exposure measurement

The final set of epigenetic articles from our systematic search included one study that assessed the exposure of mean (global) methylation[99] and four studies that targeted genes involved in pro-inflammatory and stress-response mechanisms.[61, 96-98] In cross-sectional studies, designation of the exposure and outcome was not always made explicit. These studies examined the effect of obesity (BMI >30) on *IL6* methylation,[97] global methylation and depression,[98] continuous BMI and depression with global methylation,[99] and the effect of *SLC6A* methylation on BMI, body weight, waist circumference, and waist to hip ratio.[61] The final study, examined associations between DNA methylation of the serotonin receptor 2A (*HTR2A*) gene and waist circumference, BMI, and body weight, in addition to longitudinal changes in fat mass or changes in depressive symptoms.[96]

Genetic factors

DNA methylation is known to be allele-specific,[35, 100-102] which is a limitation that is routinely addressed in epigenetic studies. Speculating that polymorphisms associated with methylation could be playing a confounding role, Perez-Cornago attempted to validate methylation levels with genetic expression levels, although results were not significant. Another method of dealing the genetic confounding of methylation levels was employed by Zhao et al. as this study was conducted in monozygotic twins, assuring that any paired differences in methylation were not attributable to varying genotypes. Failure to validate or replicate findings was a limitation of all of the epigenetic studies. Researchers took steps to address spurious DNA methylation findings by

validating methylation levels with more precise techniques, such as pyrosequencing,[61, 97] or by correlating methylation with gene expression levels.[96]

Main findings

In four of the five articles reviewed, a significant association was established between global or gene-specific *hyper*methylation and increases in various measures of obesity.[61, 96-98] The last study, conducted by Gomes et al., found no significant correlation between global methylation and BMI (p-value=0.12) or depression (p-value=0.83).[99] Contrary to non-significant findings of Zhang et al.,[98] Na and colleagues found a statistically significant association between *IL6* promoter *hyper*methylation and obesity in a cross-sectional study of 284 Korean women.[97] Zhao et al. showed that hypermethylation of promoter CpGs within the serotonin transporter gene (*SLC6A*) was associated with increased BMI, body weight and waist circumference, but not WHR, a point that the authors suggest is understandable, given consistent prior research that indicates WHR is a poor measure of abdominal adiposity.[61] Additionally, unadjusted cross-sectional analyses revealed a significant relationship between serotonin receptor 2A (*HTR2A*) mean methylation and waist circumference, but not with BMI, body weight, fat mass, or depression, as measured by the Beck Depression Inventory Score.[96]

In longitudinal analyses, one study found that following a weight loss intervention program, *hyper*methylation of 6 CpGs in *HTR2A* was associated with a weaker reduction in fat mass (including 5 CpGs with BMI (p-value <0.05)), as well as an attenuated decrease in depressive symptoms (mean-*HTR2A* (p-value=0.003); *HTR2A*:cg24118521 (p-value=0.023)).[96] This would suggest that the reduced expression of *HTR2A* (by nature of methylation blocking transcription) is contributing to persistent levels of obesity and depression. The authors were unable to validate their findings, observing no association between methylation and expression levels, citing reasons that will be expanded on in the discussion section of this review.[96]

Limitations

Use of the cross-sectional study design makes it impossible to determine if methylation levels are a cause or result of obesity/increased BMI. This is a concern that Na et al. acknowledged, along with the speculative nature of what the true causal direction may be. As adipose tissue has known endocrine properties,[103] it is difficult to say with any certainty whether the methylation patterns in cross-sectional obesity studies are a cause or consequence of obesity status. Another study presented in this review, conducted cross-sectional, as well as longitudinal analyses, but could not determine if methylation was a cause of high BMI, or if methylation levels allowed BMI to persist at high levels, as repeated measures of DNA methylation were not taken.

Two of the epigenetic studies focused solely on non-Hispanic White individuals due to availability of samples.[61, 96] This echoes a similar pattern observed in the GxE literature search, and limits the generalizability of any findings to other racial/ethnic groups. The previous studies highlight a need for the field to include populations of different ancestries in order to examine the complex

interactions of psychosocial factors on DNAm and obesity in multiple populations. Additionally, studies with small sample sizes that conduct many tests are hindered by low power and limited in their ability to detect significant effects.

Similar to genome-wide analyses, these epigenetic studies were susceptible to making false discoveries from multiple comparisons of independent hypotheses. However, the Benjamini and Hochberg method of correcting for multiple comparisons[104] was employed in array-based DNA methylation studies to account for this potential limitation.[96] Additionally, using global measures instead of site-specific methylation[99] eliminated many potential tests. Likewise, employing a targeted approach by interrogating CpGs in specific genes based on *a priori* knowledge[61, 96-98] diminished multiple testing concerns. Nevertheless, the possibility of reporting false positives should be taken into account.

2.4 Discussion

Primary results

The overall goal of this study was to determine if psychosocial factors interact with genetic risk to play a role in obesity or weight gain, and in doing so, help bridge the gap in information regarding health disparities still observed today. Our systematic search of electronic articles yielded very few studies that examined interactive effects or epigenetic effects on obesity measures, with only nine studies identified in total that met all eligibility criteria. Among GxE findings, we found multiple examples that indicated a positive interactive effect of psychosocial factors with genetic factors leading to an increase in measures of hip circumference,[92] waist circumference,[93, 94] and BMI.[91, 93] Among methylations studies we found mixed results, where global methylation was not associated with BMI or depression,[99] while site-specific tests found that *hyper*methylation of sites within *HTR2A*, *SLC6A*, and *IL6* was associated with various measures of obesity.[61, 96, 97]

To our knowledge, this is first systematic review to examine the interaction of psychosocial factors and genetic risk for obesity that also examines the role of epigenetic regulation. The epigenetic search revealed little in the way of an ideal study that encompassed the sequential pathway from environment to methylation change to negative health outcome. Rather, each study either assessed the impact of DNA methylation on obesity and psychosocial factors (namely, depression) as separate outcomes,[96, 99] or focused on genes associated with stress-response mechanisms, without directly evaluating measures of psychosocial stress.[61, 97, 98]

The limitation in the number of studies available for review was anticipated as GxE and epigenetic studies are relatively new fields of research. Furthermore, given that that epigenetic remodeling occurs primarily in utero,[105] and is perhaps most fluid during embryogenesis,[106] it is not surprising that there are so few DNA methylation studies examining adults, and then specifically, adulthood obesity. Still, MZ twin studies have shown that independent of inherited genomic sequences, DNA methylation patterns can become altered throughout adulthood,[59] and thus, these studies make good candidates for elucidating the effect of the environment on adverse health outcomes.

As mentioned previously, the desire to discover the "missing heritability" of complex genetic diseases, obesity included, has led to the expansion of gene-environment interaction studies, and especially epigenetic studies. However, it should be noted that epigenetics in particular is likely a better mechanism for determining missing causality of complex disease, than missing heritability.[107] As Slatkin et al. demonstrates, epigenetic marks (e.g. DNA methylation) would have to persist for many generations in order to explain missing heritability that is not uncovered from candidate gene or GWAS,[107] and at this time, that information is not available.

The current review has shown that there are many difficulties that come with trying to elucidate GxE and epigenetic effects not directly inherited from genetics alone. These challenges include poor study design, measurement discrepancies, and limited available data. We found that the potential for false discoveries in GxE and epigenetic studies was a substantial concern of multiple investigators. As mentioned previously, no one single polymorphism can explain a large proportion of observed BMI variance.[35] A method that can address both multiple testing concerns and the need for polygenic risk assessment, is the use of genetic risk scores, which are gaining prominence in the field of genetic epidemiology,[108, 109] including studies of obesity.[38, 39]

Future studies should employ study design and analytical methods that mitigate the possibility of false positives from conducting many statistical tests. An ideal example of both these techniques can be found in the work of Needham et al. who conducted a study on SES trajectories from childhood to adulthood and DNA methylation in a multi-ethnic cohort.[110] The study used a targeted approach by selecting CpGs within specific stress-response and inflammatory pathways. Analytically, the investigators assessed methylation at the gene level, with specific CpG sites acting as "repeated measures" in a multi-level model, which also allowed for intra-gene correlation of DNA methylation to be taken into account.[110]

Studies that are conducted with repeated measures (i.e. longitudinally) to assess temporality as well as the mediatory role of DNA methylation from psychosocial factor to obesity would be ideal. Cao-Lei et al. conducted one such study, but it was not included in the final set of relevant articles, having been excluded for not being in the population of interest (i.e. adults).[111] Evaluating the mediating effect of DNA methylation from prenatal maternal stress to obesity in offspring at age 13.5 years, the researchers identified CpGs using a targeted approach for genes that were in pathways associated with Type 1 and Type 2 diabetes mellitus. The study discovered a slightly protective effect of DNA methylation, meaning it limited the role of objective stress on adverse metabolic outcomes. Future research should employ similar methods to those employed by Cao-Lei et al.,[111] so that DNA methylation can be assessed as a mediator between an environmental stimulus and negative health outcome.

The growing interest in epigenetic regulation is supported by the fact that non-genetic exposures such as diet, smoking and environmental toxins are directly associated with altering DNA methylation levels.[49-51] Increased DNA methylation leads to differential expression of genetic material by blocking transcription[52]; inhibiting transcription prevents a gene from being expressed or "turned on." Additionally, methylation is a recurring process that is believed to be dynamic throughout the life course, with evidence showing that methylation is indeed reversible[53, 54].

Non-representative sample populations

The benefit of limiting a study to a genetically homogenous population is ensuring that any discoveries are truly due to causal polymorphisms and thereby not related to an artifact of genetic ancestry.[112] This is especially true for smaller studies utilizing a candidate gene approach that do not have the benefit of genome-wide data to adjust for ancestral principal components. However, we know that the effects of stressors are experienced and responded to differently among racial/ethnic subgroups and can lead to differing health outcomes.[113] Additionally, obesity genetic risk variants differ by racial/ethnic subgroups[36]. Therefore, future researchers would do well to expand their study populations to non-whites in order t address all potential social and biological mechanisms by which the "Gene by psychosocial interactions" contribute to obesity.

A comment on twin studies

It should be noted that the concept of an environmental impact on genetic predisposition for disease is not by any means novel. Two canonical divisions of epidemiology, genetic and social, are often pitted against each as a representation of the age-old debate of whether poor health outcomes are derived from biological inheritance or adaptations in response to one's environment (i.e. nature vs. nurture). What is ultimately lacking in the current literature are studies that integrate genetic and environmental factors together to explain a joint effect on negative health outcomes (e.g. obesity). Attempts at addressing the nature vs. nurture debate are exemplified in twin studies. While no twin studies made it into the final set of relevant studies, it should be noted that prior research has attempted to assess the impact of environmental factors while also considering the role that genetics plays in the etiology of obesity.

As early as 1986, Stunkard et al. conducted a twin study in order to produce heritability estimates of obesity that could be attributed to genetics and the environment.[75] The study found that 80% of heritability, as determined by concordance rates among monozygotic (MZ) and dizygotic (DZ) twins respectively, was accounted for by genetics. The variable discordance rate in MZ twins was then assumed to represent environmental (non-genetic) factors contributing to obesity. Although, the authors recognized that a potential bias in heritability estimates exists due to non-additive variance, GxE interactions, and unequal shared-environments among MZ and DZ twins. [75]

Many twin studies have sought to eliminate the effect that genes play by selecting only MZ twins. One particular twin study that was retrieved from our initial search showed the effect of psychosocial factors in identical twins discordant for obesity, with obese co-twins having significantly higher levels of two measures of psychological distress.[114] This study was excluded however, because while genetics was taken into account, the monozygotic twin design is effectively eliminates the genetic component in order to isolate the psychosocial component.

In addition to the limitations presented thus far, the current systematic review acknowledges the potential threats to validity, including publication bias, article selection bias, and the use of studies with unrepresentative sample populations.[115] To our knowledge, the current systematic review is in accordance with PRISMA guidelines for reporting scientific studies.[116]

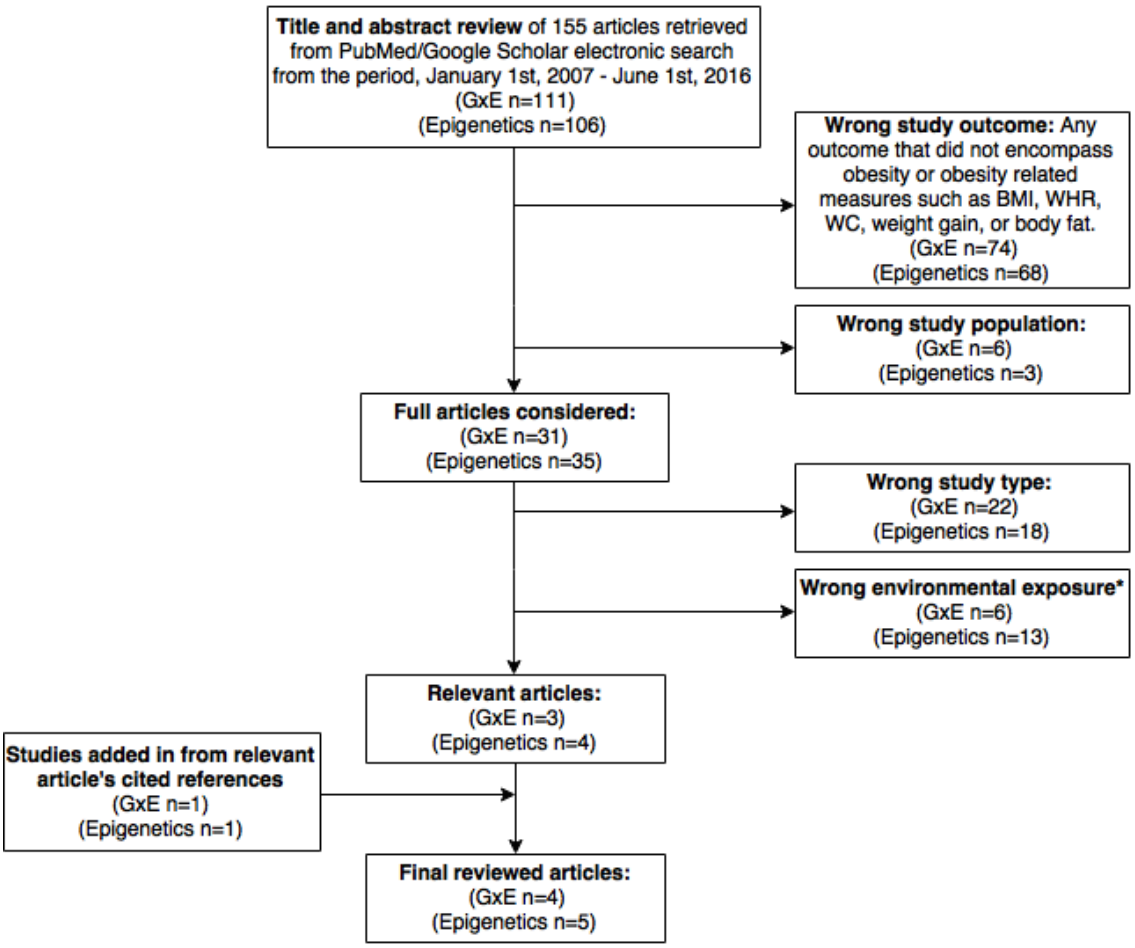
We have attempted to summarize the body of work that has been conducted with respect to the effect of psychosocial factors on obesity, in a way that integrates genetic predisposition or epigenetic regulation. We observed that there are few studies that address the joint effect of genetic and psychosocial contributors to obesity. There were multiple examples where each factor was examined separately, or by studies in monozygotic twin studies that attempt to isolate one effect over another. Also, there are only a few examples of GxE interaction studies or epigenetic studies that show clear temporal ordering from an environmental stimulus to mediating effect on DNA methylation to adverse health outcome. Further, we observed that genome-wide epigenetic and GxE studies are limited due to design complications, or lack of power driven by inadequate sample sizes and multiple testing issues.

GxE interaction and epigenetic studies also have practical applications for points of intervention. The findings lend themselves as diagnostic tools for individuals at higher risk for metabolic conditions. Gene-environment interaction studies are ideal for assessing a point of intervention for obesity reduction, as blocking the environmental trigger of the genetic association (i.e. stress reduction) can still prevent disease, despite a polymorphism lacking functional or biological relevance. Similarly if DNA methylation is being altered by the environment, then the environment can be intervened on to prevent disease when we would otherwise consider the heritable risk factor to be something that was unchangeable.

Ideally, researchers going forward will conduct longitudinal studies using well-defined population-based cohorts, and employ statistical methods that limit the total number of tests performed. As more data become available, replication studies can help confirm already established associations, and expand into previously unavailable populations. Finally, validation of findings by multiple technologies for genetic and epigenetic studies will ensure that results can be interpreted directly and not in a speculative manner. The discovery of new pathways that can increase the risk for obesity allows for the implementation of new techniques for treatment and intervention. A key benefit of establishing gene-environment interactions and epigenetic mechanisms is that it allows for the intervention on a modifiable risk factor, e.g. psychosocial stress, to eliminate the effect of an unchangeable genomic sequence. As obesity remains a pervasive issue in the United States, researchers must examine multiple mechanisms by which interventions can be put into place to reduce the overall public health burden of the disease.

2.5 Tables and Figures

Figure 1. Flowchart of electronic search strategy along with steps included for article exclusion criteria.



Articles that fell into multiple exclusion categories were prioritized in the following order: wrong study outcome, wrong study population, wrong study type, or wrong environmental exposure. Thus, no article has been counted twice in the flowchart.

* Examples of "wrong environmental exposure" include dietary measures or behavioral factors such as physical activity, and smoking.

Table 1: Significant results from gene-psychosocial factor interactions with various measures of obesity.

Author/Year	No. of Subjects	Study pop.	Psychosocial factor	Obesity measure	SNP	CHR	Gene	Effect estimate	GxE P-value	Replicated
Singh, 2014	2,460	White Men and Women	Chronic burden of stress	Hip Circ. [a]	rs4704963	5	*EBF1*	*B* = 2.98	7.14E-09	N/A
					rs17056278	5	*EBF1*	Not Reported	7.14E-09	Yes
					rs17056298	5	*EBF1*	Not Reported	1.30E-08	Yes
					rs10077799	5	*EBF1*	Not Reported	1.71E-08	N/A
					rs17056318	5	*EBF1*	Not Reported	2.33E-08	Yes
Kring, 2010	341	White Women	Chronically stressed caregivers	Waist Circ.	rs439401	19	*APOE*	*F* = 3.10 (additive model)	0.047	Yes
								F = 5.00 (dominant model)	0.03	Yes
Kring, 2011	1,184	White Men	Cumulative stress score	Waist Circ.	rs439401	19	*APOE*	OR = 1.09[c] 95%CI = (1.02-1.17)	0.01[c]	N/A
				BMI				OR = 1.09[c] 95%CI = (1.01-1.17	0.02[c]	N/A
Marmorstein, 2010	903	Men	Receipt of public assistance	BMI	Short allele	X	*MAOA*	*B* = 0.42 (SE=0.14)	<0.01[b]	N/A

[a] Primary results for BMI and waist circumference GxE interactions reached the p-value <0.005 for all 5 SNPs, but not genome-wide significance (p<5.0E-08).

[b] Result was only significant in a 3-way interaction with Age. The 2-way interaction with the psychosocial and genetic factor did not reach statistical significance.

[c] This is not the GxE interaction odds ratio and p-value, but the significant stratified result within stressed individuals. The odds ratio represents the odds of being homozygous for the minor allele in *APOE*:rs439401, given high levels of waist circumference/BMI and chronic stress.

Supplementary Table S1. Detailed search terms for GxE and Epigenetic electronic articles

Gene by environments interaction studies (GxE) (N=111 results)	**PubMed search:** ("gene-environment interaction"[All Fields] OR "gene-stress"[All Fields] OR "genetic interaction"[All Fields] OR "gene variants"[All Fields]) AND ("SES"[All Fields] OR "race"[All Fields] OR "education"[All Fields] OR "income"[All Fields] OR "depression"[All Fields] OR "discrimination"[All Fields] OR "hassle"[All Fields] OR "stress"[All Fields]) AND ("obesity/genetics"[All Fields] OR "obesity"[All Fields] OR "obese"[All Fields] OR "obesity/epidemiology"[All Fields] OR "body mass index"[All Fields] OR "waist hip ratio"[All Fields] OR "waist circumference"[All Fields] OR "body fat"[All Fields] OR "adiposity"[All Fields] OR "BMI"[All Fields] OR "WHR"[All Fields]) AND (("2007/01/01"[PDAT] : "2016/06/01"[PDAT]) AND "humans"[MeSH Terms] AND English[lang]) **Google Scholar:** gene-stress interaction bmi obese obesity
Epigenetic studies (N=106 results)	**PubMed search:** ("stressor"[All Fields] OR "HPA"[All Fields] OR "stress response"[All Fields] OR "serotonin"[All Fields] OR "stress"[All Fields] OR "psychosocial"[All Fields] OR "depression"[All Fields] OR "hassle"[All Fields] OR "discrimination"[All Fields]) AND ("epigenetic"[All Fields] OR "methylation"[All Fields]) AND ("obesity/genetics"[All Fields] OR "obesity"[All Fields] OR "obesity/epidemiology"[All Fields] OR "body mass index"[All Fields] OR "waist hip ratio"[All Fields] OR "waist circumference"[All Fields]) AND (("2007/01/01"[PDAT] : "2016/06/01"[PDAT]) AND "humans"[MeSH Terms] AND English[lang])

Search terms for electronic retrieval of articles are presented here.

Chapter 3

Genetic modification of the effect of self-reported psychosocial factors on obesity

3.1 Introduction

Obesity continues to be a public health concern, with the prevalence in the US remaining at an all-time high (35.5% among adult men and 35.8% among adult women).[3] There are significant disparities in obesity, with racial/ethnic minorities more likely to be obese compared to no-Hispanic White individuals. Higher age-adjusted prevalences are observed among non-Hispanic black women (58.5%) and Hispanic populations (39.1%), compared to non-Hispanic white men and women (34.3%).[3] Obesity is a concern as it is a risk factor for various chronic conditions, such as cardiovascular disease, as well as an independent predictor of overall mortality.[5] The current situation with regard to obesity shows that there is a sustained need for discovering all possible causes contributing to the onset or persistence of the condition.

Over the past 30 years, many genetic epidemiologic studies, including candidate gene studies, genome linkage, fine mapping, and genome-wide association studies (GWAS) have identified numerous single nucleotide polymorphisms (SNPs) associated with BMI. GWAS findings have been encouraging, yet limited in identifying very strong predictors of obesity. For example, variation in the strongest genetic predictor, fat mass and obesity gene (*FTO*), confers at best an increase in BMI of 0.39 for each copy of the risk allele (explaining only 0.34% of total genetic variance).[35] Additionally, replication of *FTO* has yielded mixed results in various racial and ethnic populations, with some showing weak or no association,[117] and others showing definitive associations in populations of Chinese[118] or African origin.[119]

Weak genetic findings highlight the limited ability of individual loci to predict common outcomes. In order to synthesize genetic discoveries in a way that incorporates a broader polygenic contribution to obesity, Belsky and colleagues developed a method for examining total genetic risk using an obesity genetic risk score (oGRS) with promising results.[38] Polygenic risk scores allow for the incorporation of the most recently discovered SNPs available for research. The most recent meta-analysis for obesity risk genes was conducted in 2014 by Locke et al., which identified 97 genetic predictors of high BMI (56 novel loci), from studies that included 339,224 individuals.[36] These new methods for assessing total genetic risk are timely and warranted and can provide added insight for obesity risk to supplement the non-genetic risk factors that have predominantly led obesity research.

