Gene-environment interactions in obesity: current evidence and future directions

by

Hudson Reddon

A thesis submitted in partial fulfillment of the requirements for the degree of doctor of philosophy

Health Research Methodology
McMaster University

Doctor of Philosophy (2017); Health Research Methodology; McMaster University; Hamilton, Ontario

Title: Gene-environment interactions in obesity: current evidence and future directions

Author: Hudson Reddon

Supervisor: Dr. David Meyre

Pages: 185

ABSTRACT

Background: Obesity is a multifactorial disease caused by the interplay of environmental and genetic risk factors. With the prevalence of obesity more than doubling since 1980, this disease has become a global epidemic. The objectives of this research were to (1) review the current evidence of gene-environment interactions (GEI) in the field of obesity, (2) investigate novel GEI involving sedentary behaviour, sleep duration and alcohol consumption, (3) assess GEI using a cumulative environmental risk score, and (4) provide an overview of methodological weaknesses in GEI studies and provide suggestions for future directions.

Methods: The data for the gene-environment interaction analyses were collected from the EpiDREAM study: a cohort study including participants of six ethnic backgrounds from 17 countries worldwide. A subset of 17 423 participants with complete genotype and phenotype information was included in the analysis. Twenty-three obesity predisposing single nucleotide polymorphisms (SNPs) were analyzed independently and as a genetic risk score (GRS). Linear regression models were used to analyze these interactions.

Results: Heritability, monogenic and polygenic obesity studies provide converging evidence that obesity-predisposing genes interact with a variety of environmental exposures including physical activity and diet patterns. In the EpiDREAM cohort, we found that increased sedentary time did not interact with obesity predisposing SNPs or the GRS to modulate BMI. The interaction between sedentary time and physical activity was also not significant. We observed a U-shaped association between sleep duration and BMI and sleep duration did not appear to moderate the impact of the obesity predisposing SNPs or the GRS. However, we did observe an alcohol x *FTO* rs1421085 interaction, whereby increased alcohol consumption attenuated the impact of *FTO* rs1421085 variation on BMI. We also found that the combined effect of several environmental risk factors significantly modified the effect of *FTO* rs3751812 on BMI. Specifically, we found that the effect of the *FTO* rs3751812 SNP on BMI was over two times greater among those in the highest quartile of environmental risk compared to those in the lowest quartile. The GRS did not interact with any of the exposures tested.

Discussion: Our results indicate that sedentary behaviour did not moderate the impact of obesity predisposing genes, while alcohol consumption decreased the impact of variation in *FTO* rs3751812 on BMI. We also observed that variation in *FTO* rs3751812 interacted with a cumulative environmental risk score to moderate BMI. The growing body of GEI evidence has provided a deeper understanding of obesity aetiology and may have tremendous applications in the emerging field of personalized medicine and individualized lifestyle recommendations. Although the number of gene-environment interaction analyses has increased rapidly across multiple disciplines, addressing methodological concerns such as statistical modeling, confounding, biological assumptions and measurement precision will be necessary to fully exploit the potential of the GEI field. With the development of new methodological and measurement techniques such as hypothesis-free genome wide interaction studies and deep phenotyping, it may be possible to translate the information from GEI studies into public health policy and personalized medicine for obesity and other complex human diseases.

ACKNOWLEDGMENTS

I would like to extend my gratitude to my supervisor, Dr. David Meyre, for his unwavering support, constructive feedback, and warm encouragement over the past three years. Without his guidance, consistent help, and engaging teaching style, this thesis would not have been possible.

I would also like to extend my appreciation to my Advisory Committee, Dr. John Cairney, and Dr. Mark Loeb, for their constructive feedback and helpful support throughout this project. I am very thankful to have worked with such knowledgeable and insightful faculty throughout this project.

It would be remiss of me to fail to mention the generous help and encouraging feedback from all members of our research group over the past two years: Jackie Hudson, Dipika Desai and my colleagues Akram Alyass, Alexandra Mayhew, Sébastien Robiou du Pont, Arkan Abadi and Aihua Li. More specifically, Jackie and Dipika provided administrative support and the students in our group provided thoughtful feedback regarding statistical analysis and research design.

Finally, I would like to express my sincerest gratitude to my parents, my partner Amanda and close friends who have provided me with endless support and encouragement throughout this program.