Although much of the literature on non-genetic determinants of obesity has focused on health behaviors (e.g. altered diet or physical activity), there is a greater recognition of the potential role of psychosocial factors in the etiology of the condition. Two of the largest nationally representative, population-based longitudinal studies from the Coronary Artery Risk Development in Young Adults (CARDIA) cohort and the Midlife in the United States (MIDUS) study have previously

examined how social and psychosocial factors, can lead to weight gain and increase the likelihood of obesity.[20, 21] These studies found that psychosocial factors, such as difficulty paying bills and job-related demands, increased the likelihood of obesity.[20] An additional psychosocial factor, depression,[69, 70] has also been strongly implicated in weight and obesity.

Current arguments emphasize that a key mechanism by which stress may lead to obesity is through altering behavioral patterns related to diet or total caloric intake.[28] Alternative biological pathways have been explored in animal models where chronic exposure to stressors led to excess levels of glucocorticoids in the blood, which can influence fat storage.[74] Though similar mechanisms have not been tested thoroughly in humans, biological plausibility exists for the role of psychosocial factors on obesity onset, not directly attributed to behavioral patterns. Given the multifactorial nature of obesity, new mechanisms warrant exploration. As biological pathways are interconnected with genetic factors, a unique model system that can link psychosocial factors with obesity is the gene-environment interaction.

Gene by environment interaction studies (GxE) suggest that a genetic predisposition for obesity works synergistically or antagonistically with non-genetic factors to modify the risk of obesity. GxE studies have the potential to build on weak findings in previous studies looking at genetics or psychosocial factors alone, by revealing cases that occur only in the presence of both exposures simultaneously. While the current literature primarily focuses on the interactions of dietary factors or socioeconomic status,[40] there are several examples of how psychosocial stress/stressors interact with genetic risk variants (e.g. *APOE[120]*, *MAOA[120]*, and *EBF1[120]*) to increase the likelihood of obesity or weight gain. GxE findings thus far have been either weak, or restricted to homogenous populations of European ancestry because of sample availability or to avoid confounding by population stratification. Concerns regarding population stratification can be mitigated by the use of oGRS scores, which are effective in multi-ethnic populations[120] and can maximize sample size to detect main effects of psychosocial factors.

The overall goal of the current study was to examine whether psychosocial factors and an obesity genetic risk score act independently or synergistically to affect the risk of obesity. The specific research questions were: 1) is there an independent association between psychosocial factors and genetic factors in relation to incident obesity, 2) is the association between psychosocial factors and obesity modified by genetic factors, and 3) do any significant associations vary by race/ethnicity. We hypothesized that individuals with higher levels of baseline psychosocial factors will have a higher risk of obesity over the study follow-up period. Additionally, associations between psychosocial factors and obesity will be modified by an obesity genetic risk score, such that associations will be the strongest among those with high genetic risk and high psychosocial stress/depression.

3.2 Methods

Subjects

The Multi-Ethnic Study of Atherosclerosis (MESA) is a prospective cohort study of 6,814 individuals recruited between July 2000 and August 2002 at six field centers around the United States: Columbia University, New York; Johns Hopkins University, Baltimore; Northwestern University, Chicago; UCLA, Los Angeles; University of Minnesota, Twin Cities; Wake Forest University, Winston Salem.[121] Participants were from diverse racial and ethnic backgrounds (non-Hispanic White, African-American, Hispanic, Asian-American) and free from clinical CVD at baseline.[121] Data collection included study questionnaire and clinical measurements taken at baseline and four follow-up exams (approximately every two years through January 2012). Detailed methods and recruitment procedures have been described elsewhere.[122]

Of all MESA participants recruited at baseline, 6,429 consented to providing DNA samples for genotyping. A total of 68 subjects were excluded for poor genotyping quality (call rate <95%). Subjects were then removed for missing data on psychosocial factors (total N=100; CSS(N=27), EHS(N=37), CES-D(N=5), missing two exposures (N=6), missing all three exposures (N=25)). No participants were missing outcome data, but an additional 469 were removed due to missing information on one or more of the key study covariates. Implementing these exclusion criteria yielded a final analytic sample of 5,792. Institutional Review Board (IRB) approval was obtained at each study site, with written informed consent given by all MESA study participants.

Outcome assessment

The primary outcome was obesity. Anthropometric measures such as height and weight were taken at each clinical examination and BMI was calculated as weight/height (kg/m^2). Obesity was defined as having a BMI greater than or equal to 30 kg/m^2. Prevalent cases of obesity at baseline were not excluded in primary analyses, but accounted for by analytical procedures described below.

Psychosocial factors

Three psychosocial variables were assessed via study questionnaire at baseline: chronic psychological stress (the chronic burden scale developed for the Healthy Women Study);[123] everyday hassles, developed from perceptions of discrimination and unfair treatment;[124, 125] and depression.[126]

The chronic burden of stress scale (CSS) was measured as the sum of the number of times a subject answered yes to the following ongoing stressful problems for the past six months (own health, close person health, job, financial, or relationship). [86] Participants were given an option to

classify the level of stress for each domain on a scale from 1-3; 1 being "not very stressful," 3 being "very stressful." A chronic burden score was analyzed by summing the number of domains in which moderate-to-very stressful was identified. Possible scores of CSS ranged from 0 to 5 and were analyzed categorically as high or low chronic stress (2+ vs. <2; referent), based on previous methods.[127]

The Everyday Hassles Scale (EHS) is a validated nine-item scale that measures day-to-day incidents of unfair treatment, based on the frequency of encounters in which someone perceived that he/she was treated unfairly.[128] Examples of perceived unfairness included being called names, implied to be a liar or unintelligent, etc. Each item ranged from 1-6 (1=never to 6=almost every day), with the final score calculated as a sum of all nine items. A higher score on the scale represented a higher frequency of hassles each day and, therefore, more stressors. The scale was examined based on previous literature as a dichotomous variable: high discrimination (19+) versus moderate to no discrimination (<19; referent).[87] Internal reliability/consistency estimates for the EHS measure were good in the final set of complete cases used in all analyses (Cronbach's alpha= 0.88, 95% CI (0.87, 0.89)).

We additionally assessed the possibility of there being an increased effect of having high psychosocial stress in both chronic burden and everyday hassles that might not have been captured by assessing each variable individually. This new composite variable was called "high stress-both" (or HS-both) and was examined as a dichotomous variable.

Depression was analyzed as a dichotomous variable (Yes/No) based on depressive symptoms measured using the Center for Epidemiological Studies-Depression (CES-D) inventory,[129] which ranged in MESA from 0 to 53. An affirmative for depression entails scoring at least 16 on the CES-D scale, based on previous methods.[130] The CES-D has been a reliable measure for depression in multiple ethnic groups, including European-Americans, African-Americans, Mexican-Americans, and Chinese populations.[129, 131, 132] This measure has also been previously associated with waist circumference in the MESA population.[133] Internal reliability/consistency estimates for the CES-D were reasonable, but not ideal, based on certain accepted standards[134] (Cronbach's alpha=0.68, 95% CI (0.67,0.70)).

Genotyping and imputation QC

All MESA participants were genotyped using the Affymetrix Human SNP array 6.0 (Affymetrix Inc., Santa Clara, CA), which provided information on more than 900,000 informative SNPs. Additional genotypes were imputed to the 1000 genomes Phase I integrated variant set for all ancestries by each ethnic group in MESA separately.[135] Imputation was performed using IMPUTE v2.2.2.[136] Multiple quality control measures were employed to ensure SNP and sample quality, as described previously.[137] Briefly, prior to imputation, SNP exclusion criteria included the removal of monomorphic SNPs, heterozygosity greater than 53%, or a missing rate greater than 5% across all samples. Samples with call rates less than 95%, duplicate samples and sex mismatches were also removed. Analyses were only performed on the combination of genotyped and imputed SNPs together. Genetic risk scores are less sensitive to SNPs with low minor allele frequency (MAF), and the more SNPs that are included in the risk score, the more

likely the score will approach a normal distribution.[38] Thus, no SNPs were filtered out due to low MAF. All SNPs had an imputation quality score greater than 0.3.

As there is uncertainty in the exact call of imputed genotypes, a "dosage" level was calculated when necessary. Imputed SNPs are given a dosage score based on the probability of having a certain number of copies at each variant, as opposed to an outright number of 0, 1 or 2 alleles and these analytical methods have proven effective when accounting for the uncertainty of imputed genetic data.[138] For example, if an individual is given a 0.30 probability of having one copy and 0.70 probability of having two copies of the risk allele at an imputed SNP, rather than defer to the allele count that has a higher probability, we calculated a total risk dosage of the variant by summing up the weighted allele-specific probabilities (i.e 0.30*1 + 0.70 *2).

Obesity genetic risk score

Of the 97 BMI-associated SNPs that were established from prior GWAS and a comprehensive meta-analysis conducted by Locke et al.,[36] we selected a total of 93 available QC'd SNPs to calculate a weighted obesity genetic risk score (oGRS) for all study participants, based on established validated methods developed by Belsky and colleagues.[38] The oGRS is a weighted measure that derives weights from high BMI susceptibility loci beta values determined from the largest and most recent meta-analysis on obesity genes to date.[36] At each locus, the number of risk alleles is multiplied by the weights for said variant. The final score was calculated by summing weighted values across all 93 loci. A higher weighted score translates to greater genetic susceptibility. The current study used weights derived from the Locke et al. meta-analysis conducted for "all ancestries," which encompassed people who self-identified as European-American, African-American, or Chinese-American. The weights from all ancestries created the score used for the main oGRS variable and will be hereafter referred to as the oGRS-meta.

The benefit of using this weighted risk score instead of individual variants is that the oGRS can be used for trans-ethnic replication, even if the risk variants were established in populations of European ancestry.[39] Testing a three-way interaction between each psychosocial factor by continuous oGRS-meta by race revealed no significant interaction, and thus race-stratified analyses were not considered as part of the main findings (supplemental table S1). However, due to lingering concerns of heterogeneity for oGRS-meta and obesity within racial/ethnic sub-populations, race-stratified supplementary analyses were performed using racial/ethnic specific risk scores for subgroups of MESA's European-Americans, Chinese-Americans and African-Americans, employed in multi-ethnic populations by Domingue and colleagues.[39] These sub-scores were comprised of the 93 European-by-descent derived SNPs for MESA's non-Hispanic White participants (oGRS-EU), 75 SNPs for Asian-American participants (oGRS-CHN), and 72 SNPs for African-American participants (oGRS-AFA), based on directional consistency with European-only analyses conducted by Locke et al.[36] The weights used for sub-scores are based on beta estimates from populations of European, African, and Chinese ancestry respectively.[36] (see supplementary table S2 for list of genetic predictors and weights that comprise each subset of oGRS). All oGRS scores were made into a dichotomous variable by a median split for ease of interpretation using hazard ratios.

It should be noted that each racial/ethnic specific sub-score was weighted by the beta values derived by Locke et al. in racial/ethnic subpopulations separately. Therefore, the weights applied to each supplementary oGRS do not match the oGRS-meta weights and cannot be compared across groups. To our knowledge, there is no consensus in the current literature regarding an obesity genetic risk score that is specific to Hispanic adults, and therefore, no Hispanic-specific risk score was evaluated in the current study. However, in race-stratified analyses the oGRS-meta was used in the Hispanic-American sub-population.

Additional study covariates:

Initial models (Model 1) accounted for a "basic" set of individual-level socio-demographic confounders, including age, sex, and self-reported race/ethnicity. Model 2 additionally adjusted for the covariates of smoking status (ever vs. never smoked), education level (see table 1 for details), and per capita adjusted household income per $10,000 or (continuous income/# people supported)/10,000); hereafter this will be referred to as the "fully adjusted" set of covariates. Both Models 1 and 2 accounted for population stratification in order to determine direct genetic contributions from obesity genes, independent of variation due to genetic origin. This was done using three principal components that explained at least 95% of total observed genetic variation in the population and was calculated from methods described previously.[139, 140] Briefly, principal components in MESA were derived from genotyped SNPs after pruning SNPs in linkage disequilibrium based on an R^2>0.2. The initial set of SNPs was determined within self-reported Hispanic-Americans and further refined in the remaining ethnic groups from African-Americans, to European-Americans, to Chinese-Americans, yielding a final set of 76,804 SNPs for PC analysis.

Additionally, physical activity was controlled for in the fully adjusted models as it has a direct impact on energy expenditure and is a known correlate of stress.[141] Physical activity was measured as the number of active minutes in a typical week using a detailed questionnaire adapted from the Cross-Cultural Activity Participation Study and validated for a multiethnic cohort.[142] The influence on weight gain caused by energy intake is directly caused by one's nutritional diet. Prior research has indicated a potential interactive effect between genes and diet on BMI.[143] However, a recent study of 68,317 individuals of European ancestry showed no evidence of an interaction between diet score and obesity genetic risk score on BMI, although it did show a nominally significant effect with waist-hip ratio.[144] Given the potential relationship between stress and eating patterns, fully adjusted models also accounted for caloric intake, measured in kcal/day.

Statistical analysis

In descriptive analyses we examined the univariate distribution of all study variables and bivariate associations between the study outcome and all study covariates (Table 1). We also calculated the effect of oGRS on BMI at baseline, adjusting for principal components, for all subjects and subsequently stratified by race to determine if the oGRS was performing as expected. This was

assessed continuously and by oGRS quartiles to show the dose-response effect on BMI levels by increasing genetic risk categories.

The prevalence of obesity in the overall MESA cohort was 31.8% at baseline, with 517 incident cases (11.9% cumulative incidence) by the end of exam 5. The overall rate of incident obesity in MESA by exam 5 was 17/1000 person-years (Table 2). The potential for the relatively low incidence of obesity to inhibit statistical power led us to conduct our main analysis using both prevalent and incident cases. We employed parametric, interval censored survival analysis to test main hypotheses, modeling associations between psychosocial factors, genetic risk scores, and their interactive effect on obesity. We used age as the time scale in order to incorporate prevalent obesity cases at baseline. This method, detailed previously,[145, 146] accounts for the subjects' varying age-based entry points into the study and allowed us to maximize the amount of information our data can provide on the relationship between our exposures of interest and obesity.

Prevalent cases were treated as left censored with baseline age as the upper interval boundary, and incident cases were censored within the interval defined by age at last obesity-free exam point to age at obesity onset. All remaining participants free of obesity or lost to follow-up were considered right censored at the final exam in which they participated. Analysis was conducted using a Weibull distribution for the hazard based on graphical evidence (plotting the log(-log(Survival)) and model fit (AIC values).[147] The log(-log(survival)) plots showed no clear violation in the proportional hazards assumption for binary oGRS-meta or any psychosocial factor variable. This indicated that the Weibull model was, in fact, appropriate.

Psychosocial factors (CSS, EHS, HS-both, and Depression) were the "environment" exposures of interest, and gene-environment interaction (GxE) was assessed with obesity on both the multiplicative and additive scales. Each model included main effects for both genetic and psychosocial factors, along with a cross product term for oGRS and psychosocial factor to determine interaction on a multiplicative scale. Multiplicative interactions were assessed using likelihood ratio tests, with Wald statistic p-values <0.05 considered statistically significant, after correcting for multiple comparisons, using a non-conservative false discovery rate.[104] Multiple testing was performed for main effect and interaction p-values for models 1 and 2 across all exposure variables in primary analyses (N=24 tests). In supplementary analyses stratified by race/ethnicity, results were corrected for a total of 60 tests.

We used hazard ratios to calculate the Relative Excess Risk due to Interaction (RERI), as described by Knol, et al,[148] which determined if there was an interactive effect on the additive scale. Confidence intervals for the RERI were obtained from the 2.5% and 97.5% quantiles of 2000 bootstrapped samples. The current study had power to detect an interactive effect of 1.8 or greater, assuming certain dependencies of main genetic and environmental effects, based on methods detailed by VanderWeele et al.[149] These dependencies included a dichotomous designation for oGRS, an assumed marginal prevalence of 25% for both E and G exposures, and main effect sizes of 1.5. All regression models were conducted in R using the *Survival* package.[150] Effect estimates were reported as hazard ratios, with confidence intervals calculated by the delta method[151] using the *SurvRegCensCov* package.[152]

3.3 Results

Descriptive statistics

Demographic information on all MESA subjects can be found in Table 1. There was a total of 5,792 subjects available for analyses, with roughly 30% already obese at the initiation of the study. Of the 4,003 non-obese subjects at baseline, 486 (12%) went on to become obese at some point, with the total cohort contributing 28,380 person-years of follow-up. Obese cases contributed, on average, 4.03 person-years of follow-up time (range=1-11), with non-obese subjects contributing roughly 7.47 person-years (range=0-11). An unadjusted incidence rate of 16 obesity cases per 1000 person-years was observed in the current study.

The overall MESA population was comprised of older adults at baseline (mean=62.0, sd=10.2), with non-obese subjects slightly older than prevalently obese subjects (62.6 vs 61.0; p-value <0.001). The cohort was a majority female (52.0%) and the breakdown by self-reported racial/ethnic categories showed that most participants identified as non-Hispanic White (40.4%), with 13.2% Chinese-American, 23.1% African-American, and 23.2% Hispanic-American. Race/ethnicity was significantly associated with prevalent obesity (p-value <0.001) and incident obesity (p-value <0.001). The cohort was relatively well educated, with roughly 92% of the population having received at least a high school diploma and 18.7% having received post-college training. Education showed a significant trend with prevalent and incident obesity, with higher education associated with a lower likelihood of obesity.

The weighted obesity genetic risk score was normally distributed in the population (mean=2.17, range=1.61-2.76). We assessed the unadjusted effects of oGRS-meta with prevalent obesity both continuously and in quartiles of genetic risk. Statistically higher cell counts were observed in non-obese subjects with low risk and prevalently obese subjects with high risk (p-value <0.001). This trend was evident in incident cases as well (p-value=0.008).

Each of the three psychosocial factors examined (CSS, EHS, Depression) was associated with prevalent and incident obesity in bivariate analyses. An apparent dose-response relationship by categories of increasing CSS was observed. Among prevalent cases of obesity, the frequencies of low, moderate, and high levels of chronic stress were 8.6%, 9.6% and 12.7%, respectively. This effect appeared to become attenuated among incident cases (low=3.8%, moderate=3.7%, high=4.2%). Every covariate shown in Table 1 was significantly associated with prevalent obesity at the alpha 0.05 level, in unadjusted bivariate analyses, except for smoking status. Other than income, sex, physical activity, and depression, all other covariates were also associated with incident obesity.

We assessed the impact of the oGRS-meta with continuous BMI at baseline using simple linear regression, to ensure that the score was performing as expected. We found that oGRS-meta was significantly associated with BMI in the overall cohort (beta=0.56, 95%CI (0.43, 0.70)) and within non-Hispanic White, Chinese, and Hispanic racial/ethnic subsets, but not in African-Americans (supplementary table S3). This was done primarily to ensure that a risk score using meta-analysis weights from European, African, and Chinese ancestries could apply to the MESA population,

which included Hispanic subjects as well. These results also revealed that the highest quartile of genetic risk was significantly associated with increased BMI in MESA, regardless of race/ethnicity. The oGRS-meta performed better than the race/ethnicity specific sub-scores and weights for baseline BMI, indicating that oGRS-meta is perhaps a better measure for determining genetic risk for obesity in a multi-ethnic population.

Survival analysis

In Model 1, adjusting for basic covariates (age, sex, race/ethnicity, and principal components), there were statistically significant associations between each of the main exposures CSS/oGRS-meta and obesity; oGRS-high vs. low (HR 1.23, 95%CI (1.10, 1.37); p-value=0.008), CSS-high vs. low (HR 1.42, 95%CI (1.25, 1.61); p-value=0.010) (Table 2, Model 1). No significant multiplicative or additive interaction was observed. Main effects were attenuated but remained statistically significant upon additional adjustment for smoking, education, income, exercise, and caloric intake. (Table 2, Model 2). The main effect for oGRS-high in fully adjusted models was a hazard ratio of 1.20, (95%CI (1.08,1.34); p-value=0.004), and for CSS-high, a HR of 1.18 (95%CI (1.04,1.35); p-value=0.026). Similar to Model 1, no interaction was detected on the multiplicative or additive scale in the fully adjusted model.

No model comparisons for EHS, HS-both, or depression yielded a statistically significant association with obesity. This was true when adjusting for the basic or full set of covariates. The main effect for oGRS remained statistically significant for both adjustment sets, regardless of which psychosocial variable was included in the model (Model 1: p-value <0.05; Model 2: p-value <0.05). There was no evidence of interaction on either the multiplicative or additive scale. All models were subsequently examined comparing those in the highest quartile of genetic risk versus the lower three quartiles. While main effects for oGRS-meta were consistently stronger across each comparison and all models failed to show a significant interaction on either scale (data not shown).

Supplementary analyses

We tested the potential for heterogeneity of the main effects variables and obesity by sex or race. In tests for three-way interactions, we observed no statistically significant effect, and thus stratifying by these third variables was not necessary. However, as mentioned in the methods section, due to the nature of how oGRS SNPs are discovered in populations of certain ancestries and weighted according to those populations as well, we felt it necessary to supplement our main findings with race-stratified results. For each racial/ethnic subgroup in MESA, we analyzed the effects of each psychosocial variable and oGRS-meta, as well as the race/ethnicity weighted oGRS sub-scores defined in the methods section. Results were only presented for the fully adjusted models (Table 3).

In African-American subjects, the main effects of oGRS-meta and oGRS-AFA were not significant in any model, regardless of which psychosocial factor was being compared. However, we observed

a significant negative interaction on the additive scale with oGRS-AFA and CSS (RERI= -0.27, 95%CI (-0.55, -0.01). Also, we observed a sub-additive interaction for oGRS-meta and EHS (RERI= -0.31, 95%CI (-0.55, -0.06), but not with oGRS-AFA. This additive interaction persisted when looking at the composite measure of CSS/EHS and oGRS-meta (RERI= -0.42, 95%CI (-0.75, -0.10). Evaluating depression and oGRS-AFA, we observed negative interactive effects additive scale (RERI= -0.58, 95%CI (-1.03, -0.22).

Hispanic subjects did not have a Hispanic-specific oGRS sub-score to evaluate, so we elected to use the oGRS-meta instead. There appeared to be a significant interaction with EHS on the multiplicative scale (HR=1.76, 95 % CI (1.12, 2.78)). However, after correcting for multiple comparisons, the multiplicative effect was no longer significant at the 0.05 level (adjusted p-value=0.08). A positive additive interaction with EHS was observed (RERI= 0.56, 95%CI (0.24, 0.90)), despite null main effects in the fully adjusted models. This interactive effect was carried through to the measure looking at high stress in both CSS/EHS as well (HS-both RERI= 0.42, 95%CI (0.03,0.80)). Additive effects trended more strongly in the positive direction when comparing those in the highest quartile of genetic risk with those in the bottom three quartiles (data not shown). No other comparisons in Hispanic Americans reached statistical significance.

For non-Hispanic White participants, significant main effects for oGRS-meta and oGRS-EU with obesity were detected for all comparisons, regardless of which psychosocial factor was being examined. No significant interactions were observed on either the multiplicative or additive scale. In the Asian-American subset, sample sizes were too small to draw meaningful conclusions. All models had difficulty with convergence, likely due to the low counts of obesity in this population.

3.4 Discussion

The current study assessed whether genetic and psychosocial factors act independently or synergistically to increase the risk of obesity. In fully adjusted models, there was an 18% to 20% increased risk of obesity in those with high genetic risk, with an 18% increased risk of obesity in those with high levels of chronic burden of stress (CSS). The main effect for EHS in the model adjusting for a basic set of covariates did not reach statistical significance, nor did a combined measure of CSS and EHS looking at those with high levels of both exposures. Therefore, it would seem that CSS is capturing a unique effect of stress on obesity in the MESA cohort. There was no evidence to support an independent effect of depression on obesity risk. Regardless of which psychosocial factor was being examined, there was no evidence of multiplicative or additive interaction in any model comparison for the overall cohort.

Only the effects of oGRS-meta and CSS were robust to model specifications. The oGRS-meta finding provides support for the use of polygenic risk scores in determining obesity risk from a large set of SNPs previously associated with increased BMI. While the CSS was designed to measure more objective instances of stressors that occur over a long period of time (e.g. job-related demands, difficulty paying bills, or social relationship demands), the EHS measure incorporates topics of *perceived* hassles or *perceived* unfair treatment.[153] Researchers have long tried to determine the best measures for quantifying stress, and the debate regarding objective versus subjective scales has persisted for years. Prior research has shown how perceived stress can

correlate with biomarkers representative of the stress response and obesity[20, 154] or continuous weight gain.[24] An important omission from these studies was the contribution of genetic risk factors, and future studies should incorporate some measure of genetic risk, as the current study has demonstrated.

Showing how psychosocial variables operate within strata of genetic risk (supplementary table S4), it is apparent that regardless of genetic risk level, a high level of chronic burden of stress is independently associated with obesity. We did not observe the same effect when looking at other psychosocial factors. For example, comparing the doubly-exposed participants with high genetic risk and high everyday hassles to subjects with neither exposure, we saw a significant relative hazard of 1.29 (95%CI 1.09, 1.45). The genetic risk predictor is likely driving this effect, given that we observe no effect of everyday hassles within specific stratum of genetic risk (oGRS-meta$_{low}$: HR= 1.04, 95%CI (0.90, 1.21); oGRS-meta$_{high}$: HR= 1.08, 95%CI (0.93, 1.25)). Thus, the oGRS-meta and CSS are stronger risk factors for obesity than everyday hassles. It is likely that the objective factors measured by the chronic burden of stress scale might be simply stronger predictors of the biological response that leads to obesity than the everyday hassles scale, when also adjusting for genetic risk.

To our knowledge, this is one of the first studies to examine gene by environment interactions of comprehensive obesity genetic risk using the validated oGRS and detailed measures of psychosocial factors in a large multi-ethnic dataset. Previous studies have focused on individual candidate genes. In a recent study using a multi-ethnic population, Singh et al. found that SNPs within the Early B-cell Factor 1 gene (*EBF1*) interacted with the chronic burden of stress score for hip circumference (p-value <9.46E-08), but only in White participants.[92] Further, a case-control study conducted by Kring et al. showed that SNPs within the *APOE* gene (rs439401) interacted with caregiver stress, leading to higher waist circumference.[93] This effect was replicated in an independent dataset using a self-reported seven-item chronic stress score.[93] The current study adds to this literature by providing evidence of interaction on the additive scale in race/ethnicity specific sub-sets.

Additive interaction is commonly assessed to approximate what is commonly referred to as "biologic interaction."[148] This occurs when there is a subset of the population that only develop the outcome when both exposures are present, potentially interacting biologically to cause disease that would have otherwise not occurred.[148] That form of synergy would support the diathesis stress model and explain why some people with a genetic predisposition may not develop the disease, when the second exposure is not present.[155, 156] This type of interaction would support the hypothesis that weak effects for obesity by either genetic risk or psychosocial measures alone were due to the fact that some subset of the population would be observed as obese only if both exposures were present together. This was the case for Hispanic-Americans in MESA with respect to the everyday hassles scale.

While no evidence of multiplicative interaction was present in the overall cohort or race/ethnic specific subsets, it remains the case that oGRS by psychosocial stress interact on at least one level. Multiplicative interaction is often easiest to detect analytically, where we can use cross-product terms to extract a p-value directly from the model, giving us a commonly accepted measure of statistical significance.[157] Although there isn't a consensus that interaction on the multiplicative

scale has any clear advantage over the additive scale, except to say that the two are different, and depending on the data structure, can reveal instances of interaction on one scale when there are null or opposite effects on the other.[157] Therefore, evidence of interaction in Hispanic Americans should not be underestimated, and the significance of this finding will be explored in the future.

We did not find an interaction between depression and genetic risk, and depression was not independently associated with obesity after adjustment for study covariates. Three recent studies have examined the interaction and/or effect modification of genetic risk scores by depression for obesity. The studies were conducted using a weighted 32 SNP score in a population of European ancestry,[158] an un-weighted 21 SNP score in a multiethnic population,[159] and five polygenic risk scores of varying SNP count in a European population.[160] Each study revealed significant, albeit modest, interactive effects between obesity genetic risk and depression on BMI/obesity. Each of these studies, however, used the DSM-IV designation of Major Depressive Disorder (MDD) for defining depression.

The lack of a significant main effect in the current study may not be attributed to depression itself but rather the use of the CES-D scale, as opposed to clinically based definitions of depressive disorders. Multiple forms of the scale are employed in current research, depending on the investigator's *a priori* decision to use the entire scale or a sub-scale.[161, 162] A 2013 review in PLOS One highlighted numerous concerns that certain components of the CES-D scale for measuring depression might be assessing other constructs, such as social anxiety disorders; Furthermore, particular components are sensitive to sex and cultural biases.[163] Additionally, the CES-D internal reliability/consistency estimate in the current study was not ideal, indicating that the CES-D may not be approximating a single latent construct. Altogether, this indicates that the CES-D score may be inconsistent for assessing clinical depression in certain populations.

There also appears to be a discrepancy regarding the effect of depression on obesity, depending on how obesity is measured. Our findings are consistent with those of a prior study in the MESA cohort that found negligible effects of the CES-D on continuous BMI.[133] This previous study did, however, establish an association between depression and waist circumference, with the authors positing that depression was perhaps more indicative of visceral adiposity than the general adiposity that BMI measures. Thus, multiple issues of exposure and outcome measurement cannot be disregarded for why an effect may not have been observed here.

The strength of the main effect for oGRS-meta in the overall cohort highlights the utility of using composite risk scores for assessing obesity risk, as opposed to the minimal effects detected when using individual candidate genes. Very few studies have assessed obesity genetic risk in non-European ancestry populations; large scale GWAS SNPs are frequently discovered in subjects with European ancestry. This is primarily due to methodological constraints of the need for large sample sizes in genetically homogenous populations, which maximize statistical power for SNPs to reach significance at the genome wide level.[72] Additionally, early examples of the oGRS score showed weaker findings in non-European populations,[38] making it an unattractive measure until more SNPs could be discovered in subjects of varying ancestries. As the oGRS can now be comprised of 97 SNPs discovered from recent meta-analyses, and has been previously shown to be effective in trans-ethnic populations,[36, 39] the current study was able to successfully examine

comprehensive genetic risk in multiple racial/ethnic populations. We were able to show that an oGRS using weights from the Locke et al. meta-analysis of "all-ancestries"[36] was associated with higher BMI levels at baseline in the overall MESA cohort, as well as in subsets stratified by self-reported race (except for African-Americans), adjusting for population stratification, age, and sex.

Despite the strength of the oGRS using meta-analyzed weights in the overall cohort, we elected to perform a stratified analysis by self-reported race to determine if there was any additional variation in effect that could be explained by race/ethnicity-specific GRS subs-scores and weights. We were unable to detect multiplicative or additive interaction in the entire MESA cohort. The lack of multiplicative interaction for the overall cohort indicates that the hazard ratio that we observed did not exceed the expected joint probability of having the combination of each exposure together. In other words, the rate of obesity in the doubly exposed did not exceed the expected joint probability of having high genetic risk and high psychosocial stress together.

We observed several instances of additive interaction in African-Americans and Hispanic-Americans separately. Sub-additive interaction was present in African-American subjects for each of the three psychosocial factors, as well as in the composite measure for CSS/EHS, indicating a protective effect of having high genetic risk and high levels of stress, compared with what was expected. The results for African-American subjects may not be very informative, given that neither the oGRS-meta, nor the oGRS-AFA, was significantly associated with BMI levels at baseline. It is possible that these measures were not accurately capturing true genetic risk of obesity in African-Americans. That being the case, any evidence of interaction should be interpreted cautiously.

Positive additive interaction was observed in stratified analyses for Hispanic participants for the everyday hassles measure. It should be noted that this was only performed using the oGRS-meta, given that there is no current specific risk score developed for Hispanic-American adults. However, given that the oGRS-meta was associated with BMI at baseline in Hispanic subjects when adjusting for population stratification, it is possible that this result is, in fact, a true positive finding. To our knowledge, no studies have assessed the meta-analysis weighted oGRS in Hispanic populations. We observed null main effects for both exposures in adjusted models, yet a positive interaction on the additive scale. This was supported by the potential of a multiplicative interaction, although this result diminished after correcting for multiple testing. Nevertheless, interaction on the additive scale provides evidence that for Hispanic individuals with high genetic risk and high levels of everyday hassles, there is a greater likelihood of obesity than we would otherwise expect from simply adding the two main effects together. Thus, there is a subset of Hispanic-Americans who only became obese when exposed to both high genetic risk and high everyday hassles together.

The limited findings between psychosocial factors and genetic risk with obesity in other groups suggests that further research is needed to fully elucidate the interaction of genes and psychosocial factors in non-Hispanic populations. For example, there is a growing interest in epigenetic regulation of genetic risk factors. As genetic risk is predicated on the idea that risk genes are being over expressed in the body and thus directly biologically causing obesity, if there were factors that could regulate gene expression, then that might explain why people with high genetic risk and high stress are not experiencing obesity. This was true for African-American subjects in the current

study, where the protective interactive effect might be explained by exogenous factors that are regulating genetic expression of obesity risk genes (e.g. psychosocial stress). Evidence for the intersection of stress, epigenetics, and obesity can be seen in recent studies that have examined how stress reactivity is associated with DNA methylation levels, a known regulator of genetic expression.[45] Many non-genetic factors other than psychosocial risk, such as diet, smoking and environmental toxins have been associated with DNA methylation.[49-51] Additionally, an epigenome wide analysis conducted by Zhao et al. in a 2014 twin study examined methylation patterns in stress genes and found a significant association with obesity at the promoter region of *SLC6A4*.[61] Future studies should consider the role of epigenetic regulation between psychosocial factors and obesity risk genes.

Results that differed across each *type* of psychosocial factor were particularly noteworthy. While the CSS was designed to measure more objective instances of stressors that occur over a long period of time (e.g. job-related demands, difficulty paying bills, or social relationship demands), the EHS measure incorporates topics of *perceived* hassles or *perceived* unfair treatment.[153] Researchers have long been trying to determine the best measures for quantifying stress, and the debate between using objective versus subjective scales has persisted over the years. The argument exists that self-reported measures such as perceived stress are known to correlate with stress response biomarkers, which have been previously associated with obesity[20, 154] and with continuous weight gain.[24] Therefore, it is difficult to say with certainty that one measure is superior over another. Nevertheless, future studies would benefit from supplementing or validating self-reported measures with biomarkers of stress in the same population.

Strengths and limitations

The current study was strengthened by the use of robust statistical methods for survival analysis,[145, 146] detailed questionnaires that measured multiple psychosocial factors, and expansive coverage of obesity genetic risk factors. This is the largest and, to our knowledge, first study of its kind that assesses the interaction of a validated obesity genetic risk score with psychosocial factors on obesity in a population-based multi-ethic cohort.

Several limitations warrant comment. A low number of incident cases of obesity necessitated the use of interval censored survival analysis in lieu of traditional Cox models. This analysis allowed for the inclusion of left-censored prevalent cases at baseline to boost power. Although interval censored survival analysis was an improvement upon traditional Cox models, the current study was still underpowered to detect small main effect sizes. Additionally, the use of prevalent cases limited our ability to establish clear temporal ordering of psychosocial factors and obesity. Additionally, the MESA cohort was comprised of an older adult population, with some participants as old as 94 years of age at the end of follow-up. Older subjects who completed the study follow-up without becoming obese were considered right-censored, with the potential of becoming obese in the future, when in all likelihood this was not possible at such advanced ages.

There is a slight potential for spurious findings in race-stratified analyses given that the SNPs and weights used for oGRS-meta were derived from ancestries that did not include Hispanic populations. However, Hispanic ancestry has been found to be a mixture of European, Native

American, and African ancestries,[164] which could translate to obesity risk genes coming from a mixture of ancestries as well. The oGRS is gauging an overall genetic risk that is less susceptible to bias when including SNPs with weak effects.[38] If European derived SNPs had weak effects in Hispanic Americans, the overall score itself remained unbiased. To mitigate these concerns, we tested the association between oGRS-meta and increased BMI in Hispanic subjects at baseline, and results indicated that it was a useful predictor of genetic risk for obesity.

The score for oGRS-meta/oGRS-AFA in African-American subjects was less effective in determining increased BMI levels at baseline. Race/ethnic specific scores are perhaps not the best measures for genetic risk in this population, presenting a significant limitation of the current study. As mentioned previously, results in this population should be interpreted cautiously.

Our decision to use self-reported measures was primarily to assess if perception of stress could be effective in determining obesity risk. In the long run, the efficacy of self-reported measures of psychosocial factors will allow for larger studies with more comprehensive assessments of stress that are less expensive and easier to measure. Maximizing sample size by the use of less expensive methods will also increase the power to detect significant interactions.

The purpose of this study was to evaluate the hypothesis that genetic and psychosocial factors might interact to cause obesity. By assessing interaction, the current study established whether or not intervening on a modifiable exposure (psychosocial stress) to reduce obesity risk would be useful. Using comprehensive analytical methods, we have shown that an obesity genetic risk score and chronic burden of stress can independently impact obesity in a large multiethnic cohort. We found no clear evidence of an oGRS by psychosocial factor interaction when looking at the entire MESA cohort. There was evidence that positive interaction was present on the multiplicative and additive scales for the everyday hassles measure and oGRS-meta in Hispanic-Americans. Future studies would benefit from examining this relationship in an independent set of Hispanic-American adults. While an overall causal mechanism by which these factors biologically interact was not established, future studies should investigate how epigenetic mechanism might explain weak interactive effects in various populations.

3.5 Tables and Figures

Table 1. Demographic information on complete cases of MESA participants at baseline and follow-up (N=5,792)

	Baseline			Follow-up			
	Overall	**Non Obese**	**Prevalent Obese**	**Overall**	**Never Obese**	**Incident Obese**	**Incidence Rate[a]**
N # (%)	5,792	4,003 (69.1)	1,789 (30.1) **p<0.001**	4,003	3,535 (88.3)	468 (11.7) **p<0.001**	16
oGRS-meta[b]	2.17 (0.16)	2.16 (0.16)	2.19 (0.16) **p<0.001**	2.16 (0.16)	2.16 (0.16)	2.18 (0.16) **p<0.001**	
Quartile 1	1,448	1,090 (18.8)	358 (6.18)	1,090 (18.8)	980 (24.4)	110 (2.74)	14
Quartile 2	1,448	1,010 (17.4)	438 (7.56)	1,010 (17.4)	901 (22.5)	109 (2.72)	15
Quartile 3	1,448	994 (17.2)	454 (7.83)	994 (17.2)	879 (22.0)	115 (2.87)	17
Quartile 4	1,448	909 (15.7)	539 (9.30) **p<0.001**	909 (15.7)	775 (19.4)	134 (3.35) **p=0.008**	21
Chronic Burden of Stress[c]							
Low (0)	1,998 (34.5)	1,500 (25.9)	498 (8.60)	1,500 (25.9)	1,348 (33.7)	152 (3.79)	14
Moderate (1)	1,823 (31.4)	1,268 (21.9)	555 (9.58)	1,268 (21.9)	1,118 (27.9)	150 (3.74)	17
High (2+)	1,971 (34.0)	1,235 (21.3)	736 (12.7) **p<0.001**	1,235 (21.3)	1,069 (26.7)	166 (4.15) **p=0.027**	19
Everyday Hassles[c]							
Low (9)	1,587 (27.4)	1,177 (20.4)	410 (7.07)	1,177 (20.4)	1,050 (26.2)	127 (3.17)	16
Moderate (10-18)	3,028 (52.2)	2,081 (35.9)	947 (16.4)	2,081 (35.9)	1,849 (46.2)	232 (5.80)	15
High (19+)	1,177 (20.3)	745 (12.9)	432 (7.46) **p<0.001**	745 (12.9)	636 (15.9)	109 (2.72) **p=0.021**	19
High Stress Both CSS/EHS							
No	5,201 (89.7)	3,651 (63.0)	1,550 (26.8)	3,651 (63.0)	3,239 (80.9)	412 (10.3)	16
Yes	591 (10.3)	352 (6.08)	239 (4.12) **p<0.001**	352 (6.08)	296 (7.39)	56 (1.40) **p=0.013**	21
Depression							
No	5,072 (87.6)	3,542 (61.2)	1,530 (26.4)	3,542 (61.2)	3,136 (78.3)	406 (10.1)	16
Yes	720 (12.4)	461 (7.96)	259 (4.47) **p=0.002**	461 (7.96)	399 (9.97)	62 (1.55) p=0.241	21

Baseline Age[b]	62.0 (10.2)	62.6 (10.4)	61.0 (9.73) **p<0.001**	62.6 (10.4)	63.1 (10.4)	59.1 (9.23) **p<0.001**	
Sex							
Female	3,013 (52.0)	1,970 (34.0)	1,043 (18.0)	1,970 (34.0)	1,740 (43.5)	230 (5.75)	16
Male	2,779 (48.0)	2,033 (35.1)	764 (12.9) **p<0.001**	2,033 (35.1)	1,795 (44.8)	238 (5.95) p=1.00	17
Race/ethnicity							
Non-Hispanic White	2,342 (40.4)	1,687 (29.1)	655 (11.3)	1,687 (29.1)	1,483 (37.0)	204 (5.10)	16
Chinese-American	767 (13.2)	734 (12.7)	33 (0.57)	734 (12.7)	706 (17.6)	28 (0.70)	5
African-American	1,338 (23.1)	746 (12.9)	592 (10.2)	746 (12.9)	638 (15.9)	108 (2.70)	22
Hispanic-American	1,345 (23.2)	836 (14.4)	509 (8.78) **p<0.001**	836 (14.4)	708 (17.7)	128 (3.20) **p<0.001**	23
Baseline BMI[b]	28.2 (5.42)	25.4 (2.88)	34.6 (4.25) **p<0.001**	25.4 (2.88)	24.9 (2.75)	28.6 (1.32) **p<0.001**	
Physical Activity (Median(SD) MET hrs/week)	14.0 (39.1)	15.8 (40.7)	10.5 (33.3) **p<0.001**	15.8 (40.7)	15.9 (40.5)	15.9 (44.7) p=0.257	
Energy intake (Median(SD) kcal/day)	1,476 (863)	1,430 (819)	1,593 (949) **p<0.001**	1,430 (819)	1,408 (804)	1,540 (888) **p<0.001**	
Smoking status							
Never	2,963 (51.2)	2,075 (35.8)	888 (15.3)	2,075 (35.8)	1,856 (46.4)	219 (1547)	15
Ever	2,829 (48.8)	1,928 (33.3)	901 (15.6) p=0.129	1,928 (33.3)	1,697 (41.9)	249 (6.22) **p=0.023**	19
Income Median (SD)	2.18 (2.08)	2.25 (2.16)	2.08 (1.86) **p<0.001**	2.25 (2.16)	2.25 (2.18)	2.19 (2.03) p=0.944	
Education							
Less than high school diploma	1,039 (17.9)	710 (12.3)	329 (5.68)	710 (12.3)	619 (15.5)	91 (2.27)	21
High school grad/some college	1,936 (33.4)	1,254 (21.7)	682 (11.8)	1,254 (21.7)	1,088 (27.2)	166 (4.15)	19
Technical/associate/bachelor	1,736 (30.0)	1,223 (21.1)	513 (8.86)	1,223 (21.1)	1,085 (27.1)	138 (3.45)	15
Post college training	1,081 (18.7)	816 (14.0)	265 (4.58) **p<0.001**	816 (14.0)	743 (18.6)	73 (1.82) **p=0.019**	11

[a] Incidence density rates are reported as the number of cases per 1000 person-years.

[b] Continuous variables reported as a mean (standard deviation).

[c] While primary analyses for psychosocial variables were performed using dichotomous variables, tertiles are presented here to show a clear dose-response effect with increasing levels of the Chronic Burden of Stress Scale (CSS) and Everyday Hassles Scale (EHS) measures. Categorical variables were reported as the total number of subjects and proportion of all participants at baseline or follow-up. P-values for continuous variables were reported from a student's T-test for comparison of obesity cases to non-cases. P-values for categorical variables were reported from a Chi-squared test.

Table 2. Primary results from parametric interval censored survival analysis using Weibull regression (N=5,792)

		Model 1		Model 2	
		HR (95%CI)	**P-value[a]**	**HR (95%CI)**	**P-value[a]**
Chronic Burden of Stress (CSS)	**oGRS-high**	**1.18 (1.06, 1.32)**	**0.008**	**1.20 (1.08, 1.34)**	**0.004**
	CSS-high	**1.21 (1.07, 1.37)**	**0.010**	**1.18 (1.04, 1.35)**	**0.026**
	Interaction	0.99 (0.83, 1.17)	0.990	0.97 (0.82, 1.16)	0.990
	RERI	0.02 (-0.13, 0.17)	---	0.00 (-0.21, 0.20)	---
Everyday Hassles (EHS)	**oGRS-high**	**1.17 (1.06, 1.29)**	**0.007**	**1.18 (1.07, 1.30)**	**0.004**
	EHS-high	1.03 (0.89, 1.19)	0.990	1.04 (0.90, 1.21)	0.986
	Interaction	1.03 (0.84, 1.25)	0.990	1.02 (0.84, 1.24)	0.990
	RERI	0.04 (-0.13, 0.19)	---	0.03 (-0.2, 0.24)	---
High stress in Both CSS and EHS	**oGRS-high**	**1.17 (1.07, 1.28)**	**0.004**	**1.19 (1.08, 1.30)**	**0.004**
	HS-both	1.09 (0.91, 1.31)	0.732	1.09 (0.91, 1.30)	0.732
	Interaction	1.03 (0.81, 1.33)	0.990	1.00 (0.78, 1.29)	0.990
	RERI	0.06 (-0.15, 0.27)	---	0.02 (-0.28, 0.32)	---
Depression	**oGRS-high**	**1.17 (1.07, 1.29)**	**0.004**	**1.19 (1.08, 1.30)**	**0.004**
	Dep.	1.06 (0.89, 1.27)	0.934	1.03 (0.86, 1.23)	0.990
	Interaction	1.01 (0.79, 1.29)	0.990	1.00 (0.78, 1.27)	0.990
	RERI	0.03 (-0.18, 0.22)	---	0.00 (-0.26, 0.27)	---

Results from interval censored survival regression using a Weibull distribution are presented here. Model 1 adjusts for principal components, age, race, and sex. Model 2 represents the fully adjusted model that controls for all core covariates described in the methods section. The referent group is participants below the median level of the obesity genetic risk score (oGRS-meta), and low to moderate psychosocial factor. CSS stands for the chronic burden of stress scale. EHS stands for the everyday hassles scale. High stress in both CSS/EHS is compared to a referent population that has never experienced both psychosocial factors at a high level. Depression represents the comparison of subjects scoring above and below 16 on the CES-D scale. Hazard ratios and confidence intervals for main effects and multiplicative interaction were calculated using the delta method. RERI confidence intervals were calculated using model runs from 2000 bootstrapped samples. P-values were computed from Wald test z-statistics.

[a] P-values have been corrected for multiple testing using a false discovery rate across all categories of psychosocial factor (N=24 tests). Results in bold font are statistically significant at the FDR alpha 0.05 level.