Table of Contents

Abstract 2

Acknowledgments 3

List of Abbreviations 8

Chapter 1-Introduction 9

Chapter 2-Literature Review 11

ABSTRACT 11

INTRODUCTION 12

DEFINITIONS 13

Heritability estimates are influenced by the environment 15

Obesity predisposing gene variants interact with the environment 18

Obesity predisposing gene variants interact with non-modifiable biological factors 18

Obesity predisposing gene variants interact with lifestyle factors 21

Obesity predisposing gene variants interact with disease status/response to treatment 29

Biological processes underlying statistical gene-environment interactions 34

Conclusion 37

Chapter 3- Sedentary behaviour and genetic predisposition to obesity in a multiethnic study 40

ABSTRACT 40

INTRODUCTION 41

METHODS 45

Study Participants 45

Genotyping 45

Phenotyping 46

Statistical Analyses 48

RESULTS 49

Characteristics of the studied cohort 49

Effect of Sedentary Behaviour on BMI 50

Effect of SNPs/GRS on Sedentary behaviour 50

Effect of SNPs/GRS on BMI 50

Interaction Analyses 51

DISCUSSION 51

Chapter 4-Sleep duration and body mass index in a multiethnic study: evidence from observational and genetic epidemiology 70

ABSTRACT 70

INTRODUCTION 71

METHODS 73

Study Participants 73

Genotyping 74

Phenotyping 75

Statistical Analysis 76

RESULTS 77

Characteristics of the studied cohort 77

Effect of Sleep duration on BMI 77

Effect of SNPs/GS on Sleep duration 78
Effect of SNPs/GRS on BMI 78
Sleep duration x SNP/GRS Interaction Analysis 78
DISCUSSION 79

Chapter 5-Alcohol intake and obesity: evidence of sex and gene interactions in a multiethnic study 89
ABSTRACT 89
INTRODUCTION 90
METHODS 93
Study Participants 93
Genotyping 93
Phenotyping 95
Statistical Analysis 95
RESULTS 97
Characteristics of the studied cohort 97
Effect of Alcohol Intake on BMI 97
Effect of SNPs/GS on Alcohol Intake 98
Effect of SNPs/GRS on BMI 98
Alcohol Intake x SNP/GRS Interaction Analysis 99
DISCUSSION 99

Chapter 6-cumulative environmental risk and genetic predisposition to obesity 111
Abstract 111
INTRODUCTION 112
METHODS 114
Study Participants 114
Genotyping 115
Phenotyping 116
Statistical Analyses 117
RESULTS 118
Characteristics of the studied cohort 118
Effect of Lifestyle Factors on BMI 119
Effect of FTO rs3751812 and GRS on BMI 119
Interaction Analyses 120
DISCUSSION 120

Chapter 7-Methodological considerations in gene-environment interaction studies of obesity: recent developments and future options 130
Abstract 130
Introduction 131
Statistical modeling issues in gene-environment interaction research 132
Confounding issues in gene-environment interaction research 136
Considering time of exposure in gene-environment interaction research 138
Measurement issues in gene-environment interactions research 139
Future directions for gene-environment interactions and obesity 143
Conclusion 147

References 150

List of Tables

Table 1. Baseline characteristics by sedentary level in the EpiDREAM study.....56
Table 2. Effect of sedentary behaviour on BMI (adjusted for sex, age, ethnicity and glycemic status). 57
Table 3. Effect of SNPs/GRS on sedentary behaviour (adjusted for sex, age, ethnicity, glycemic status and BMI).....58
Table 4. Effect of SNPs/GRS on BMI (adjusted for sex, age, ethnicity and glycemic status).....59
Table 5. Interaction analyses between sedentary behaviour and obesity predisposing SNPs/GRS (adjusted for sex, age, ethnicity, relatedness, glycemic status, physical activity and physical activity x SNP/GRS interaction). 60
Table 6. Baseline characteristics of the EpiDREAM study stratified by sleep duration...82
Table 7. Linear vs. U-shaped models of the association between sleep duration and BMI. 83
Table 8. Effect of SNPs/GRS on sleep duration.....84
Table 9. Effect of SNPs/GRS on BMI.....85
Table 10. Interaction analysis between SNPs/GRS and sleep duration on BMI. 86
Table 11. Baseline characteristics stratified by alcohol consumption in the EpiDREAM study. 104
Table 12. Effect of alcohol consumption on BMI. 105
Table 13. Effect of SNPs/GRS on alcohol consumption (adjusted for sex, age, ethnicity, glycemic status and BMI).....106
Table 14. Main association between SNPs/GRS and BMI, and interaction effects with alcohol consumption (adjusted for sex, age, ethnicity and glycemic status).....107
Table 15. Baseline characteristics of the EpiDREAM study. 124
Table 16. Effect of lifestyle factors on BMI (adjusted for sex, age, ethnicity and glycemic status).....126
Table 17. Main effect and interaction between cumulative environmental score and *FTO* rs3751812/GRS and BMI (adjusted for sex, age, ethnicity and glycemic status)...127
Table 18. Summary of methodological issues and solutions for gene-environment interaction studies in obesity. 149