Table 3. Race-stratified analyses for Model 2, fully adjusted model controlling for core covariates from parametric interval censored survival analysis using Weibull regression

	Overall (N=5,792)	Race Specific: White (N=2,342)		Black (N=1,338)		Hispanic (N=1,345)	Chinese[a] (N=767)
	oGRS-meta	**oGRS-meta**	**oGRS-EU**	**oGRS-meta**	**oGRS-AFA**	**oGRS-meta**	**oGRS-meta/-CHN**
CSS							
oGRS-high	**1.20 (1.08, 1.34)**	**1.41 (1.17, 1.69)**	**1.41 (1.17, 1.70)**	1.07 (0.87, 1.31)	1.04 (0.85, 1.28)	1.12 (0.91, 1.37)	---
CSS-high	**1.18 (1.04, 1.35)**	1.17 (0.92, 1.48)	1.18 (0.93, 1.49)	**1.33 (1.08, 1.64)**	**1.40 (1.12, 1.73)**	1.03 (0.82, 1.32)	---
Interaction	0.97 (0.82, 1.16)	0.96 (0.71, 1.28)	0.94 (0.70, 1.26)	0.86 (0.63, 1.18)	0.80 (0.59, 1.09)	1.13 (0.80, 1.58)	---
RERI	0.00 (-0.21, 0.20)	-0.01 (-0.27, 0.25)	-0.03 (-0.3, 0.23)	-0.17 (-0.47, 0.09)	**-0.27 (-0.55, -0.01)**	0.15 (-0.13, 0.41)	---
EHS							
oGRS-high	**1.18 (1.07, 1.30)**	**1.35 (1.15, 1.58)**	**1.35 (1.15, 1.58)**	1.11 (0.91, 1.34)	0.96 (0.79, 1.16)	1.05 (0.88, 1.25)	---
EHS-high	1.04 (0.90, 1.21)	1.03 (0.76, 1.38)	1.03 (0.77, 1.38)	1.14 (0.92, 1.41)	1.01 (0.81, 1.27)	0.72 (0.50, 1.02)	---
Interaction	1.02 (0.84, 1.24)	1.13 (0.79, 1.61)	1.12 (0.79, 1.59)	0.75 (0.54, 1.03)	0.99 (0.72, 1.36)	1.76 (1.12, 2.78)	---
RERI	0.03 (-0.2, 0.24)	0.18 (-0.11, 0.47)	0.18 (-0.12, 0.45)	**-0.31 (-0.55, -0.06)**	-0.01 (-0.25, 0.2)	**0.56 (0.24, 0.90)**	---
HS-both							
oGRS-high	**1.19 (1.08, 1.30)**	**1.37 (1.18, 1.59)**	**1.38 (1.18, 1.60)**	1.07 (0.90, 1.27)	0.96 (0.81, 1.14)	1.10 (0.93, 1.30)	---
HS-both	1.09 (0.91, 1.30)	1.13 (0.78, 1.62)	1.19 (0.83, 1.70)	1.33 (1.04, 1.71)	1.16 (0.89, 1.52)	0.68 (0.44, 1.05)	---
Interaction	1.00 (0.78, 1.29)	1.11 (0.71, 1.74)	1.02 (0.66, 1.59)	0.69 (0.47, 1.03)	0.96 (0.65, 1.40)	1.61 (0.91, 2.86)	---
RERI	0.02 (-0.28, 0.32)	0.23 (-0.21, 0.67)	0.1 (-0.35, 0.52)	**-0.42 (-0.75, -0.10)**	-0.06 (-0.36, 0.24)	**0.42 (0.03, 0.80)**	---
Dep.							
oGRS-high	**1.19 (1.08, 1.30)**	**1.39 (1.19, 1.62)**	**1.39 (1.19, 1.61)**	1.03 (0.87, 1.21)	1.01 (0.86, 1.20)	1.11 (0.93, 1.33)	---
Dep.	1.03 (0.86, 1.23)	1.15 (0.79, 1.66)	1.16 (0.80, 1.69)	1.17 (0.87, 1.58)	1.42 (1.02, 1.97)	0.86 (0.64, 1.16)	---
Interaction	1.00 (0.78, 1.27)	0.96 (0.61, 1.52)	0.94 (0.60, 1.47)	0.80 (0.51, 1.27)	0.60 (0.38, 0.95)	1.19 (0.81, 1.77)	---
RERI	0.00 (-0.26, 0.27)	0.01 (-0.40, 0.42)	-0.04 (-0.46, 0.36)	-0.23 (-0.59, 0.14)	**-0.58 (-1.03, -0.22)**	0.17 (-0.13, 0.44)	---

Race/ethnicity specific results are reported for the obesity genetic risk score (oGRS-meta) as well as Racial/ethnic oGRS sub-scores. The referent group is participants below the median level of oGRS, and low to moderate psychosocial factor. CSS stands for the chronic burden of stress scale. EHS stands for the everyday hassles scale. High stress in both CSS/EHS is compared to a referent population that has never experienced both psychosocial factors at a high level. Depression represents the comparison of subjects scoring above and below 16 on the CES-D scale. Hazard ratios and confidence intervals for main effects and multiplicative interaction were calculated using the delta method. RERI confidence intervals were calculated using model runs from 2000 bootstrapped samples. Non-RERI results in **bold** font are statistically significant at the FDR corrected 0.05 level (N=60 tests). Bolded RERI values are based on confidence intervals, as p-values were not calculable.

[a] Due to low sample size, we were unable to draw meaningful conclusions and therefore, results will not be presented. Due to the very low incidence of obesity among Chinese participants, these models had difficulties with convergence.

Supplementary Table S1. Hazard ratios and p-values from testing 3-way interactions for continuous main exposures by race/ethnicity and Sex (N=5,792)

	HR (95% CI)	**P-value**
oGRS-meta x race/ethnicity x CSS	1.01 (0.98, 1.04)	0.386
oGRS-meta x race/ethnicity x EHS	1.00 (0.996, 1.01)	0.408
oGRS-meta x race/ethnicity x Dep	0.99 (0.91, 1.09)	0.909
oGRS-meta x Sex x CSS	1.01 (0.98, 1.04)	0.694
oGRS-meta x Sex x EHS	1.00 (0.99, 1.01)	0.166
oGRS-meta x Sex x Dep	1.08 (0.83, 1.42)	0.548

Table S1 shows the three-way interaction results performed in order to determine if models needed to be stratified by race or sex. All models adjusted for pc1-3 to account for the influence of population stratification, as well as baseline measures of age, race, and sex. Models represent the hazard of obesity given 1 standard deviation increase in continuous oGRS using "meta" weights. CSS stands for the Chronic Burden of Stress Scale, EHS stands for the Everyday Hassles Scale, and Dep stands for depression.

Supplementary Table S2. Features of obesity risk genes and weights used for oGRS measures (N= 97 SNPs)

rsID	Gene located at or near loci	Chr	Risk/non-risk allele (overall risk allele frequency in MESA)	oGRS weights (non-Hispanic White)	oGRS weights (African-American)	oGRS weights (Chinese-American)	oGRS Meta weights
rs1000940	*RABEP1*	17	G/A (0.326)	0.0192	0.0149	0.0104	0.018
rs10132280	*STXBP6*	14	C/A (0.631)	0.0230	0.0104	0.033	0.022
rs1016287	*LINC01122*	2	T/C (0.258)	0.0229	-0.0076	0.0051	0.023
rs10182181	*ADCY3*	2	G/A (0.566)	0.0307	0.0427	0.0335	0.031
rs10733682	*LMX1B*	9	A/G (0.473)	0.0174	0.0061	0.0161	0.019
rs10938397	*GNPDA2*	4	G/A (0.338)	0.0402	0.0534	0.0366	0.040
rs10968576	*LINGO2*	9	G/A (0.226)	0.0249	0.0374	0.0104	0.025
rs11030104	*BDNF*	11	A/G (0.825)	0.0414	0.0904	0.0478	0.042
rs11057405	*CLIP1*	12	G/A (0.953)	0.0307	0.0339	Monomorphic / rare (<1%)	0.030
rs11126666	*KCNK3*	2	A/G (0.263)	0.0207	-0.0124	0.0100	0.020
rs11165643	*PTBP2*	1	T/C (0.482)	0.0218	0.0216	0.0053	0.022
rs11191560	*NT5C2*	10	C/T (0.117)	0.0308	0.0260	0.0271	0.031
rs11583200	*ELAVL4*	1	C/T (0.578)	0.0177	0.0175	0.0083	0.017

rs1167827	*HIP1*	7	G/A (0.589)	0.0202	0.0132	0.0063	0.020
rs11688816	*EHBP1*	2	G/A (0.574)	0.0172	-0.0032	0.0053	0.015
rs11727676[a]	***HHIP***	**4**	**T/C (NA)**	**0.0358**	**-0.0081**	**Monomorphic / rare (<1%)**	**0.037**
rs11847697	*PRKD1*	14	T/C (0.153)	0.0492	-0.0041	Monomorphic / rare (<1%)	0.037
rs12016871 [a]	***MTIF3***	**13**	**T/C (NA)**	**0.0298**	**0.0127**	**0.0103**	**0.030**
rs12286929	*CADM1*	11	G/A (0.526)	0.0217	0.0049	0.0149	0.021
rs12401738	*FUBP1*	1	A/G (0.219)	0.0211	0.0126	Monomorphic / rare (<1%)	0.020
rs12429545	*OLFM4*	13	A/G (0.148)	0.0334	0.0400	0.0309	0.032
rs12446632	*GPRC5B*	16	G/A (0.903)	0.0403	0.0134	Monomorphic / rare (<1%)	0.040
rs12566985	*FPGT*	1	G/A (0.633)	0.0242	0.0104	0.0211	0.024
rs12885454	*PRKD1*	14	C/A (0.735)	0.0207	0.0127	0.0058	0.020
rs12940622	*RPTOR*	17	G/A (0.556)	0.0182	0.0194	0.0164	0.018
rs13021737	*TMEM18*	2	G/A (0.853)	0.0601	Insufficient data available	0.0589	0.060
rs13078960	*CADM2*	3	G/T (0.129)	0.0297	0.0018	Monomorphic / rare (<1%)	0.029
rs13107325	*SLC39A8*	4	T/C (0.059)	0.0477	0.0530	Monomorphic / rare (<1%)	0.047
rs13191362	*PARK2*	6	A/G (0.924)	0.0277	0.0366	Monomorphic / rare (<1%)	0.029
rs13201877	*IFNGR1*	6	G/A (0.084)	0.0233	0.0197	0.0067	0.024
rs1441264	*MIR548A2*	13	A/G (0.659)	0.0175	-0.0122	-0.0022	0.017

rs1460676	*FIGN*	2	C/T (0.200)	0.0197	0.0163	0.0211	0.021
rs1516725	*ETV5*	3	C/T (0.873)	0.0451	0.0328	0.0547	0.045
rs1528435	*UBE2E3*	2	T/C (0.626)	0.0178	0.0030	0.0097	0.018
rs1558902	*FTO*	16	A/T (0.253)	0.0818	0.0658	0.0734	0.081
rs16851483	*RASA2*	3	T/G (0.095)	0.0483	0.0817	0.0171	0.048
rs16907751	*ZBTB10*	8	C/T (0.911)	0.0350	0.0005	0.0454	0.033
rs16951275	*MAP2K5*	15	T/C (0.607)	0.0311	0.0260	0.0273	0.030
rs17001654	*SCARB2*	4	G/C (0.172)	0.0306	-0.0013	0.0237	0.030
rs17024393	*GNAT2*	1	C/T (0.038)	0.0658	0.0101	Monomorphic / rare (<1%)	0.061
rs17094222	*HIF1AN*	10	C/T (0.178)	0.0249	0.0175	0.0138	0.025
rs17203016	*CREB1*	2	G/A (0.121)	0.0210	-0.0199	0.0225	0.021
rs17405819	*HNF4G*	8	T/C (0.767)	0.0224	0.0094	0.0199	0.022
rs17724992	*PGPEP1*	19	A/G (0.743)	0.0194	0.0225	0.0187	0.020
rs1808579	*C18orf8*	18	C/T (0.492)	0.0167	-0.0013	0.0019	0.016
rs1928295	*TLR4*	9	T/C (0.559)	0.0188	-0.0086	0.0138	0.018
rs2033529	*TDRG1*	6	G/A (0.199)	0.0190	0.0126	-0.0037	0.018
rs2033732	*RALYL*	8	C/T (0.791)	0.0192	-0.0117	0.0047	0.018
rs205262	*C6orf106*	6	G/A (0.390)	0.0221	0.0159	0.0115	0.021

rs2075650	*TOMM40*	19	A/G (0.887)	0.0258	Insufficient data available	0.0156	0.026
rs2080454	*CBLN1*	16	C/A (0.462)	0.0168	0.0146	0.0108	0.017
rs2112347	*POC5*	5	T/G (0.563)	0.0261	0.0021	0.0214	0.025
rs2121279 [a]	***LRP1B***	**2**	**T/C (NA)**	**0.0245**	**-0.0315**	**Monomorphic / rare (<1%)**	**0.024**
rs2176040	*LOC646736*	2	A/G (0.227)	0.0141	0.0122	0.0037	0.015
rs2176598	*HSD17B12*	11	T/C (0.320)	0.0198	0.0126	-0.0031	0.019
rs2207139	*TFAP2B*	6	G/A (0.175)	0.0447	0.0554	0.0296	0.045
rs2245368	*DTX2P1*	7	C/T (0.259)	0.0317	-0.0084	0.0212	0.029
rs2287019	*QPCTL*	19	C/T (0.859)	0.0360	0.0659	0.0252	0.035
rs2365389	*FHIT*	3	C/T (0.367)	0.0200	Insufficient data available	0.0059	0.020
rs2650492	*SBK1*	16	A/G (0.147)	0.0207	0.0145	0.0167	0.021
rs2820292	*NAV1*	1	C/A (0.414)	0.0195	0.0134	0.0051	0.018
rs2836754	*ETS2*	21	C/T (0.467)	0.0164	0.0045	0.0176	0.017
rs29941	*KCTD15*	19	G/A (0.683)	0.0182	-0.0002	-0.0018	0.018
rs3101336	*NEGR1*	1	C/T (0.662)	0.0334	0.0036	0.0203	0.032
rs3736485	*DMXL2*	15	A/G (0.541)	0.0176	0.0094	0.0047	0.016
rs3810291	*ZC3H4*	19	A/G (0.451)	0.0283	0.0189	0.0338	0.029
rs3817334	*MTCH2*	11	T/C (0.353)	0.0262	0.0083	0.0215	0.026

rs3849570	*GBE1*	3	A/C (0.399)	0.0188	-0.0075	-0.0044	0.018
rs3888190	*ATP2A1*	16	A/C (0.319)	0.0309	0.0174	0.0252	0.031
rs4256980	*TRIM66*	11	G/C (0.560)	0.0209	0.0200	0.0137	0.021
rs4740619	*C9orf93*	9	T/C (0.482)	0.0179	0.0033	-0.0042	0.017
rs4787491	*INO80E*	16	G/A (0.501)	0.0159	0.0222	0.0137	0.015
rs492400	*USP37*	2	C/T (0.446)	0.0158	0.0025	0.0022	0.015
rs543874	*SEC16B*	1	G/A (0.202)	0.0482	0.0571	0.0648	0.050
rs6091540	*ZFP64*	20	C/T (0.270)	0.0188	0.0167	0.0124	0.019
rs6465468	*ASB4*	7	T/G (0.190)	0.0166	-0.0037	0.0005	0.016
rs6477694	*EPB41L4B*	9	C/T (0.405)	0.0174	0.0057	0.0100	0.017
rs6567160	*MC4R*	18	C/T (0.196)	0.0556	0.0621	0.0487	0.056
rs657452	*AGBL4*	1	A/G (0.485)	0.0227	-0.0056	0.0142	0.023
rs6804842	*RARB*	3	G/A (0.520)	0.0185	0.0110	0.0070	0.018
rs7138803	*BCDIN3D*	12	A/G (0.276)	0.0315	0.0470	0.0193	0.032
rs7141420	*NRXN3*	14	T/C (0.564)	0.0235	0.0183	0.0114	0.023
rs7164727	*LOC100287559*	15	T/C (0.500)	0.0180	0.0057	-0.0033	0.019
rs7239883	*LOC284260*	18	G/A (0.369)	0.0164	0.0041	0.0058	0.015
rs7243357	*GRP*	18	T/G (0.817)	0.0217	0.0243	0.0186	0.022

rs758747	*NLRC3*	16	T/C (0.403)	0.0225	0.0046	0.0229	0.023
rs7599312	*ERBB4*	2	G/A (0.734)	0.0220	0.0078	0.0078	0.021
rs7715256	*GALNT10*	5	G/T (0.518)	0.0163	0.0410	0.0437	0.017
rs7899106	*GRID1*	10	G/A (0.070)	0.0395	0.0303	Monomorphic / rare (<1%)	0.038
rs7903146	*TCF7L2*	10	C/T (0.738)	0.0234	0.0305	0.0255	0.024
rs9374842	*LOC285762*	6	T/C (0.781)	0.0187	0.0066	-0.0026	0.020
rs9400239	*FOXO3*	6	C/T (0.529)	0.0188	-0.0003	0.0209	0.017
rs9540493 [a]	***MIR548X2***	**13**	**A/G (NA)**	**0.0172**	**0.0180**	**0.0130**	**0.018**
rs9641123	*CALCR*	7	C/G (0.259)	0.0191	0.0037	0.0069	0.019
rs977747	*TAL1*	1	T/G (0.573)	0.0167	-0.0077	0.0133	0.017
rs9914578	*SMG6*	17	G/C (0.328)	0.0201	-0.0095	0.0268	0.020
rs9925964	*KAT8*	16	A/G (0.647)	0.0192	0.0127	Monomorphic / rare (<1%)	0.020
				97 total SNPs	74 w/ directional consistency	77 w/ directional consistency	

This table present a full list of BMI associated genes, along with SNPs that are specific to multiple race/ethnicities. Weights were derived from a meta-analysis conducted by Locke et al. 2015. The oGRS-meta weights are derived from the full meta-analysis in subjects of "all ancestries." Weights derived from homogenous populations of European, African, or east Asian ancestry were used in race/ethnicity specific sub-scores for MESA's non-Hispanic White, African-American, and Chinese-American participants respectively.

[a] The **bolded** SNPs were not available in MESA after quality control measures were implemented.

Supplementary Table S3. Continuous baseline BMI by oGRS-meta, adjusting for pc1-3, age, and sex at baseline.

	Overall (N=5,792) Beta (95%CI)	**Non-Hispanic White (N=2,342) Beta (95%CI)**	**African-American (N=1,338) Beta (95%CI)**	**Asian-American (N=767) Beta (95%CI)**	**Hispanic (N=1,345) Beta (95%CI)**
oGRS-meta	**0.56 (0.42,0.69)**	**0.65 (0.46, 0.86)**	0.29 (-0.06, 0.65)	**0.32 (0.05,0.59)**	**0.65 (0.38,0.93)**
oGRS-EU		**0.66 (0.46, 0.86)**			
oGRS-AFA			-0.03 (-0.35, 0.30)		
oGRS-CHN				0.20 (-0.03, 0.43)	
oGRS-meta					
Q2	**0.49 (0.13, 0.86)**	0.22 (-0.41, 0.87)	**0.87 (0.08, 1.67)**	0.41 (-0.14, 0.96)	0.55 (-0.25, 1.36)
Q3	**0.63 (0.26, 1.00)**	**0.92 (0.30, 1.54)**	0.26 (-0.57, 1.09)	0.53 (-0.14, 1.19)	0.42 (-0.37, 1.20)
Q4	**1.45 (1.08, 1.84)**	**1.60 (1,01, 2.19)**	**0.99 (0.01, 1.96)**	0.63 (-0.28, 1.53)	**1.61 (0.83, 2.40)**

Results are shown for the effect on continuous BMI for every 1 standard deviation increase in oGRS in the overall MESA cohort and race/ethnicity specific subsets. Items in **bold** are statistically significant at the alpha 0.05 level. Weights for obesity genetic risk score (oGRS) measures were derived from a meta-analysis conducted by Locke et al. 2015. The oGRS-meta weights are derived from the full meta-analysis in subjects of "all ancestries." Weights derived from homogenous populations of European, African, or east Asian ancestry were used in race/ethnicity specific sub-scores for MESA's non-Hispanic White (oGRS-EU), African-American (oGRS-AFA), and Chinese-American (oGRS-CHN) participants respectively. Quartiles of the oGRS-meta score are presented comparing the top three quartiles (Q2,Q3,Q4) to those in the lowest quartile of genetic risk (Q1, not shown).

Supplementary Table S4. Primary results from parametric interval censored survival analysis using Weibull regression, stratified by oGRS categories (N=5,792)

oGRS-meta

		Below Median		Above Median		
		N with/without obesity	HR (95%CI); p-value	N with/without obesity	HR (95%CI); p-value	E1 vs. E0 within strata of G
Chronic Burden of Stress (CSS)	**<2**	599/1329	1.00 (Ref)	756/1137	**1.20 (1.08, 1.34)**	
	2+	416/552	**1.18 (1.04, 1.35)**	486/517	**1.39 (1.22, 1.57)**	**1.19 (1.04, 1.36)**
					1.15 (1.02, 1.29)	
Everyday Hassles (EHS)	**<19**	764/1541	1.00 (Ref)	952/1358	**1.18 (1.07, 1.30)**	
	19+	251/340	1.04 (0.90, 1.21)	290/296	**1.29 (1.09, 1.45)**	1.17 (0.99, 1.40)
					1.08 (0.93, 1.25)	
High stress in Both CSS and EHS	**No**	874/1716	1.00 (Ref)	1088/1523	**1.19 (1.08, 1.30)**	
	Yes	141/165	1.09 (0.91, 1.30)	154/131	**1.29 (1.05, 1.55)**	1.15 (0.90, 1.46)
					1.1 (0.92, 1.31)	
Depression	**No**	871/1676	1.00 (Ref)	1065/1460	**1.19 (1.08, 1.30)**	
	Yes	144/205	1.03 (0.86, 1.23)	177/194	**1.22 (1.03, 1.45)**	1.19 (0.95, 1.49)
					1.04 (0.88, 1.23)	

Results are from models fully adjusted for the core covariates. The hazard ratio (HR) comparing psychosocial factor high to low within high obesity genetic risk scores (oGRS-meta) is extracted from the model restricted to only those with high oGRS-meta. This means that the HR and confidence intervals calculated for that measure are derived from the delta method in the restricted model, and not a direct calculation of the high stress/high genetic risk HR compared to the low stress/high genetic risk HR. The same can be said for the HR of oGRS-meta within strata of psychosocial factor. The number of people with and without obesity is a combination of prevalent and obese cases. Confidence intervals for the HR in the doubly-exposed was derived from 2000 bootstrapped samples.

Chapter 4

Psychosocial factors and DNA methylation of obesity risk genes

4.1 Introduction

Obesity is a growing public health concern as it is a major risk factor for chronic diseases such as cardiovascular disease and an independent predictor of overall mortality.[5] The prevalence of obesity in the US has been increasing over time, and there are significant racial/ethnic differences in obesity such that racial and ethnic minorities have a higher prevalence compared to whites.[3] While the exact causes of obesity remain unknown, numerous biological, behavioral, and social factors have been linked to obesity in previous research.

Recently, there has been a strong argument for a genetic contribution to obesity. Over the past 30 years, many genetic epidemiologic studies including candidate gene studies, genome linkage, and fine mapping have identified several single nucleotide polymorphisms (SNPs) associated with BMI. However, now with genome wide association studies (GWAS), which improve upon other approaches by allowing for the examination of a wide range of common variants associated with disease, new discoveries have been made. GWAS studies have identified up to 97 genetic predictors of higher BMI, including, but not limited to, *FTO*, *TMEM18*, and *MC4R (B)*.[36]

Despite recent advancements in the genetics of obesity, genetic variation in even the strongest predictor, *FTO*, confers at best an increase in BMI of 0.39 for each copy of the risk allele (explaining only 0.34% of total genetic variance).[35] Researchers have been tasked with incorporating genetic information in ways that can explain a larger proportion of the variance in BMI. DNA methylation (DNAm), the most common epigenetic mechanism explored currently, serves as a method for elucidating the effects of genetic risk factors on BMI that are not directly caused by the heritable genomic sequence. This is due to the regulatory nature of DNAm, the addition of a methyl group to cytosine/guanine base pairs (CG dinucleotides or CpGs), which affects genetic expression most prominently by either directly blocking protein binding or by long term inactivation of gene promoter regions.[165]

As DNAm is affected globally and site-specifically by non-heritable obesity risk factors (e.g. diet,[166] physical activity[120] and smoking[51]), it is an attractive approach for exploring the effects of the "environment" on genetic pre-disposition for obesity phenotypes or BMI more generally. Many studies have examined the effect of behavioral factors such as exercise or smoking on obesity epigenetics in a multitude of ways.[120, 167] However, to our knowledge, very few have examined the role of psychosocial factors on methylation,[62, 63] and none has used a candidate gene approach based on all BMI associated genes discovered from GWAS.

This is a curious omission from the current literature given the established role of multiple psychosocial stressors/factors known to affect weight in adults, such as job-related demands, lack of decision authority, perceived constraints in life, and strain in family relations.[20] Further, prior research has highlighted how early life exposures to stress can result in observable epigenetic

modifications later in life, both globally and at specific genes.[62, 63] Less is known, however, about modifications that occur due to psychosocial factors experienced in adulthood. Although the mechanisms by which this occurs are not exactly clear, understanding that genes can be affected by non-genetic factors may elucidate new pathways by which psychosocial factors can cause weight gain and ultimately, obesity.

To address the aforementioned gaps in the current literature, we used the Multi-Ethnic Study of Atherosclerosis (MESA) cohort study to investigate the associations between psychosocial factors in adulthood and DNAm in BMI-associated genes using a targeted gene approach. The overall goal was to examine the effects of stress on DNAm of BMI associated genes in order to establish a new mechanism of obesity etiology. Based on the evidence presented thus far regarding the complex relationship between stress, DNAm, and obesity, we hypothesized that there would be statistically significant associations between psychosocial factors and DNAm levels of candidate genes. Further, as DNAm is a known regulator of genetic expression,[52] we hypothesized that there would be statistically significant inverse association between DNAm levels in candidate genes of interest and genetic expression.

4.2 Methods

Subjects

MESA study subjects were recruited from six field centers in the US: Columbia University, New York; Johns Hopkins University, Baltimore; Northwestern University, Chicago; UCLA, Los Angeles; University of Minnesota, Twin Cities; Wake Forest University, Winston Salem. The population-based longitudinal cohort, comprised of 6,814 participants free of clinical CVD at baseline (approximately 1,100 from each field center), was designed for the purpose of investigating the prevalence, correlates, and progression of subclinical CVD and the risk factors that contribute to it. (http://mesa-nhlbi.org/). Participants were recruited from each field center's geographic boundaries by a random selection of names within each census tract, as well as telephone recruitment, church membership rosters, DMV registries, etc. (please see www.mesa-nhlbi.org, paragraph 4.2 for detailed recruitment procedures).[122] Each study participant who consented to enroll in the cohort had baseline information collected via questionnaire in July 2000, and followed up at four subsequent time points for a total of five exams to assess clinical outcomes and mortality, ending in January 2012.

Epigenetic information was collected between April 2010 and February 2012 (MESA Exam 5) as part of a genetic sub-study of 1,264 MESA participants who gave informed consent for DNA extraction and had available DNA for methylation analysis. Epigenetic study participants were adults ranging from 44 to 83 years of age (mean=60.2; sd=9.49) and were balanced equally by sex and by racial/ethnic subgroup. The current study utilizes two time points for all primary analyses (Exam1: baseline and Exam5: end of follow-up). Institutional Review Board (IRB) approval was obtained from each of the participating study sites in MESA.

Methylation assay and QC measures

DNA from peripheral blood leukocytes was used to interrogate DNA methylation markers, referred to as CpGs. Using the Illumina Infinium HumanMethylation450 BeadChip microarray (450k) (Illumina, Inc., San Diego, CA USA) and processed according to manufacturer protocols. This assay measures signal intensities of methylated and unmethylated probe sequences, and reports Average Beta values, a proxy measure for percent methylation of cytosine's at >480,000 CpGs. Extensive quality control (QC) measures were performed in order to ensure the reliability and quality of DNAm results using GenomeStudio software developed by Illumina.

QC was performed in two phases: pre-processing and CpG filtering. Pre-processing of all 450k data included quantile normalization, as well as background subtraction and internal control normalization according to default settings in the *lumi* R package.[168] Additionally, checks were made for sex and race/ethnicity incongruity along with outlier identification by multidimensional scaling (MDS) plots. CpG filtering was conducted according to the following criteria: 1) removed probes with sub-optimal detection (p-value >0.05 in at least 10% of samples) (N=695 CpGs); 2) removed probes containing SNPs within 10 base pairs of the CpG of interest (N=36,359 CpGs) along with 65 "rs" coded highly polymorphic SNPs that are not related to DNAm.[169]; 3) removed non-autosomal CpGs (N=10,923).[169] Supplementary Figure S1 shows a detailed flowchart of CpG removal procedures. Of the 437,600 post-QC CpGs available for analysis, 2,528 were selected from BMI associated genes established in the most recent GWAS (see Table 1 for list of gene descriptions).[35, 36, 170] CpGs were assigned to genes based on Illumina annotation.

Each gene-specific CpG was given an M-value, calculated as the $\log_2$ ratio of the methylated to the unmethylated intensities.[171] This measures ranges from completely unmethylated (-6) to completely methylated (6). M-values were used to meet the assumptions of normality and homoscedasticity required for the analytical methods employed in the current study.

Prior evidence has indicated that *hyper*methylation in the promoter region of a gene is associated with reduced gene expression, whereas *hyper*methylation in the gene body is associated with increased gene expression [172]. Therefore, incorporating information on promoter status was essential for all analyses. Based on methods previously developed by Whitaker et al, each CpG was designated as being within a promoter region if the site was 2500 bases upstream or 500 bases downstream of a transcription start site.[173] Transcription start site was identified by the RefSeq Genes track in the UCSC Genome Browser.[174] Shore/Shelf status at each CpG was determined according to Illumina annotation, which defines CpG island shores or shelves as being located 4000 bp outside of CpG island boundaries (North or South).

During sample processing, in order to mitigate a potential bias from batch effects, samples were randomly allocated to chips and chip positions using stratified sampling. Additionally, residual chip and chip position effects were adjusted prior to analysis. This was done by first running a linear regression of chip number and chip position on DNAm by each CpG separately. The residual differences from each model were then added to the mean value of the CpG for that model, yielding chip-corrected DNAm estimates.

Gene expression

Available DNA from MESA Exam 5 was used for transcriptomic (expression) data processed with the Illumina HumanHT-12 v4 Expression BeadChip, which interrogates approximately 48,000 transcripts across the genome. QC measures for expression data included the following: 1) background correction performed in GenomeStudio; 2) Bead-type summarization using the *beadarray* package in R[175]; 3) Estimated non-negative signal, performed quantile normalization, log transformation, eliminated control probes, and detected outliers using the *limma* package in R.[176-178] Transcripts were filtered out according to the following criteria: 1) poorly detected in at least 10% of all samples (p-value >0.01); 2) probes that contain a SNP; 3) probes with low variance across samples (<10^{th} percentile). 4) probes that overlap with non-unique regions. A more detailed description of implemented QC protocol for DNAm and expression data can be found in Liu et al.[179] Chip effects were adjusted for expression data prior to analysis, using identical methods to the DNAm correction.

Psychosocial factors

The chronic burden of stress scale (CSS) is measured as the sum of the number of times a subject answered yes to the following ongoing stressful problems (own health, close person health, job, financial, or relationship).[180] This problem needed to last more than six months to be considered affirmative. Possible scores of CSS range from 0 to 5 and CSS was analyzed categorically as high (2 or more), medium (1) and low (0; referent) chronic stress, based on previous methods.[127]

The Everyday Hassles Scale (EHS) is a validated nine-item scale that measures day-to-day incidents of unfair treatment based on the frequency of encounters in which the subject was treated unfairly.[128] Examples of perceived unfairness include being called names, implied as a liar or unintelligent, etc. Each item ranges from 1-6 (1=never to 6=almost every day), with the final score calculated as a sum of all nine items. A higher score on the scale represents a higher frequency of hassles each day and therefore more stressors. This scale has been previously validated for use in multi-ethnic populations.[181, 182] The scale will be examined based on previous literature as a categorical variable: 9 = no discrimination (low), 10 to 18 = moderate discrimination (med), or >18 = high discrimination (high).[87, 124, 181] Internal reliability /consistency estimates of the EHS measure were fairly good based on commonly accepted standards (Cronbach's alpha= 0.88, 95%CI (0.85, 0.88)).

The validated, four-item Cohen's perceived stress scale (PSS-4) [31, 183] was used as the third psychosocial factor. The scale was only available at MESA Exam 5, and calculated as the sum of four questions related to control, confidence, things going your way, and insurmountable difficulties (possible values range from 0-never to 5-always). The PSS-4 had questionable internal reliability/consistency (Cronbach's alpha= 0.65, 95%CI (0.67, 0.70)), this is consistent with previous studies examining the validity of the PSS-4 with more extensive versions of the scale.[184] As the Cronbach measure is sensitive to the number of components included in the test,[134] the estimate may be attributed to the PSS only being comprised of four questions. Given the suboptimal level of reliability in the measure and to maximize statistical power, the PSS-4 was analyzed as a dichotomous variable by the midpoint of the scale (Low ≤8 vs. high >8). A value

greater than 8 on the PSS-4 scale ensures that an individual experienced stress "fairly often" for at least one component. For information on the subcomponents of each psychosocial factor, please refer to supplementary table S1.

Covariate descriptions

DNAm models were adjusted for the following core set of covariates: age, sex, BMI, race/ethnicity, education, income, and enrichment scores for each of four major white blood cell types (neutrophils, B cells, T cells, and natural killer cells) to correct for differential peripheral blood composition. All covariates were baseline measures evaluated at MESA Exam 1, with the exception of cell composition, which was assessed at MESA Exam 5, when sample collection for the Epigenetic study was conducted.

Height and weight measurements for calculating continuous BMI were taken at clinical MESA Exam 1 and reported in kg/m^2. There were three racial/ethnic groups in the current study (non-Hispanic White, African-American, and Hispanic). Education was measured categorically by the following designations: 1) Less than high school diploma; 2) High school graduate or some college, but no degree; 3) Technical/associate/bachelor degree; and 4) Post college training. Income was calculated as a continuous measure, representing per capita adjusted household income per $10,000 or (continuous income/# people supported)/10,000).

Supplementary analyses adjusted for physical activity (total metabolic equivalent of task (MET) minutes/week of all light, moderate, and vigorous activity). Total physical activity was self-reported at MESA Exam 1 using a detailed semi-quantitative questionnaire based on the Cross-Culture Activity Participation Study.[121] Additionally, total caloric intake (kcal/day) was assessed at MESA Exam 1 using a food frequency questionnaire that was adapted from the Insulin Resistance Atherosclerosis Study to ensure cross-cultural validity.[121] Physical activity and caloric intake were incorporated in order to account for stress related behavioral factors that may affect DNAm.

Lastly, for any significant findings, we explored if additionally controlling for saturated, monounsaturated, and polyunsaturated fat intake mitigated the effects. This was performed to account for information in the current literature that found Western-like diet patterns (high saturated fat intake) exacerbated cardiovascular and hypothalamic-pituitary-adrenal (HPA) responses to chronic stress.[185] Therefore, we intended to control for the indirect effect of stress on DNAm by fat consumption.

Statistical analyses

Initial analyses included descriptive distributions, as well as bivariate tests between each covariate and each psychosocial measure. Bivariate tests were conducted for continuous variables using a Student's T-test for the PSS-4 and an analysis of variance (ANOVA) for CSS and EHS measures. Categorical comparisons were tested using a standard Chi-squared test. Tests were considered statistically significant at the alpha 0.05 level.

To avoid multiple testing concerns associated with conducting thousands of statistical tests for all CpGs of interest, we employed a two-level model for each unique gene (87 genes from 97 obesity GWAS SNPs) treating the CpG as the level-1 unit and the individual as the level-2 unit. Hierarchical models allow for repeated measures in longitudinal data, and analogously, we treated multiple CpGs within each gene as a "repeated measure" within an individual. The two-level model allowed us to fit a regression to the individual measurements while accounting for systematic unexplained variation among CpGs in the each gene. Each individual had a unique random intercept, which allowed for "within-subject" correlation of gene-specific CpGs to vary across individuals, along with fixed effects for the association between psychosocial factor and gene-specific DNAm. Such techniques have proven to be effective in prior DNAm studies of similar design.[110, 186]

The two-level model also allowed for CpG specific variables such as shore/shelf status and promoter status, along with an interaction term of CpG type (shore/shelf or promoter indicator) and psychosocial factor, to be included in all gene-level analyses.

The basic model was as follows:

$$Y_{ij} = \beta_0 + \beta_1 PF_i + \beta'_2 Covariates_i + \beta'_3 I_j + \beta'_4 PF_i \cdot I_j + \eta_i + \epsilon_{ij}$$

Y_{ij}: The methylation value for site j for individual i within each given gene

PF_i: Psychosocial Factor measure on individual i

$Covariates_i$: Covariate measures on individual i

I_j: Promoter region indicator ($I_j = 1$ affirmative; else $I_j = 0$)

β_0: Overall mean of M-value when all predictors are 0 (model intercept)

β_1: Difference in M-value between psychosocial factor categories

β'_2: Vector of parameter estimates by demographic covariates

β'_3: Vector of parameter estimates by CpG type (promoter or shore/shelf)

β'_4: Vector of parameter estimates for the difference in psychosocial factor effect on the M-value between promoter and non-promoter sites or between shore/shelf and non-shore/shelf sites/

η_i : Random effect for individual i, $\eta_i \sim N(0, \sigma^2_{individual})$

ϵ_{ij}: Residual site-level error $\epsilon_{ij} \sim N(0, \sigma^2_{error})$

The errors ϵ_{ij} with variance σ^2_{error} represent the gene-specific "within-subject-variation", which may include DNAm measurement error, natural variation in DNAm between unique CpGs, and variation between subjects beyond what is explained by the stress level indicator. The random effect for an individual η_i with variance $\sigma^2_{individual}$ represents variation between individuals, beyond what is explained by gene-level DNAm.

Prior genomic studies have confirmed that DNAm can be allele-specific,[35, 100-102] and thus, independent tests of association should account for SNPs within a pre-defined region of the CpG site. A standard practice in array-based DNA methylation studies is to remove any CpG from the analysis that may be affected by the presence of a SNP within or near the probe interrogating the CpG of interest. As mentioned previously, the current study filtered out CpGs based on Illumina annotation of a SNP being within 10 base pairs of the CpG of interest. We additionally flagged a larger set of CpGs that might be affected by SNPs based on a recent study by Liu et al. using 450k data, which identified 97,658 CpG-SNP pairings where methylation was significantly associated with genotype.[187]

Supplementary models were examined to ensure that we were not missing potential effects that were being masked by various model specifications: 1) initial models stratified by sex, 2) initial models stratified race, 3) initial models stratified ever/never obese, 4) using a restricted set of the core covariates (no longer adjusting for baseline BMI) and 5) creating a composite score for stress for those with exposure to high levels of stress in at least two of the three psychosocial variables.

For genes associated with any one of the psychosocial variables, we assessed whether genetic expression is also associated with DNAm. Cross-sectional associations between DNA methylation and gene expression levels were assessed at the end of the study follow-up period (exam 5). We fit a single model for each gene transcript, which treated transcript level as the dependent variable, adjusting for a core set of covariates (age, sex, race/ethnicity, and enrichment scores for neutrophils, B cells, T cells and natural killer cells) as the predictors. These methods have been used previously in a similar study by Needham et al.[110]

All p-values for fixed effects in multilevel models were obtained by Kenward and Roger's methods for mixed model comparisons. [188] Despite minimizing all potential test that could have been performed, results were additionally adjusted for multiple testing using a non-conservative false discovery rate, as detailed by Benjamini and Hochberg.[189] All analyses were performed using R version (R v3.3.1 (2016-06-21)).[190] Several R packages were used for data pre-processing (*lumi*)[168] and multilevel modeling (*lme4, pbkrtest*).[191, 192] Flowchart of DNAm QC and CpG filtering process was created using the Google drive web tool, draw.io diagrams.

4.3 Results

Of the 1,264 participants in the methylation study, 284 were removed due to missing data on study covariates, leaving a final sample size of N=1,180. There were a total of 2,528 CpGs within 87 unique obesity genes available for analysis after performing quality control measures and CpG reduction. The median number of CpGs per gene was 19 (mean~29), ranging from as low as 3 sites to 395 sites per gene (Table 1). There was good coverage of promoter regions, as most genes had at least a third of all CpGs located within the promoter.

Demographic information can be found in Table 2. The MESA participants are an older population of adults, with a mean age of 59.9 years at baseline. A majority of the participants were female (51.5%) and identified as non-Hispanic White (48.0%). The average BMI level in the overall population was just towards the high end of the overweight range at 29.2 kg/m^2. An overwhelming majority were non-smokers (86.5%) and tended to be educated, with only 13.8% not having received at least a high school degree.

Psychosocial variables are presented either categorically or dichotomously as detailed in the methods section. Approximately 38.7% of the cohort identified as having a high chronic burden of stress, with another ~30% claiming to have at least some level of moderate (medium) stress. The CSS was associated with age, sex, baseline BMI, and adjusted income in bivariate analyses (p-values <0.001), but not with race, smoking status or education level (p-value >0.05). For the everyday hassles measure, 54.2% reported a medium level, while 23.6% reported experiencing high levels. This measure was associated with age, race/ethnicity, smoking status, adjusted income, and categorical education (p-values <0.05). BMI level and sex distribution did not significantly differ across categories of EHS (Table 1). Cohen's perceived stress scale indicated that roughly 40% of the population perceived their lives to be stressful. Education was significantly associated with the PSS-4 (p-value=0.001). Also, a significant amount of women and those with higher income reported higher levels of perceived stress (p-values <0.001). No other covariates significantly varied by PSS-4 level.

Hierarchical models for gene-level DNAm analysis

For multilevel models adjusting for core covariates, we found two genes (*C9orf93* and *FIGN*) that were significant on some level of the CSS psychosocial factor at an uncorrected alpha threshold of 0.05. For the gene *C9orf93*, those with high CSS had a higher level of methylation compared to those with low CSS (beta= 0.055, p-value=0.008). Additionally, we observed an interaction between high CSS and sites within the promoter region of *C9orf93* showing lower levels of methylation compared to the referent population (beta= -0.064, p-value=0.006). For *FIGN*, the only potentially significant finding was with the interaction between high CSS and sites within the shore/shelf region (beta=0.133, p-value=0.004). CSS comparisons were also robust to additional adjustments for physical activity and caloric intake (not shown).

We recognized the potential for CpGs to vary within a gene, depending on the type of site (i.e. promoter or shore/shelf),[193] and therefore conducted a sensitivity analysis for potentially significant results by including a random slope for specific site types. Results did not change

significantly for either *C9orf93* or *FIGN* in these sensitivity models. Most notably, however, none of these findings was significant after correcting for multiple comparisons using FDR, indicating potentially spurious results. No marginally significant effects were observed for the EHS or PSS-4 measures.

Supplementary analyses

In order to expand on the findings with the CSS score, we attempted to use several different models to determine whether or not there was a true effect of CSS on DNAm levels. The supplementary models included a restricted set of covariates (no longer adjusting for baseline BMI), a composite score for stress for those with exposure to high levels of stress in at least two of the three psychosocial variables, and subsequently, models stratified by sex, race, or ever/never obese. The stratified models were employed primarily to account for a differential effect of the exposures on DNAm by race/ethnicity categories and sex. Models that stratified for ever/never obese were to account for the potential feedback loop of obesity causing DNAm of obesity risk genes and not in the assumed temporal ordering of psychosocial factor to stress measure to DNAm.

These supplementary findings showed minimal improvement on the genes that were significantly associated any of the psychosocial factors and DNAm, with only *FIGN* and *C9orf39* still coming up in various comparisons. *FIGN* was significant for the interaction between CSS-high and shore/shelf sites only in non-Hispanic White participants. Additionally, we observed an interaction between CSS-high and promoter sites only in obese subjects. Similar to the main *FIGN* results, at the uncorrected alpha 0.05 level, the composite score of stress that incorporated EHS and PSS-4 showed that higher levels of stress interacted with shore/shelf sites within *FIGN*. In stratified analyses, this finding persisted only in obese and non-Hispanic White subjects.

For *C9orf93*, the original finding of an interaction between high CSS and promoter sites was significant only in subjects who were not obese at baseline or during follow-up. However, once again, none of these supplementary results passed multiple testing corrections, indicating that earlier results were likely to be false positives. Further, neither *C9orf93* nor *FIGN* had expression transcripts available to bolster confidence by validating the original findings.

Site-specific tests

The motivation for using multi-level models was to maximize statistical power, by reducing the total number of hypothesis tests performed. However, given the minimal findings at the gene-level, a supplementary analysis looked at all CpG sites individually to account for the possibility that a large number of non-significant CpGs were masking some potentially strong effects. This analysis was done by performing linear regression of each psychosocial variable by DNAm at each CpG, totaling 12,640 tests. The top ten findings, as presented in table 4, show two unique CpGs that reached statistical significance at the FDR corrected threshold of 0.20 (*NAV1*:cg17753974 comparing CSS-med to CSS-low and *SMG6*:cg17648080 comparing CSS-high to CSS-low). The

FDR corrected q-values for *NAV1*:cg17753974 and *SMG6*:cg17648080 were 0.15 and 0.17, respectively.

The effect in *NAV1*:cg17753974 remained unchanged in a supplementary model that additionally adjusted for dietary factors of saturated, monounsaturated, and polyunsaturated fats (N=1,160; beta= -0.048, unadjusted p-value=2.07e-05). However, when adding in lifestyle factors such as exercise (measure in total metabolic equivalent of task (MET) minutes per week), the effect became slightly attenuated (beta= -0.043, unadjusted p-value=0.002). Further, examining *NAV1*:cg17753974 in CSS-high, we observed a similar result to CSS-med that simply did not pass FDR correction (beta= -0.03, p-value=0.015, q-value= 0.998).

SMG6:cg17648080 also remained relatively unchanged when additionally adjusting for dietary factors of saturated, monounsaturated, and polyunsaturated fats (N=1,160; beta= -0.097, p-value=2.55e-05). Similarly, the effect became slightly attenuated when adding exercise to the linear model (beta= -0.088, p-value=0.002).

In order to determine the functional relevance of the *NAV1* finding, we examined the association between DNAm and genetic expression, utilizing one available transcript for the *NAV1* gene. We found that DNAm was significantly associated with expression in the model adjusting for the core covariates of age, sex, baseline BMI, race/ethnicity, and the four major cell type compositions (probeID: ILMN_1703374; estimate > 0, p-value=0.008). Expression data were not available for *SMG6*.

4.4 Discussion

The appeal of epigenetic studies is in discovering the impact of non-genetic factors on how genes may be expressed in the body and cause negative health conditions. The multifactorial nature of obesity risk made it a prime outcome for discovering sites that might be affected by environmental or social factors, including psychosocial stress. To our knowledge, this study is the first of it's kind to assess multiple measures of psychosocial stress on DNA methylation of the most comprehensive set of GWAS-derived obesity risk genes in a large multi-ethnic cohort. Using robust methods to interrogate gene-level DNAm, we were also able to incorporate CpG level information to determine if there were interactions between psychosocial measures and sites within gene promoter regions or CpG island shores and shelves. This study also benefited from the availability of genetic expression data.

We found mixed evidence to support the relationship between psychosocial stress and DNAm. Our analysis found that only one of the three psychosocial stress measures, chronic burden of stress, appeared to have a modest effect on methylation at the gene-level and at individual CpGs. Specifically, we found that chronic burden of stress measure appeared to have some effect at the gene-level for *FIGN* and *C9orf93*. The positive direction of effect in *FIGN* indicated that subjects with higher levels of CSS had greater DNAm. Conversely, in *C9orf93* we found that subjects with higher levels of CSS had lower levels of DNAm, and an apparent interaction between CSS-high and promoter CpGs. While the results appeared promising, after correcting for multiple comparisons the effects were no longer statistically significant. While these genes are derived from

recent GWAS and meta-analyses for obesity, we found minimal information about the functional importance of *FIGN* and *C9orf93* with respect to their role in obesity or psychosocial stress.

FIGN, or Figetin, Microtubule Severing Factor, is know to be involved in cellular and developmental processes, including embryogenesis,[194] and is a member of the AAA family, functioning as a chaperone to ATPase.[195] Beyond that, its function with respect to obesity remains unknown. There is relatively little information about how *C9orf93,* also known as *CCDC171*, or coiled-coil domain containing 171,[196] might play a role in obesity directly or be affected by stressors. Given the weak associations observed with both of these genes, there may be little reason to move forward with them as candidates for identifying the effect of psychosocial stress on DNAm.

A similar study conducted recently by Needham and colleagues examined DNAm for genes in stress and inflammatory pathways, some of which had been previously associated with psychosocial factors.[110] Their study showed that trajectories of socioeconomic status (defined by educational attainment) from childhood into adulthood were significantly associated with gene-level DNAm and consequent expression in *CD1D*, *F8*, *FKBP5*, *KLRG1*, and *NLRP12*.[110] Needham et al. presented a strong case for a relationship between social factors and DNAm for genes directly related to the stress response, and posited psychosocial factors as a potential mechanism by which the stress response could lead to altered DNAm levels.

Findings from Needham et al. showing the strength of using social measures on DNAm are supported by the results of another recent study of psychosocial stress and DNAm of genes related to the stress-response system.[197] The study, conducted by Unternaehrer et al., examined the role of the Trier Social Stress Test (TSST) on DNAm of the *OXTR* and *BDNF* genes. While a significant effect was reported for *OXTR* methylation, no effect was observed in *BDNF*, an obesity-associated gene that was also not significantly associated with psychosocial factors in the current study. Unternaehrer and colleagues discussed previous examples where changes in *BDNF* methylation were associated with childhood adversity.[198] This indicated that perhaps methylation in the *BDNF* gene is more susceptible to long-term effects of stress experienced in early life. This would also support the finding from Needham et al., which showed a significant effect of childhood measures of SES and SES trajectories on DNAm in adulthood. Overall, the mixed findings from the current study are likely indicative of two possibilities: 1) Long-term measures of psychosocial factors that incorporate childhood exposures are more relevant for DNAm changes in adulthood, or 2) the effect of psychosocial factors on gene-level methylation of obesity risk genes may simply be weaker.

The greatest effect estimate from site-specific tests was observed in comparisons of the chronic burden of stress variable with methylation of the Neuron Navigator 1 gene (*NAV1*). *NAV1*:cg17753974 is located on chromosome 1q32.1, within the gene body, and is a Shore/Shelf site (South Shelf) not associated with any SNPs thought to influence DNAm levels.[187] *NAV1* is a protein-coding gene that is expressed primarily in the nervous system, as well as the heart, and thought to be involved in neuronal development and regeneration.[199, 200] This gene is located in a region that contains a cluster of genes (*MYF4*, *FMOD*, *REN*, *PMCA4*, *TNNI1*) thought to be associated with heart function[199] and was found to be over-expressed in cardiovascular tissue in mice.[201] To our knowledge, there is currently no information on how the gene functions with

respect to psychosocial stressors or obesity. However, its expression in cardiovascular tissue may indicate a more complex mechanism by which the stress response impacts cardio-metabolic systems to contribute to overall obesity. The *NAV1* gene appears to be a noteworthy candidate for exploration of epigenetic influence on cardio-metabolic activity.

Prior studies have shown that DNAm from single CpGs can be significantly correlated with genetic expression levels.[202] Over-expression of genetic risk factors examined for obesity or increased BMI in the current study would indicate an increased likelihood for obesity at a later time point. However, it is difficult to say with certainty what functional role DNAm plays in blocking transcription of a gene not located within the promoter region, as this is biologically what is assumed to be most directly responsible for the association.[81] Although, a recent study by Wagner and colleagues found that CpG location was less relevant for the sign of the correlation with gene expression compared to chromatin state, highlighting that the complex mechanisms by which DNAm can affect expression are still not well understood.[202]

Results from the current study revealed *hypo*methylation of the *NAV1* site comparing those with medium level of chronic burden of stress to those with a low level. While this site was not located within the promoter region of the gene, we observed a significant effect of DNAm at this site with genetic expression of the *NAV1* transcript, ILMN_1703374. The direction of the effect was in the expected inverse direction, with *hypo*methylation leading to increased expression, as the ILMN_1703374 regression beta value was significantly above 0 at a p-value of 0.002. The relationship between psychosocial factors and DNAm of obesity genes is only meaningful if it translates to an increased risk of obesity at a later time point. Given that DNAm was assessed at the end of the follow-up period, it is difficult to say whether or not *hypo*methylation of *NAV1* would lead to more cases of obesity or vice versa (potential reverse-causation). Based on our results, we have determined that the association between the chronic burden of stress and DNAm warrants further investigation. Future studies would benefit from examining the mediatory role of *NAV1* methylation in the pathway from chronic burden of stress to obesity at a later time point.

Strengths and limitations

The current study was strengthened by several key features. To our knowledge, this was the largest study of its kind to assess the impact of multiple psychosocial factors with DNA methylation of 87 obesity risk genes. Further, the use of a large multi-ethnic population improves upon generalizability of results compared to other studies that are restricted to small homogenous populations. Using high throughput array based technology for DNAm (specifically, with 450k data) was also a major strength allowing us to obtain expansive genomic coverage of obesity risk genes. We were able to supplement evidence of DNAm differences with expression analyses, despite relatively small effects for our most significant DNAm findings. Lastly, we were able to adjust for cell composition of four major white blood cell types found in peripheral blood, which is an improvement upon studies that fail to do so. This reduces a major source of criticism of epigenetic studies that use DNA from peripheral blood, as DNAm is highly cell type and tissue specific with respect to methylation levels.[203]

A few limitations warrant comment. When correcting for epigenetic "noise," it was necessary to impose multiple quality control normalization techniques, which can hinder the ability to detect significant differences. Though not available in the current study, supplementing 450k findings with other methylation technologies (e.g. pyrosequencing) would help validate our results, although the current study benefitted from the availability of genetic expression data, validating the CpG site-specific results.

The psychosocial factors examined were intended to assess multiple forms of stress, including objective (CSS), perceived (PSS-4), and discriminatory measures (EHS). While the internal reliability of the EHS measure was quite good, the same could not be said of the PSS-4. As mentioned in the discussion section, the use of stress measures that are not accurately capturing some latent construct of more general stress limits our interpretation of results. The PSS-4 is generally easier to collect as the questions can be administered via telephone, but future studies should consider using a version of the PSS that includes more components so as to strengthen the consistency of the measure.

Another possible explanation for the minimal findings in the PSS-4 is that it was assessed cross-sectionally with DNAm at MESA Exam 5. Given that we observed several cases of the impact chronic measures of stress on DNAm, it is possible that there was not enough time for PSS measures to have an effect on DNAm. While there isn't a clear consensus on the exact timeframe required for stressful life experiences to affect DNAm changes, earlier studies showing effects of longer-term exposures indicate that cross-sectional assessment of stress and methylation may not detect significant differences.

We elected to employ a more rigorous approach for our CpG filtering criteria, primarily to ensure that methylation changes were reflective of responses to stress and not merely due to the presence of a SNP within the Illumina probe. This approach is common in DNAm studies that do not have the benefit of adjusting for genotypic variation among subjects. The current study was also limited in this respect. Future studies that can incorporate genetic data to account for allele-specific methylation changes would strengthen statistically significant findings.

Genetic data would also allow for the inclusion of genetic ancestry markers to adjust for variation that might be explained by population stratification. While we were to able to control for self-reported measures of race/ethnicity, recent studies have shown mixed results with respect to how well self-reported race matches populations of homogenous ancestry. There have been reports of a strong correlation between self-report of being non-Hispanic White and European ancestry,[204] but weaker evidence of clustering of self-reported race and genetic ancestry for African-Americans[204, 205] and Asian-Americans.[206] The possibility of residual confounding by population stratification remains a concern that should be addressed in future studies.

As mentioned previously, DNAm is highly cell-type and tissue specific. A potential limitation in the current study is the use of DNA from peripheral blood leukocytes in the first place. There are examples of how DNAm is significantly altered in adipose tissue after intervening on factors related to stress, such as physical activity,[120] indicating that there may be biological systems that are more reflective of DNAm changes in the short term, depending on the tissue examined. Ideally, studies should have tissues specific to the disease available for DNAm analysis, but in the context of a large population-based cohort study, this may not be economical or practically feasible.

It is important to note that the current study was not well powered to detect small effects. Consequently, we did not employ a more stringent cutoff for FDR significance, consistent with recent DNAm studies.[110] The FDR q-value of 0.2 ensured that there was less than a 20% chance this discovery was, in fact, a false positive.[207] This cutoff was used to prevent limiting the study from discovering sites that may be of interest for future research. However, a more stringent cutoff would further minimize the concern of reporting a false positive finding.

The current study aimed to investigate the relationship between psychosocial factors and DNAm of obesity risk genes. We were able to show that a specific CpG within *NAV1* was significantly lower in DNAm for those with some level of chronic burden of stress, compared to those with no stress, and this result was associated with a higher level of expression in MESA participants. While we observed minimal evidence of small effects at the gene-level for *FIGN* and *C9orf93*, these findings ultimately did not withstand FDR correction, indicating that they were likely false positives. As prior studies have established significant effects of social factors on genes from various other pathways, more research is warranted. Future studies would benefit from further exploration of the full network of genes that might be affected by psychosocial factors, and if the sites within *NAV1* can be replicated in independent populations.

4.5 Tables and Figures

Table 1. Features of CpG sites in obesity risk genes (N=2,528 CpGs)

Gene	# of CpG sites	# In promoter region (%)	# In shore/shelf region (%)	# associated with SNPs (%)	Expression data available
ADCY3	29	11 (37.9)	9 (31.0)	2 (6.90)	Yes
AGBL4	52	27 (51.9)	17 (32.7)	1 (1.90)	No
ASB4	17	8 (47.1)	3 (17.6)	0 (0.00)	No
ATP2A1	27	12 (44.4)	16 (59.3)	0 (0.00)	No
BCDIN3D	14	12 (85.7)	8 (57.1)	0 (0.00)	Yes
BDNF	74	59 (79.7)	38 (51.4)	0 (0.00)	No
C18orf8	10	8 (80.0)	2 (20.0)	0 (0.00)	No
C6orf106	21	9 (42.9)	11 (52.4)	0 (0.00)	No
C9orf93	7	6 (85.7)	2 (28.6)	0 (0.00)	No
CADM1	47	13 (27.7)	10 (21.3)	0 (0.00)	No
CADM2	27	9 (33.3)	12 (44.4)	0 (0.00)	No
CALCR	34	29 (85.3)	18 (52.9)	0 (0.00)	No
CBLN1	15	9 (60.0)	4 (26.7)	0 (0.00)	No
CLIP1	25	7 (28.0)	7 (28.0)	0 (0.00)	Yes
CREB1	16	9 (56.2)	4 (25.0)	0 (0.00)	Yes
DMXL2	12	7 (58.3)	3 (25.0)	1 (8.30)	Yes
EHBP1	35	17 (48.6)	11 (31.4)	0 (0.00)	Yes
ELAVL4	23	20 (87.0)	2 (8.70)	0 (0.00)	No
EPB41L4B	12	3 (25.0)	6 (50.0)	0 (0.00)	Yes
ERBB4	33	6 (18.2)	11 (33.3)	1 (3.00)	No
ETS2	16	7 (43.8)	7 (43.8)	0 (0.00)	Yes

ETV5	28	17 (60.7)	10 (35.7)	0 (0.00)	No
FHIT	56	12 (21.4)	10 (17.9)	0 (0.00)	Yes
FIGN	22	10 (45.5)	7 (31.8)	0 (0.00)	No
FOXO3	42	16 (38.1)	6 (14.3)	0 (0.00)	Yes
FPGT	14	13 (92.9)	10 (71.4)	0 (0.00)	No
FTO	36	14 (38.9)	4 (11.1)	0 (0.00)	Yes
FUBP1	15	7 (46.7)	8 (53.3)	0 (0.00)	Yes
GALNT10	40	7 (17.5)	11 (27.5)	1 (2.50)	Yes
GBE1	17	11 (64.7)	9 (52.9)	0 (0.00)	Yes
GNAT2	7	6 (85.7)	0 (0.00)	0 (0.00)	No
GNPDA2	10	7 (70.0)	5 (50.0)	0 (0.00)	Yes
GPRC5B	20	8 (40.0)	3 (15.0)	0 (0.00)	Yes
GRID1	59	21 (35.6)	15 (25.4)	0 (0.00)	No
GRP	16	13 (81.2)	7 (43.8)	0 (0.00)	No
HHIP	18	15 (83.3)	9 (50.0)	0 (0.00)	No
HIF1AN	11	10 (90.9)	0 (0.00)	0 (0.00)	Yes
HIP1	42	13 (31.0)	13 (31.0)	0 (0.00)	Yes
HNF4G	3	1 (33.3)	0 (0.00)	0 (0.00)	No
HSD17B12	19	11 (57.9)	5 (26.3)	0 (0.00)	No
IFNGR1	12	8 (66.7)	6 (50.0)	0 (0.00)	Yes
INO80E	18	14 (77.8)	3 (16.7)	0 (0.00)	Yes
KCNK3	29	5 (17.2)	13 (44.8)	0 (0.00)	Yes
KCTD15	14	6 (42.9)	4 (28.6)	0 (0.00)	Yes
LINGO2	3	0 (0.00)	0 (0.00)	0 (0.00)	No
LMX1B	46	8 (17.4)	19 (41.3)	0 (0.00)	No
LRP1B	37	9 (24.3)	11 (29.7)	0 (0.00)	No

MAP2K5	23	8 (34.8)	4 (17.4)	1 (4.30)	Yes
MC4R	6	5 (83.3)	0 (0.00)	0 (0.00)	No
MIR548A2	10	0 (0.00)	0 (0.00)	0 (0.00)	No
MTCH2	13	8 (61.5)	6 (46.2)	0 (0.00)	Yes
MTIF3	26	22 (84.6)	4 (15.4)	0 (0.00)	Yes
NAV1	67	29 (43.3)	12 (17.9)	0 (0.00)	Yes
NEGR1	22	13 (59.1)	4 (18.2)	0 (0.00)	No
NLRC3	24	7 (29.2)	5 (20.8)	0 (0.00)	No
NRXN3	67	13 (19.4)	14 (20.9)	0 (0.00)	No
NT5C2	24	10 (41.7)	5 (20.8)	2 (8.30)	Yes
OLFM4	8	6 (75.0)	0 (0.00)	0 (0.00)	No
PARK2	75	18 (24.0)	15 (20.0)	3 (4.00)	No
PGPEP1	16	12 (75.0)	8 (50.0)	0 (0.00)	No
PRKD1	20	10 (50.0)	4 (20.0)	0 (0.00)	No
PTBP2	13	10 (76.9)	6 (46.2)	0 (0.00)	Yes
QPCTL	19	16 (84.2)	6 (31.6)	0 (0.00)	Yes
RABEP1	15	7 (46.7)	7 (46.7)	0 (0.00)	Yes
RALYL	31	21 (67.7)	12 (38.7)	0 (0.00)	No
RARB	28	17 (60.7)	0 (0.00)	1 (3.60)	No
RASA2	19	8 (42.1)	4 (21.1)	0 (0.00)	Yes
RPTOR	395	11 (2.80)	214 (54.2)	15 (3.80)	Yes
SBK1	19	6 (31.6)	11 (57.9)	1 (5.30)	Yes
SCARB2	18	12 (66.7)	5 (27.8)	0 (0.00)	Yes
SEC16B	9	4 (44.4)	0 (0.00)	0 (0.00)	No
SLC39A8	23	20 (87.0)	4 (17.4)	0 (0.00)	Yes
SMG6	44	13 (29.5)	8 (18.2)	0 (0.00)	No

STXBP6	17	4 (23.5)	2 (11.8)	0 (0.00)	No
TAL1	23	0 (0.00)	13 (56.5)	0 (0.00)	No
TCF7L2	79	11 (13.9)	24 (30.4)	0 (0.00)	Yes
TDRG1	9	8 (88.9)	0 (0.00)	0 (0.00)	No
TFAP2B	43	8 (18.6)	24 (55.8)	0 (0.00)	No
TLR4	3	2 (66.7)	0 (0.00)	0 (0.00)	Yes
TMEM18	16	10 (62.5)	6 (37.5)	1 (6.20)	Yes
TOMM40	12	10 (83.3)	3 (25.0)	0 (0.00)	Yes
TRIM66	10	7 (70.0)	0 (0.00)	0 (0.00)	Yes
UBE2E3	14	7 (50.0)	4 (28.6)	0 (0.00)	Yes
USP37	16	12 (75.0)	2 (12.5)	0 (0.00)	Yes
ZBTB10	14	6 (42.9)	7 (50.0)	0 (0.00)	No
ZC3H4	16	8 (50.0)	5 (31.2)	0 (0.00)	Yes
ZFP64	42	14 (33.3)	18 (42.9)	0 (0.00)	No

CpGs included in this table are based on the number of total CpGs after quality control filtering was employed. The column labeled “Expression data available” indicates that there is at least one transcript within the gene, available for functional analysis. The column for “# of CpGs associated with SNPs” is based on associations detailed by Liu et al.

Table 2. Demographic information on MESA participants (N=1,180)

		Chronic Burden of Stress Score			Everyday Hassles Score			Cohen's PSS-4	
	Overall	**Low**	**Med**	**High**	**Low**	**Med**	**High**	**Low**	**High**
N	1,180	372 (31.5)	351 (29.7)	457 (38.7)	261 (22.1)	639 (54.2)	280 (23.7)	708 (60.0)	472 (40.0)
Age Mean (SD)	59.9 (9.49)	61.5 (9.46)	61.0 (9.40)	57.8 (9.22)	63.1(9.69)	60.4 (9.27)	55.7 (8.34)	60.3 (9.30)	59.4 (9.77)
				p<0.001			**p<0.001**		p=0.119
Sex									
Female	608 (51.5)	151 (12.8)	180 (15.3)	277 (23.5)	124 (10.5)	346 (29.3)	138 (11.7)	334 (28.3)	274 (23.2)
Male	572 (48.5)	221 (18.7)	171 (14.5)	180 (15.3)	137 (11.6)	293 (24.8)	142 (12.0)	374 (31.7)	198 (16.8)
				p<0.001			p=0.135		**p<0.001**
Race/ethnicity									
Non-Hispanic White	566 (48.0)	190 (16.1)	164 (13.9)	212 (18.0)	104 (8.81)	352 (29.8)	110 (9.32)	346 (29.3)	220 (18.6)
African-American	237 (20.1)	64 (5.42)	73 (6.19)	100 (8.47)	31 (2.63)	129 (10.9)	77 (6.52)	146 (12.4)	91 (7.71)
Hispanic	377 (31.9)	118 (10.0)	114 (9.66)	145 (12.3)	126 (10.7)	158 (13.4)	93 (7.88)	216 (18.3)	161 (13.6)
				p=0.477			**p<0.001**		p=0.426
BMI (kg/m^2) Mean (SD)	29.2 (5.26)	28.4 (4.71)	29.4 (5.32)	29.8 (5.55)	28.8 (4.80)	29.4 (5.27)	29.4 (5.62)	29.2 (5.20)	29.3 (5.35)
				p<0.001			p=0.352		p=0.947
Smoking status									
No	1035 (86.5)	320 (27.1)	313 (26.5)	387 (32.8)	237 (20.1)	551 (46.7)	232 (19.7)	616 (52.2)	404 (34.2)
Yes	162 (13.5)	52 (4.41)	38 (3.22)	70 (5.93)	24 (2.03)	88 (7.46)	48 (4.07)	92 (7.80)	68 (5.76)
				p=0.174			**p=0.026**		p=0.544
Income Median (SD) [a]	2.25 (1.89)	2.25 (2.05)	2.25 (1.87)	2.08 (1.74)	1.80 (1.66)	2.25 (2.00)	2.22 (1.78)	2.25 (1.99)	1.88 (1.68)
				p<0.001			**p<0.001**		**p<0.001**
Education [b]									
Category 1	163 (13.8)	57 (4.83)	45 (3.81)	61 (5.17)	69 (5.85)	75 (6.36)	19 (1.61)	81 (6.86)	82 (6.95)
Category 2	432 (36.6)	118 (10.0)	141 (11.9)	173 (14.7)	95 (8.05)	238 (20.2)	99 (8.39)	257 (21.8)	175 (14.8)
Category 3	369 (31.3)	118 (10.0)	110 (9.32)	141 (11.9)	57 (4.83)	214 (18.1)	98 (8.31)	220 (18.6)	149 (12.6)
Category 4	216 (18.3)	79 (6.69)	55 (4.66)	82 (6.95)	40 (3.39)	112 (9.49)	64 (5.42)	150 (12.7)	66 (5.59)
				p=0.243			**p<0.001**		**p=0.001**

Statistical tests for continuous covariates were done using a Student's T-test for PSS-4 and ANOVA for CSS and EHS measures. Categorical covariates were tested using a standard Chi-squared test. P-values less than 0.05 are in **bold**.

[a] Income is a continuous measure calculated by per capita adjusted household income per $10,000 (continuous income/# of people supported)/10,000).

[b] Categories 1-4 for the education variable are coded as follows: 1) Less than high school diploma; 2) High school graduate or some college, but no degree; 3) Technical/associate/bachelor degree; 4) Post college training.

Table 3. Multivariate regression of M-value on psychosocial factors in obesity risk genes (N=1,197)

	C9orf93				*FIGN*			
	Estimate	**SE**	***P*-val**	***q*-val**	**Estimate**	**SE**	***P*-val**	***q*-val**
CSS med	0.026	0.040	0.107	0.999	-0.031	0.027	0.247	0.999
CSS high	0.036	0.022	**0.008**	0.999	-0.016	0.025	0.521	0.999
Promoter	-5.596	0.017	0.000	-	-5.120	0.032	0.000	-
Shore/Shelf	-0.027	0.013	0.045	-	-0.634	0.035	0.000	-
CSS med * Promoter	-0.038	0.025	0.127	0.999	0.001	0.047	0.984	0.999
CSS high * Promoter	-0.064	0.023	**0.006**	0.999	-0.048	0.044	0.271	0.999
CSS med * Shore/Shelf	0.007	0.019	0.697	0.999	0.058	0.050	0.246	0.999
CSS high * Shore/Shelf	1.54e-03	0.018	0.931	0.999	0.133	0.047	**0.004**	0.999

This tables reports DNAm findings for the chronic burden of stress scale measure (CSS) and potentially significant genes at the unadjusted alpha 0.05 level; these values are marked in bold font. The referent group for comparisons is study participants with low CSS. Results are derived from a two-level model treating the CpG as the level-1 unit and the individual as level-2 unit, adjusting for core covariates: categorical education and continuous per capita adjusted household income, sex, smoking status, race/ethnicity, and enrichment score for the four major blood cell types (B cells, T cells, NK cells, and neutrophils).

Table 4. Site-specific linear regression of M-Value on psychosocial factors in obesity risk genes (N=1,180)

Psychosocial Factor (ref: Low)	Gene	TargetID	Chr	Estimate	*P*-val	q-val	Promoter region	Shore/Shelf region	SNP associated	Expression Data
CSS med	*NAV1*	cg17753974	1	**-0.048**	**1.02e-05**	**0.152**	No	Yes	No	Yes
CSS high	*SMG6*	cg17648080	17	**-0.096**	**2.68e-05**	**0.169**	Yes	No	No	No
PSS-4 high	*CALCR*	cg05284750	6	-0.066	1.34e-04	0.405	Yes	Yes	No	No
EHS med	*GBE1*	cg14465376	9	-0.060	1.60e-04	0.405	No	No	No	Yes
PSS-4 high	*NAV1*	cg14780255	1	-0.039	1.51e-04	0.405	No	No	No	Yes
EHS med	*HIP1*	cg02713883	10	-0.041	4.46e-04	0.719	No	Yes	No	Yes
CSS med	*TCF7L2*	cg03089923	3	-0.054	5.34e-04	0.719	No	No	No	Yes
EHS med	*USP37*	cg05867885		-0.058	5.69e-04	0.719	No	Yes	No	No
PSS-4 high	*FOXO3*	cg08345465	7	-0.035	4.96e-04	0.719	No	No	No	Yes
CSS med	*NAV1*	cg23236366	1	-0.043	4.32e-04	0.719	Yes	No	No	Yes

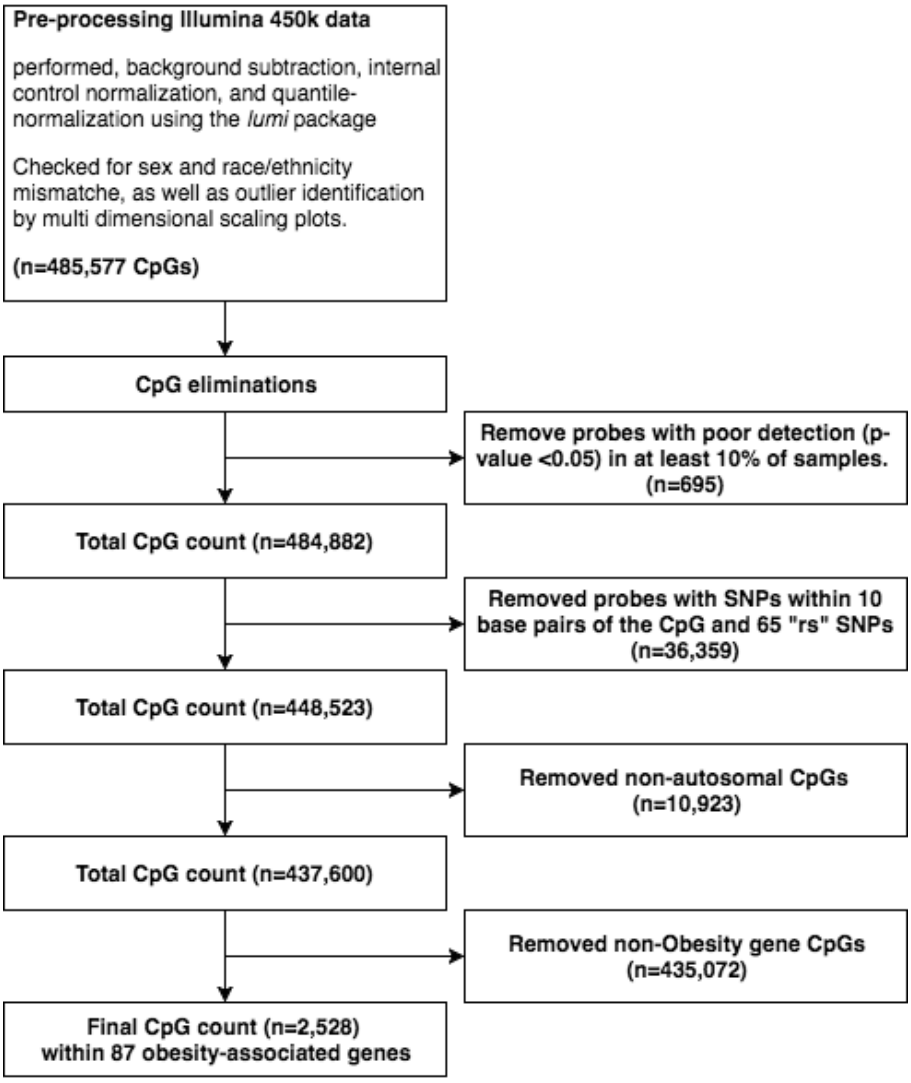
CpG site specific tests were conducted using linear regression, adjusting for core covariates: categorical education and continuous per capita adjusted household income, sex, smoking status, race/ethnicity, and enrichment score for the four major blood cell types (B cells, T cells, NK cells, and neutrophils). Results in **bold** font indicate that the test reached and FDR significance threshold of 0.20, correcting for 12,640 tests. The referent group is the low level of the psychosocial factor, for any of the three exposures examined. CSS represents the chronic burden of stress scale measure, categorized by low, medium and high stress. EHS represents the everyday hassles scale, categorized by low, medium and high discrimination. PSS-4 represents Cohen's perceived stress scale, categorized by low and high, perceived stress.

Table S1. Detailed questionnaire information for psychosocial variables

Chronic Burden of Stress Scale (CSS) Response options: ___No ___Yes	**Ongoing problems lasting greater than 6 months in one of the following domains:** 1) Personal Health 2) Health of a loved one 3) Job 4) Relationship 5) Financial
Everyday Hassles Scale (EHS) Response options: 1. Four or more times 2. Two or three times 3. Once 4. Never	6) You have been treated with less courtesy than other people 7) You have been treated with less respect than other people 8) You have received poorer service than other people at restaurants or stores 9) People have acted as if they think you are not smart 10) People have acted as if they are afraid of you 11) People have acted as if they think you are dishonest 12) People have acted as if they're better than you are 13) You have been called names or insulted 14) You have been threatened or harassed
Cohen's Perceived Stress Scale (PSS-4) Response options: ___0=never ___1=almost never ___2=sometimes ___3=fairly often ___4=very often	1) In the last month, how often have you felt that you were unable to control the important things in your life? 2) In the last month, how often have you felt confident about your ability to handle your personal problems? [a] 3) In the last month, how often have you felt that things were going your way? [a] 4) In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?

[a] Positive items were reverse coded when calculating the total values for PSS-4.

Figure S1. Flowchart of DNAm QC procedures and CpG filtering criteria at multiple stages



Chapter 5

Conclusion

5.1 Summary of key findings

Obesity is a medical condition affecting approximately 35% of all US adults. It is characterized by increased abdominal adiposity and associated with myriad negative health outcomes, including CVD and various cancers (e.g. breast, colon, and endometrial). Obesity has also been found to be an independent risk factor for overall mortality. Not only is obesity a significant public health concern in the broader population, but also, disparities in the prevalence of obesity by sex and race/ethnicity lead to disproportionate burdens placed on various sub-populations. For example, the age-adjusted prevalence of obesity among non-Hispanic Black women is approximately 58%, while in non-Hispanic White women, it is only approximately 32%. Obesity is a multifactorial condition and, as a result, it must be examined with respect to multiple pathways that incorporate biological, behavioral, and environmental risk factors. This dissertation aimed to address limitations of previous studies by examining the association between psychosocial and genetic risk factors and obesity, as well as psychosocial factors and DNA methylation of obesity risk genes.

Conducting a systematic review of publicly available studies, I found that there were few examples of gene-environment interaction studies (N=4) examining the association between psychosocial factors and genetic risk factors and obesity. These studies found the following statistically significant associations between various measures of psychosocial factors, genetic factors and obesity: 1) chronic burden of stress score and SNPs in *EBF1* with hip circumference, 2) chronically stressed caregiver status and SNPs in *APOE* with waist circumference, 3) cumulative stress score and SNPs in *APOE* with waist circumference and BMI, and 4) receipt of public assistance and the *MAOA* short allele with BMI. There were also few epigenetic studies (N=5 studies) that investigated the relationship between psychosocial factors, obesity and DNA methylation. Four of these studies reported statistically significant associations, albeit with modest effects, between various obesity measures (e.g. BMI, body weight, waist circumference, and waist to hip ratio), psychosocial measures (e.g. depression), and DNA methylation of genes in inflammatory pathways (e.g. *IL6*) or genes involved in the biological stress-response (e.g. *SLC6A* and *HTR2A*). Results of the systematic review showed that both GxE and epigenetic studies were generally limited to small sample populations of non-Hispanic Whites. The limited number of studies, and the modest effects reported, indicated that more research is required and should include populations of varied race/ethnicity, so as to provide more generalizable findings.

Results from the gene-environment interaction study in chapter 3 established statistically significant main effects in the overall MESA population for an obesity genetic risk score and for the chronic burden of stress scale with obesity. I did not find statistically significant interactive effects, on either the multiplicative or additive scale, between any of the psychosocial factors examined (e.g. chronic burden of stress, everyday hassles, and depression) and an obesity genetic risk score and obesity. In supplementary analyses, stratified by race/ethnicity, I showed evidence of statistically significant sub-additive interaction between each of the psychosocial factors with the obesity genetic risk score and obesity in African-Americans. Furthermore, additive interaction was statistically significant in Hispanic-Americans, comparing the everyday hassles measure, obesity genetic risk score and obesity.

Chapter 4 presented results from a DNA methylation study in which I found no statistically significant relationship between any of three psychosocial factors (i.e. chronic burden of stress, everyday hassles, and Cohen's perceived stress scale) and gene-level DNA methylation of 87 BMI-associated genes. At the CpG level, I found statistically significant associations between the chronic burden of stress scale measure and DNA methylation of sites within the *NAV1* and *SMG6* genes. Using a single available transcript from genetic expression data, I showed that *NAV1* DNA methylation was statistically significantly inversely correlated with *NAV1* expression (i.e. higher *NAV1* DNA methylation was associated with decreased expression). Together, chapters 4 and 5 are to my knowledge, the first studies of their kind to examine multiple psychosocial factors and comprehensive obesity polygenic risk with obesity and DNA methylation of obesity risk genes in a large, multi-ethnic, longitudinal cohort study.

5.2 Conclusion and future directions

The effect of psychosocial factors on obesity is thought to operate primarily through mediating mechanisms, such as energy intake (diet) and expenditure (physical activity). However, animal models have consistently suggested that independent of these behavioral factors, there are underlying biological mechanisms that can lead to fat-storage based on a stress-response reaction to an external stimulus of stress. As these mechanisms have not been fully elucidated in humans, this dissertation aimed to identify potentially new model systems that incorporate psychosocial risk with obesity genetic risk to cause obesity. This was examined by conducting a gene-environment interaction study and a DNA methylation study in a large multi-ethnic longitudinal cohort. Both analyses showed specific examples of statistically significant relationships between psychosocial risk factors and genetic risk factors of obesity.

As emphasized throughout this dissertation, gene-environment interactions and epigenetic mechanisms allow for intervention on a modifiable risk factor like psychosocial stress, when intervening on inherited genetic sequences may not be practical or feasible. There remains a clear genetic component of obesity risk, and therefore, establishing new points of intervention is essential to preventing specific cases of obesity where onset only occurs in the presence of a psychosocial factor or when a gene is differentially expressed due to the effect of psychosocial stress on DNA methylation. Future research should expand on these studies in order to replicate or validate the current dissertation's findings, as well as to investigate biological pathways through

which interactions between psychosocial factors and genetic factors occur, in order to understand fully all mechanisms of the risk of obesity.

Bibliography

1. Eknoyan, G., *Adolphe Quetelet (1796-1874)--the average man and indices of obesity.* Nephrol Dial Transplant, 2008. **23**(1): p. 47-51.
2. Quetelet, L.A., *A treatise on man and the development of his faculties. 1842.* Obes Res, 1994. **2**(1): p. 72-85.
3. Flegal, K.M., et al., *Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999-2010.* JAMA, 2012. **307**(5): p. 491-7.
4. Ogden, C.L., et al., *Prevalence of childhood and adult obesity in the United States, 2011-2012.* JAMA, 2014. **311**(8): p. 806-14.
5. Flegal, K.M., et al., *Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis.* JAMA, 2013. **309**(1): p. 71-82.
6. Heron, M., *Deaths: leading causes for 2009.* Natl Vital Stat Rep, 2012. **61**(7): p. 1-94.
7. A., P., *A.M.A. Recognizes obesity as a disease.*, in *New York Times*. 2013: [Internet]. 2013 [cited 2014 Nov 20]. Available from: http://www.nytimes.com/2013/06/19/business/ama-recognizes-obesity-as-a-disease.html.
8. Cawley, J. and C. Meyerhoefer, *The medical care costs of obesity: an instrumental variables approach.* J Health Econ, 2012. **31**(1): p. 219-30.
9. Finkelstein, E.A., et al., *Annual medical spending attributable to obesity: payer-and service-specific estimates.* Health Aff (Millwood), 2009. **28**(5): p. w822-31.
10. National Heart, B., and Lung, Institute. *What Causes Overweight and Obesity?* ; Available from: http://www.nhlbi.nih.gov/health/health-topics/topics/obe/causes.
11. Staff, M.C., *Risk Factors.* Diseases and Conditions Obesity, 2014.
12. Glass, T.A. and M.J. McAtee, *Behavioral science at the crossroads in public health: extending horizons, envisioning the future.* Soc Sci Med, 2006. **62**(7): p. 1650-71.
13. Huang, T.T., et al., *A systems-oriented multilevel framework for addressing obesity in the 21st century.* Prev Chronic Dis, 2009. **6**(3): p. A82.
14. Kasl, S.V., *Stress and health.* Annu Rev Public Health, 1984. **5**: p. 319-41.
15. Conway, T.L., Abbey, A., & French, J. R. P., Jr. , *Beliefs about control in different life domains*, in *American Psychological Association Convention*. (1983, August): San Francisco, California.
16. Cohen, S., D. Janicki-Deverts, and G.E. Miller, *Psychological stress and disease.* JAMA, 2007. **298**(14): p. 1685-7.
17. Everson-Rose, S.A. and T.T. Lewis, *Psychosocial factors and cardiovascular diseases.* Annu Rev Public Health, 2005. **26**: p. 469-500.
18. Hemingway, H. and M. Marmot, *Evidence based cardiology: psychosocial factors in the aetiology and prognosis of coronary heart disease. Systematic review of prospective cohort studies.* BMJ, 1999. **318**(7196): p. 1460-7.
19. Overgaard, D., F. Gyntelberg, and B.L. Heitmann, *Psychological workload and body weight: is there an association? A review of the literature.* Occup Med (Lond), 2004. **54**(1): p. 35-41.
20. Block, J.P., et al., *Psychosocial stress and change in weight among US adults.* Am J Epidemiol, 2009. **170**(2): p. 181-92.

21. Kershaw, K.N., et al., *Social relationships and longitudinal changes in body mass index and waist circumference: the coronary artery risk development in young adults study.* Am J Epidemiol, 2014. **179**(5): p. 567-75.
22. Wardle, J., et al., *Stress and adiposity: a meta-analysis of longitudinal studies.* Obesity (Silver Spring), 2011. **19**(4): p. 771-8.
23. Van Strien, T., et al., *Life events, emotional eating and change in body mass index.* Int J Obes, 1986. **10**(1): p. 29-35.
24. Harding, J.L., et al., *Psychosocial stress is positively associated with body mass index gain over 5 years: evidence from the longitudinal AusDiab study.* Obesity (Silver Spring), 2014. **22**(1): p. 277-86.
25. Brunner, E.J., T. Chandola, and M.G. Marmot, *Prospective effect of job strain on general and central obesity in the Whitehall II Study.* Am J Epidemiol, 2007. **165**(7): p. 828-37.
26. Kivimaki, M., et al., *Work stress, weight gain and weight loss: evidence for bidirectional effects of job strain on body mass index in the Whitehall II study.* Int J Obes (Lond), 2006. **30**(6): p. 982-7.
27. Patterson, Z.R. and A. Abizaid, *Stress induced obesity: lessons from rodent models of stress.* Front Neurosci, 2013. **7**: p. 130.
28. Hemmingsson, E., *A new model of the role of psychological and emotional distress in promoting obesity: conceptual review with implications for treatment and prevention.* Obes Rev, 2014. **15**(9): p. 769-79.
29. Sapolsky, R.M., L.M. Romero, and A.U. Munck, *How do glucocorticoids influence stress responses? Integrating permissive, suppressive, stimulatory, and preparative actions.* Endocr Rev, 2000. **21**(1): p. 55-89.
30. Moore, C.J. and S.A. Cunningham, *Social position, psychological stress, and obesity: a systematic review.* J Acad Nutr Diet, 2012. **112**(4): p. 518-26.
31. Cohen, S., T. Kamarck, and R. Mermelstein, *A global measure of perceived stress.* J Health Soc Behav, 1983. **24**(4): p. 385-96.
32. Karasek R, T.T., *Healthy Work—Stress, Productivity and the Reconstruction of Working Life*. 1990, New York: Basic Books.
33. Sominsky, L. and S.J. Spencer, *Eating behavior and stress: a pathway to obesity.* Front Psychol, 2014. **5**: p. 434.
34. Schatzberg, A.F., et al., *HPA axis genetic variation, cortisol and psychosis in major depression.* Mol Psychiatry, 2014. **19**(10): p. 1151.
35. Speliotes, E.K., et al., *Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index.* Nat Genet, 2010. **42**(11): p. 937-48.
36. Locke, A.E., et al., *Genetic studies of body mass index yield new insights for obesity biology.* Nature, 2015. **518**(7538): p. 197-206.
37. El-Sayed Moustafa, J.S. and P. Froguel, *From obesity genetics to the future of personalized obesity therapy.* Nat Rev Endocrinol, 2013. **9**(7): p. 402-13.
38. Belsky, D.W., et al., *Development and evaluation of a genetic risk score for obesity.* Biodemography Soc Biol, 2013. **59**(1): p. 85-100.
39. Domingue, B.W., et al., *Polygenic risk predicts obesity in both white and black young adults.* PLoS One, 2014. **9**(7): p. e101596.

40. Qi, L. and Y.A. Cho, *Gene-environment interaction and obesity.* Nutr Rev, 2008. **66**(12): p. 684-94.
41. Andreasen, C.H., et al., *Low physical activity accentuates the effect of the FTO rs9939609 polymorphism on body fat accumulation.* Diabetes, 2008. **57**(1): p. 95-101.
42. Grarup, N. and G. Andersen, *Gene-environment interactions in the pathogenesis of type 2 diabetes and metabolism.* Curr Opin Clin Nutr Metab Care, 2007. **10**(4): p. 420-6.
43. Miyaki, K., et al., *Increased risk of obesity resulting from the interaction between high energy intake and the Trp64Arg polymorphism of the beta3-adrenergic receptor gene in healthy Japanese men.* J Epidemiol, 2005. **15**(6): p. 203-10.
44. Song, Y., et al., *The interaction between the interleukin 6 receptor gene genotype and dietary energy intake on abdominal obesity in Japanese men.* Metabolism, 2007. **56**(7): p. 925-30.
45. Pishva, E., et al., *Epigenetic genes and emotional reactivity to daily life events: a multi-step gene-environment interaction study.* PLoS One, 2014. **9**(6): p. e100935.
46. Berger, S.L., *The complex language of chromatin regulation during transcription.* Nature, 2007. **447**(7143): p. 407-12.
47. Vera, E., et al., *Epigenetic regulation of telomeres in human cancer.* Oncogene, 2008. **27**(54): p. 6817-33.
48. Lennartsson, A. and K. Ekwall, *Histone modification patterns and epigenetic codes.* Biochim Biophys Acta, 2009. **1790**(9): p. 863-8.
49. Anderson, O.S., K.E. Sant, and D.C. Dolinoy, *Nutrition and epigenetics: an interplay of dietary methyl donors, one-carbon metabolism and DNA methylation.* J Nutr Biochem, 2012. **23**(8): p. 853-9.
50. Hou, L., et al., *Environmental chemical exposures and human epigenetics.* Int J Epidemiol, 2012. **41**(1): p. 79-105.
51. Zeilinger, S., et al., *Tobacco smoking leads to extensive genome-wide changes in DNA methylation.* PLoS One, 2013. **8**(5): p. e63812.
52. Mathews, H.L. and L.W. Janusek, *Epigenetics and psychoneuroimmunology: mechanisms and models.* Brain Behav Immun, 2011. **25**(1): p. 25-39.
53. McGowan, P.O. and M. Szyf, *The epigenetics of social adversity in early life: implications for mental health outcomes.* Neurobiol Dis, 2010. **39**(1): p. 66-72.
54. Szyf, M., *Mind-body interrelationship in DNA methylation.* Chem Immunol Allergy, 2012. **98**: p. 85-99.
55. Blaze, J., L. Scheuing, and T.L. Roth, *Differential methylation of genes in the medial prefrontal cortex of developing and adult rats following exposure to maltreatment or nurturing care during infancy.* Dev Neurosci, 2013. **35**(4): p. 306-16.
56. Wu, Y., et al., *Early-life stress reduces DNA methylation of the Pomc gene in male mice.* Endocrinology, 2014. **155**(5): p. 1751-62.
57. Weaver, I.C., *Epigenetic programming by maternal behavior and pharmacological intervention. Nature versus nurture: let's call the whole thing off.* Epigenetics, 2007. **2**(1): p. 22-8.
58. Sasaki, A., W.C. de Vega, and P.O. McGowan, *Biological embedding in mental health: an epigenomic perspective.* Biochem Cell Biol, 2013. **91**(1): p. 14-21.

59. Fraga, M.F., et al., *Epigenetic differences arise during the lifetime of monozygotic twins.* Proc Natl Acad Sci U S A, 2005. **102**(30): p. 10604-9.
60. Dick, K.J., et al., *DNA methylation and body-mass index: a genome-wide analysis.* Lancet, 2014. **383**(9933): p. 1990-8.
61. Zhao, J., J. Goldberg, and V. Vaccarino, *Promoter methylation of serotonin transporter gene is associated with obesity measures: a monozygotic twin study.* Int J Obes (Lond), 2013. **37**(1): p. 140-5.
62. Essex, M.J., et al., *Epigenetic vestiges of early developmental adversity: childhood stress exposure and DNA methylation in adolescence.* Child Dev, 2013. **84**(1): p. 58-75.
63. Provencal, N. and E.B. Binder, *The effects of early life stress on the epigenome: From the womb to adulthood and even before.* Exp Neurol, 2014.
64. Bird, A., *DNA methylation patterns and epigenetic memory.* Genes Dev, 2002. **16**(1): p. 6-21.
65. Kopelman, P.G., *Obesity as a medical problem.* Nature, 2000. **404**(6778): p. 635-43.
66. Kumanyika, S., *Obesity, health disparities, and prevention paradigms: hard questions and hard choices.* Prev Chronic Dis, 2005. **2**(4): p. A02.
67. Mokdad, A.H., et al., *Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001.* JAMA, 2003. **289**(1): p. 76-9.
68. Rand, C.S.W. and J.M. Kuldau, *The epidemiology of obesity and self-defined weight problem in the general population: Gender, race, age, and social class.* International Journal of Eating Disorders, 1990. **9**(3): p. 329-343.
69. Luppino, F.S., et al., *Overweight, obesity, and depression: a systematic review and meta-analysis of longitudinal studies.* Arch Gen Psychiatry, 2010. **67**(3): p. 220-9.
70. Onyike, C.U., et al., *Is obesity associated with major depression? Results from the Third National Health and Nutrition Examination Survey.* Am J Epidemiol, 2003. **158**(12): p. 1139-47.
71. Walley, A.J., A.I. Blakemore, and P. Froguel, *Genetics of obesity and the prediction of risk for health.* Hum Mol Genet, 2006. **15 Spec No 2**: p. R124-30.
72. Visscher, P.M., et al., *Five years of GWAS discovery.* Am J Hum Genet, 2012. **90**(1): p. 7-24.
73. Choquet, H. and D. Meyre, *Genetics of Obesity: What have we Learned?* Curr Genomics, 2011. **12**(3): p. 169-79.
74. Scott, K.A., S.J. Melhorn, and R.R. Sakai, *Effects of Chronic Social Stress on Obesity.* Curr Obes Rep, 2012. **1**(1): p. 16-25.
75. Stunkard, A.J., T.T. Foch, and Z. Hrubec, *A twin study of human obesity.* JAMA, 1986. **256**(1): p. 51-4.
76. Stenberg, A., *Interpreting estimates of heritability--a note on the twin decomposition.* Econ Hum Biol, 2013. **11**(2): p. 201-5.
77. Zuk, O., et al., *The mystery of missing heritability: Genetic interactions create phantom heritability.* Proc Natl Acad Sci U S A, 2012. **109**(4): p. 1193-8.
78. Maher, B., *Personal genomes: The case of the missing heritability.* Nature, 2008. **456**(7218): p. 18-21.
79. Weinhold, B., *Epigenetics: the science of change.* Environ Health Perspect, 2006. **114**(3): p. A160-7.

80. Shen, L. and R.A. Waterland, *Methods of DNA methylation analysis.* Curr Opin Clin Nutr Metab Care, 2007. **10**(5): p. 576-81.
81. Razin, A. and A.D. Riggs, *DNA methylation and gene function.* Science, 1980. **210**(4470): p. 604-10.
82. Numata, S., et al., *DNA methylation signatures in development and aging of the human prefrontal cortex.* Am J Hum Genet, 2012. **90**(2): p. 260-72.
83. Bick, J., et al., *Childhood adversity and DNA methylation of genes involved in the hypothalamus-pituitary-adrenal axis and immune system: whole-genome and candidate-gene associations.* Dev Psychopathol, 2012. **24**(4): p. 1417-25.
84. Suderman, M., et al., *Childhood abuse is associated with methylation of multiple loci in adult DNA.* BMC Med Genomics, 2014. **7**: p. 13.
85. Wellcome Trust Case Control, C., *Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls.* Nature, 2007. **447**(7145): p. 661-78.
86. Troxel, W.M., et al., *Chronic stress burden, discrimination, and subclinical carotid artery disease in African American and Caucasian women.* Health Psychol, 2003. **22**(3): p. 300-9.
87. Krieger, N., et al., *The inverse hazard law: blood pressure, sexual harassment, racial discrimination, workplace abuse and occupational exposures in US low-income black, white and Latino workers.* Soc Sci Med, 2008. **67**(12): p. 1970-81.
88. Falkenberg, V.R., et al., *Functional genomics of serotonin receptor 2A (HTR2A): interaction of polymorphism, methylation, expression and disease association.* Neuromolecular Med, 2011. **13**(1): p. 66-76.
89. Kiecolt-Glaser, J.K., et al., *Chronic stress and age-related increases in the proinflammatory cytokine IL-6.* Proc Natl Acad Sci U S A, 2003. **100**(15): p. 9090-5.
90. Torres, S.J. and C.A. Nowson, *Relationship between stress, eating behavior, and obesity.* Nutrition, 2007. **23**(11-12): p. 887-94.
91. Marmorstein, N.R. and D. Hart, *Interactions Between MAOA Genotype and Receipt of Public Assistance: Predicting Change in Depressive Symptoms and Body Mass Index.* J Res Adolesc, 2011. **21**(3): p. 619-630.
92. Singh, A., et al., *Gene by stress genome-wide interaction analysis and path analysis identify EBF1 as a cardiovascular and metabolic risk gene.* Eur J Hum Genet, 2015. **23**(6): p. 854-62.
93. Iqbal Kring, S.I., et al., *Associations between APOE variants and metabolic traits and the impact of psychological stress.* PLoS One, 2011. **6**(1): p. e15745.
94. Kring, S.I., et al., *Impact of psychological stress on the associations between apolipoprotein E variants and metabolic traits: findings in an American sample of caregivers and controls.* Psychosom Med, 2010. **72**(5): p. 427-33.
95. Feinleib, M., et al., *The Framingham Offspring Study. Design and preliminary data.* Prev Med, 1975. **4**(4): p. 518-25.
96. Perez-Cornago, A., et al., *DNA hypermethylation of the serotonin receptor type-2A gene is associated with a worse response to a weight loss intervention in subjects with metabolic syndrome.* Nutrients, 2014. **6**(6): p. 2387-403.

97. Na, Y.K., et al., *Increased methylation of interleukin 6 gene is associated with obesity in Korean women.* Mol Cells, 2015. **38**(5): p. 452-6.
98. Zhang, F.F., et al., *White blood cell global methylation and IL-6 promoter methylation in association with diet and lifestyle risk factors in a cancer-free population.* Epigenetics, 2012. **7**(6): p. 606-14.
99. Gomes, M.V., et al., *Age-related changes in the global DNA methylation profile of leukocytes are linked to nutrition but are not associated with the MTHFR C677T genotype or to functional capacities.* PLoS One, 2012. **7**(12): p. e52570.
100. Kerkel, K., et al., *Genomic surveys by methylation-sensitive SNP analysis identify sequence-dependent allele-specific DNA methylation.* Nat Genet, 2008. **40**(7): p. 904-8.
101. Schmitz, R.J., et al., *Patterns of population epigenomic diversity.* Nature, 2013. **495**(7440): p. 193-8.
102. Tycko, B., *Allele-specific DNA methylation: beyond imprinting.* Hum Mol Genet, 2010. **19**(R2): p. R210-20.
103. Kershaw, E.E. and J.S. Flier, *Adipose tissue as an endocrine organ.* J Clin Endocrinol Metab, 2004. **89**(6): p. 2548-56.
104. Benjamini, Y.H.Y., *Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing.* Journal of the Royal Statistical Society, 1995. **57**(1): p. 289-300.
105. Kanherkar, R.R., N. Bhatia-Dey, and A.B. Csoka, *Epigenetics across the human lifespan.* Front Cell Dev Biol, 2014. **2**: p. 49.
106. Cantone, I. and A.G. Fisher, *Epigenetic programming and reprogramming during development.* Nat Struct Mol Biol, 2013. **20**(3): p. 282-9.
107. Slatkin, M., *Epigenetic inheritance and the missing heritability problem.* Genetics, 2009. **182**(3): p. 845-50.
108. Andersson, E.A., et al., *Genetic risk score of 46 type 2 diabetes risk variants associates with changes in plasma glucose and estimates of pancreatic beta-cell function over 5 years of follow-up.* Diabetes, 2013. **62**(10): p. 3610-7.
109. Ripatti, S., et al., *A multilocus genetic risk score for coronary heart disease: case-control and prospective cohort analyses.* Lancet, 2010. **376**(9750): p. 1393-400.
110. Needham, B.L., et al., *Life course socioeconomic status and DNA methylation in genes related to stress reactivity and inflammation: The multi-ethnic study of atherosclerosis.* Epigenetics, 2015. **10**(10): p. 958-69.
111. Cao-Lei, L., et al., *DNA methylation mediates the impact of exposure to prenatal maternal stress on BMI and central adiposity in children at age 13(1/2) years: Project Ice Storm.* Epigenetics, 2015. **10**(8): p. 749-61.
112. Tian, C., P.K. Gregersen, and M.F. Seldin, *Accounting for ancestry: population substructure and genome-wide association studies.* Hum Mol Genet, 2008. **17**(R2): p. R143-50.
113. Jackson, J.S., K.M. Knight, and J.A. Rafferty, *Race and unhealthy behaviors: chronic stress, the HPA axis, and physical and mental health disparities over the life course.* Am J Public Health, 2010. **100**(5): p. 933-9.
114. Marniemi, J., et al., *Visceral fat and psychosocial stress in identical twins discordant for obesity.* J Intern Med, 2002. **251**(1): p. 35-43.

115. HR Rothstein, A.S., M Borenstein, *Publication bias in meta-analysis: Prevention, assessment and adjustments*, ed. A.S. HR Rothstein, M Borenstein. 2006, West Sussex PO198SQ, England: John Wiley & Sons.
116. Moher, D., et al., *Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.* Int J Surg, 2010. **8**(5): p. 336-41.
117. Tan, L.J., et al., *Replication of 6 obesity genes in a meta-analysis of genome-wide association studies from diverse ancestries.* PLoS One, 2014. **9**(5): p. e96149.
118. Tan, J.T., et al., *FTO variants are associated with obesity in the Chinese and Malay populations in Singapore.* Diabetes, 2008. **57**(10): p. 2851-7.
119. Adeyemo, A., et al., *FTO genetic variation and association with obesity in West Africans and African Americans.* Diabetes, 2010. **59**(6): p. 1549-54.
120. Ronn, T., et al., *A six months exercise intervention influences the genome-wide DNA methylation pattern in human adipose tissue.* PLoS Genet, 2013. **9**(6): p. e1003572.
121. Bild, D.E., et al., *Multi-Ethnic Study of Atherosclerosis: objectives and design.* Am J Epidemiol, 2002. **156**(9): p. 871-81.
122. Green, D., et al., *Obtaining informed consent for genetic studies: The multiethnic study of atherosclerosis.* Am J Epidemiol, 2006. **164**(9): p. 845-51.
123. Bromberger, J.T. and K.A. Matthews, *A longitudinal study of the effects of pessimism, trait anxiety, and life stress on depressive symptoms in middle-aged women.* Psychol Aging, 1996. **11**(2): p. 207-13.
124. Krieger, N. and S. Sidney, *Racial discrimination and blood pressure: the CARDIA Study of young black and white adults.* Am J Public Health, 1996. **86**(10): p. 1370-8.
125. Williams, D.R., *Race and health: basic questions, emerging directions.* Ann Epidemiol, 1997. **7**(5): p. 322-33.
126. Radloff, L.S., *The use of the Center for Epidemiologic Studies Depression Scale in adolescents and young adults.* J Youth Adolesc, 1991. **20**(2): p. 149-66.
127. Kershaw, K.N., et al., *Associations of chronic individual-level and neighbourhood-level stressors with incident coronary heart disease: the Multi-Ethnic Study of Atherosclerosis.* J Epidemiol Community Health, 2015. **69**(2): p. 136-41.
128. Williams, D.R. and S.A. Mohammed, *Discrimination and racial disparities in health: evidence and needed research.* J Behav Med, 2009. **32**(1): p. 20-47.
129. Radloff, L.S., *The CES-D Scale: A self-report depression scale for research in the general population.* Applied Psychological Measurement, 1977. **1**: p. 1385-1401.
130. Lewinsohn, P.M., et al., *Center for Epidemiologic Studies Depression Scale (CES-D) as a screening instrument for depression among community-residing older adults.* Psychol Aging, 1997. **12**(2): p. 277-87.
131. Cheung, C.K. and C. Bagley, *Validating an American scale in Hong Kong: the Center for Epidemiological Studies Depression Scale (CES-D).* J Psychol, 1998. **132**(2): p. 169-86.
132. Roberts, R.E., *Reliability of the CES-D Scale in different ethnic contexts.* Psychiatry Res, 1980. **2**(2): p. 125-34.
133. Remigio-Baker, R.A., et al., *Physical environment may modify the association between depressive symptoms and change in waist circumference: the multi-ethnic study of atherosclerosis.* Psychosomatics, 2014. **55**(2): p. 144-54.

134. Tavakol, M.D., R. , *Making sense of Cronbach's alpha.* International Journal of Medical Education, 2011. **2**: p. 53-55.
135. *An integrated map of genetic variation from 1,092 human genomes.* Nature, 2012. **491**(7422): p. 56-65.
136. Howie, B.N., P. Donnelly, and J. Marchini, *A flexible and accurate genotype imputation method for the next generation of genome-wide association studies.* PLoS Genet, 2009. **5**(6): p. e1000529.
137. Rasmussen-Torvik, L.J., et al., *Fasting glucose GWAS candidate region analysis across ethnic groups in the Multiethnic Study of Atherosclerosis (MESA).* Genetic Epidemiology, 2012. **36**(4): p. 384-91.
138. Zheng, J., et al., *A comparison of approaches to account for uncertainty in analysis of imputed genotypes.* Genet Epidemiol, 2011. **35**(2): p. 102-10.
139. Patterson, N., A.L. Price, and D. Reich, *Population structure and eigenanalysis.* PLoS Genet, 2006. **2**(12): p. e190.
140. Price, A.L., et al., *Principal components analysis corrects for stratification in genome-wide association studies.* Nature Genetics, 2006. **38**(8): p. 904-9.
141. Stults-Kolehmainen, M.A. and R. Sinha, *The effects of stress on physical activity and exercise.* Sports Med, 2014. **44**(1): p. 81-121.
142. McAuley, P.A., et al., *Physical activity, measures of obesity, and cardiometabolic risk: the Multi-Ethnic Study of Atherosclerosis (MESA).* J Phys Act Health, 2014. **11**(4): p. 831-7.
143. Doo, M. and Y. Kim, *Obesity: interactions of genome and nutrients intake.* Prev Nutr Food Sci, 2015. **20**(1): p. 1-7.
144. Nettleton, J.A., et al., *Gene x dietary pattern interactions in obesity: analysis of up to 68 317 adults of European ancestry.* Hum Mol Genet, 2015. **24**(16): p. 4728-38.
145. Cologne, J., et al., *Proportional hazards regression in epidemiologic follow-up studies: an intuitive consideration of primary time scale.* Epidemiology, 2012. **23**(4): p. 565-73.
146. Korn, E.L., B.I. Graubard, and D. Midthune, *Time-to-event analysis of longitudinal follow-up of a survey: choice of the time-scale.* Am J Epidemiol, 1997. **145**(1): p. 72-80.
147. Allison, P.D., *Survial Analysis Using SAS: A Practical Guide*. 2nd ed. 2010, Cary, NC SAS Insititute Inc. : SAS Press.
148. Knol, M.J., et al., *Estimating interaction on an additive scale between continuous determinants in a logistic regression model.* International Journal of Epidemiology, 2007. **36**(5): p. 1111-8.
149. VanderWeele, T.J., *Sample size and power calculations for additive interactions.* Epidemiologic Methods, 2012. **1**(1): p. 159-188.
150. T, T., *A Package for Survival Analysis in S*. 2015.
151. Klein, J., *Survival Analysis: Techniques for Censored and Truncated Data*. 2nd ed. 2003. Corr. 3rd printing 2005 ed. Statistics for Biology and Health Series. 2005, New York, NY: Springer New York. 537.
152. Rufibach, S.H.a.K., *SurvRegCensCov: Weibull Regression for a Right-Censored Endpoint with Interval-Censored Covariate.* 2015: R package version 1.4.
153. Williams, D.R., et al., *Racial Differences in Physical and Mental Health: Socio-economic Status, Stress and Discrimination.* J Health Psychol, 1997. **2**(3): p. 335-51.

154. Olstad, D.L., et al., *Hair cortisol levels, perceived stress and body mass index in women and children living in socioeconomically disadvantaged neighborhoods: the READI study.* Stress, 2016. **19**(2): p. 158-67.
155. Ellis, B.J., et al., *Differential susceptibility to the environment: an evolutionary--neurodevelopmental theory.* Dev Psychopathol, 2011. **23**(1): p. 7-28.
156. Rokholm, B., et al., *Increasing genetic variance of body mass index during the Swedish obesity epidemic.* PLoS One, 2011. **6**(11): p. e27135.
157. Knol, T.J.V.a.M.J. *A Tutorial on Interaction*. Epidemiologic Methods, 2014. **3**, 33-72.
158. Hung, C.F., et al., *A genetic risk score combining 32 SNPs is associated with body mass index and improves obesity prediction in people with major depressive disorder.* BMC Med, 2015. **13**: p. 86.
159. Samaan, Z., et al., *Obesity genes and risk of major depressive disorder in a multiethnic population: a cross-sectional study.* J Clin Psychiatry, 2015. **76**(12): p. e1611-8.
160. Clarke, T.K., et al., *Major depressive disorder and current psychological distress moderate the effect of polygenic risk for obesity on body mass index.* Transl Psychiatry, 2015. **5**: p. e592.
161. Irwin, M., K.H. Artin, and M.N. Oxman, *Screening for depression in the older adult: criterion validity of the 10-item Center for Epidemiological Studies Depression Scale (CES-D).* Arch Intern Med, 1999. **159**(15): p. 1701-4.
162. Kohout, F.J., et al., *Two shorter forms of the CES-D (Center for Epidemiological Studies Depression) depression symptoms index.* J Aging Health, 1993. **5**(2): p. 179-93.
163. Carleton, R.N., et al., *The center for epidemiologic studies depression scale: a review with a theoretical and empirical examination of item content and factor structure.* PLoS One, 2013. **8**(3): p. e58067.
164. Price, A.L., et al., *A genomewide admixture map for Latino populations.* Am J Hum Genet, 2007. **80**(6): p. 1024-36.
165. Jaenisch, R. and A. Bird, *Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals.* Nat Genet, 2003. **33 Suppl**: p. 245-54.
166. Milagro, F.I., et al., *High fat diet-induced obesity modifies the methylation pattern of leptin promoter in rats.* J Physiol Biochem, 2009. **65**(1): p. 1-9.
167. Rzehak, P., et al., *Maternal Smoking during Pregnancy and DNA-Methylation in Children at Age 5.5 Years: Epigenome-Wide-Analysis in the European Childhood Obesity Project (CHOP)-Study.* PLoS One, 2016. **11**(5): p. e0155554.
168. Du, P., Kibbe, W.A. and Lin, S.M.,, *lumi: a pipeline for processing Illumina microarray*. 2008, Bioinformatics 24(13):1547-1548.
169. Pidsley, R., et al., *A data-driven approach to preprocessing Illumina 450K methylation array data.* BMC Genomics, 2013. **14**: p. 293.
170. Wang, K., et al., *A genome-wide association study on obesity and obesity-related traits.* PLoS One, 2011. **6**(4): p. e18939.
171. Du, P., et al., *Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis.* BMC Bioinformatics, 2010. **11**: p. 587.
172. Portela, A. and M. Esteller, *Epigenetic modifications and human disease.* Nat Biotechnol, 2010. **28**(10): p. 1057-68.

173. Whitaker, J.W., et al., *An imprinted rheumatoid arthritis methylome signature reflects pathogenic phenotype.* Genome Med, 2013. **5**(4): p. 40.
174. Kent, W.J., et al., *The human genome browser at UCSC.* Genome Res, 2002. **12**(6): p. 996-1006.
175. Dunning, M.J., et al., *beadarray: R classes and methods for Illumina bead-based data.* Bioinformatics, 2007. **23**(16): p. 2183-4.
176. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K., *limma powers differential expression analyses for RNA-sequencing and microarray studies.* 2015, Nucleic Acids Research.
177. Ritchie, M.E., et al., *A comparison of background correction methods for two-colour microarrays.* Bioinformatics, 2007. **23**(20): p. 2700-7.
178. Shi, W., A. Oshlack, and G.K. Smyth, *Optimizing the noise versus bias trade-off for Illumina whole genome expression BeadChips.* Nucleic Acids Res, 2010. **38**(22): p. e204.
179. Liu, Y., et al., *Methylomics of gene expression in human monocytes.* Hum Mol Genet, 2013. **22**(24): p. 5065-74.
180. Pilkonis, P.A., S.D. Imber, and P. Rubinsky, *Dimensions of life stress in psychiatric patients.* J Human Stress, 1985. **11**(1): p. 5-10.
181. Krieger, N., et al., *Experiences of discrimination: validity and reliability of a self-report measure for population health research on racism and health.* Soc Sci Med, 2005. **61**(7): p. 1576-96.
182. Taylor, T.R., T.W. Kamarck, and S. Shiffman, *Validation of the Detroit Area Study Discrimination Scale in a community sample of older African American adults: the Pittsburgh healthy heart project.* Int J Behav Med, 2004. **11**(2): p. 88-94.
183. Leung, D.Y., T.H. Lam, and S.S. Chan, *Three versions of Perceived Stress Scale: validation in a sample of Chinese cardiac patients who smoke.* BMC Public Health, 2010. **10**: p. 513.
184. Andreou, E., et al., *Perceived Stress Scale: reliability and validity study in Greece.* Int J Environ Res Public Health, 2011. **8**(8): p. 3287-98.
185. Bose, M., B. Olivan, and B. Laferrere, *Stress and obesity: the role of the hypothalamic-pituitary-adrenal axis in metabolic disease.* Curr Opin Endocrinol Diabetes Obes, 2009. **16**(5): p. 340-6.
186. Brown H, P.R., *Applied Mixed Models in Medicine*. 2006, West Sussex: John Wiley & Sons Ltd. .
187. Liu, Y., et al., *GeMes, clusters of DNA methylation under genetic control, can inform genetic and epigenetic analysis of disease.* Am J Hum Genet, 2014. **94**(4): p. 485-95.
188. Kenward, M.G. and J.H. Roger, *Small sample inference for fixed effects from restricted maximum likelihood.* Biometrics, 1997. **53**(3): p. 983-97.
189. Benjamini, Y., et al., *Controlling the false discovery rate in behavior genetics research.* Behav Brain Res, 2001. **125**(1-2): p. 279-84.
190. Team, R.C., *A language and environment for statistical computing.*, R.F.f.S. Computing, Editor. 2016: Vienna, Austria.
191. Douglas Bates, M.M., Ben Bolker, Steve Walker, *Fitting Linear Mixed-Effects Models Using lme4*. 2015: Journal of Statistical Software, 67(1), 1-48.

192. Ulrich Halekoh, S.H., *A Kenward-Roger Approximation and Parametric Bootstrap Methods for Tests in Linear Mixed Models - The R Package pbkrtest.* Journal of Statistical Software, 2014. **59 (9)**: p. 1-30.
193. Zhang, W., et al., *Predicting genome-wide DNA methylation using methylation marks, genomic position, and DNA regulatory elements.* Genome Biol, 2015. **16**: p. 14.
194. Cox, G.A., et al., *The mouse fidgetin gene defines a new role for AAA family proteins in mammalian development.* Nat Genet, 2000. **26**(2): p. 198-202.
195. Yang, Y., et al., *Interaction between fidgetin and protein kinase A-anchoring protein AKAP95 is critical for palatogenesis in the mouse.* J Biol Chem, 2006. **281**(31): p. 22352-9.
196. O'Leary, N.A., et al., *Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation.* Nucleic Acids Res, 2016. **44**(D1): p. D733-45.
197. Unternaehrer, E., et al., *Dynamic changes in DNA methylation of stress-associated genes (OXTR, BDNF) after acute psychosocial stress.* Transl Psychiatry, 2012. **2**: p. e150.
198. Roth, T.L., et al., *Lasting epigenetic influence of early-life adversity on the BDNF gene.* Biol Psychiatry, 2009. **65**(9): p. 760-9.
199. Maes, T., A. Barcelo, and C. Buesa, *Neuron navigator: a human gene family with homology to unc-53, a cell guidance gene from Caenorhabditis elegans.* Genomics, 2002. **80**(1): p. 21-30.
200. Coy, J.F., et al., *Pore membrane and/or filament interacting like protein 1 (POMFIL1) is predominantly expressed in the nervous system and encodes different protein isoforms.* Gene, 2002. **290**(1-2): p. 73-94.
201. Smith, C.M., et al., *The gene expression database for mouse development (GXD): putting developmental expression information at your fingertips.* Dev Dyn, 2014. **243**(10): p. 1176-86.
202. Wagner, J.R., et al., *The relationship between DNA methylation, genetic and expression inter-individual variation in untransformed human fibroblasts.* Genome Biol, 2014. **15**(2): p. R37.
203. Lokk, K., et al., *DNA methylome profiling of human tissues identifies global and tissue-specific methylation patterns.* Genome Biol, 2014. **15**(4): p. r54.
204. Sinha, M., et al., *Self-reported race and genetic admixture.* N Engl J Med, 2006. **354**(4): p. 421-2.
205. Sucheston, L.E., et al., *Genetic ancestry, self-reported race and ethnicity in African Americans and European Americans in the PCaP cohort.* PLoS One, 2012. **7**(3): p. e30950.
206. Smith, E.N., et al., *Genetic ancestry of participants in the National Children's Study.* Genome Biol, 2014. **15**(2): p. R22.
207. Storey, J.D. and R. Tibshirani, *Statistical significance for genomewide studies.* Proc Natl Acad Sci U S A, 2003. **100**(16): p. 9440-5.