List of Supplementary Tables

Supplementary Table 1. Genotype distributions, call rate and Hardy-Weinberg Equilibrium values for the 23 SNPs analyzed. 61
Supplementary Table 2. List of 39 self-reported physical activities. 66
Supplementary Table 3. Description of the 23 SNPs selected in EpiDREAM. 109

List of Figures

Figure 1. Biological model to explain gene-environment interactions in obesity. 39
Figure 2. Mean BMI values stratified by alcohol consumption and *FTO* rs3751812 genotype. 108
Figure 3. Beta value for the effect of the *FTO* rs3751812 risk allele on BMI in each quartile of environmental risk. 128
Figure 4. Mean BMI values stratified by *FTO* rs3751812 genotype and lifestyle quartile. 129

List of Supplementary Figures

Supplementary Figure 1. Flow chart of EpiDREAM study (sedentary x gene interaction). 67
Supplementary Figure 2. Power calculation (80%) for the main effect of obesity predisposing SNPs for a P-value=0.05. 68
Supplementary Figure 3. Power calculation (80%) for the interaction effect between obesity predisposing SNPs and sedentary behaviour for a P-value=0.05. 69
Supplementary Figure 4. Flow chart of EpiDREAM study (sleep x gene interaction) 87
Supplementary Figure 5. Power calculation for the interaction between obesity predisposing SNPs and sleep duration for a P-value=0.05 88
Supplementary Figure 6. Flow chart of EpiDREAM study (alcohol x gene interaction). 110

Supplementary Figure 1. Flow chart of EpiDREAM study (sedentary x gene interaction). 67
Supplementary Figure 2. Power calculation (80%) for the main effect of obesity predisposing SNPs for a P-value=0.05. 68
Supplementary Figure 3. Power calculation (80%) for the interaction effect between obesity predisposing SNPs and sedentary behaviour for a P-value=0.05. 69
Supplementary Figure 4. Flow chart of EpiDREAM study (sleep x gene interaction) 87
Supplementary Figure 5. Power calculation for the interaction between obesity predisposing SNPs and sleep duration for a P-value=0.05 88
Supplementary Figure 6. Flow chart of EpiDREAM study (alcohol x gene interaction). 110

LIST OF ABBREVIATIONS

BMI: body mass index

BPA: basic physical activity score

FSH: follicle stimulating hormone

GEWIS: gene-environment-wide interaction studies

GRS: genetic risk score

GWAS: genome wide association study

HC: hip circumference

IGT: impaired glucose tolerance

METS: metabolic equivalent score

NGT: normal glucose tolerance

PA: physical activity

PCA: principal components analysis

SNP: single nucleotide polymorphism

T2D: type 2 diabetes

CHAPTER 1-INTRODUCTION

Obesity has become a global epidemic and is a known risk factor for a number of adverse health outcomes, including clinical depression and anxiety, osteoarthritis, type 2 diabetes, hypertension, dyslipidemia, cardiovascular disease, cancer (endometrial, breast, colon, kidney, gallbladder, liver) and 8-13 years shorter life expectancy in its more severe forms[1-3]. The economic implications of obesity have also become substantial issues, which are particularly extreme in the United States and are growing concerns in Western Europe, Canada and New Zealand[4]. The leading causes of this epidemic include excess energy intake and declines in physical activity, although additional environmental risk factors have been identified: sleep patterns, psychosocial stress, alcohol consumption and smoking patterns[5,6].

In addition to environmental determinants, a substantial portion of the inter-individual variability in obesity is caused by genetic differences. Twin and family studies provided early estimates of the genetic basis for obesity, which indicated that 40-80% of the variation in obesity-related traits observed in the population could be attributed to genetic differences[7]. Despite the large genetic contribution, identifying the specific sources of genetic variation remains a challenge. Single gene mutations that are individually sufficient to cause obesity are less common and the majority of obesity cases appear to be caused by the combined influence of several common gene variants, each with a modest effect on weight gain[8-11]. Technological advances such as genome-wide association studies (GWAS) have led to the identification of over 160 polygenic obesity loci, yet these variants only explain a portion of the variability attributed to genetic differences[12-14]. Gene-environment interactions have been proposed as a possible

explanation to account for the remaining variability not explained through classical GWAS[15]. Several environmental factors, including physical activity, diet patterns, alcohol consumption and psychosocial stress have been found to moderate the impact of obesity genes in various ethnic groups[16-19].

To build on the existing literature, we undertook the current project to advance the understanding of gene-environment interaction studies in the obesity field. The specific objectives of this thesis were to (1) summarize the findings supporting gene–environment interaction in obesity from heritability, monogenic and polygenic studies, and provide a biological hypothesis to explain these statistical interactions, (2) examine the interaction between obesity predisposing gene variants and sedentary behaviour after adjusting for physical activity, (3) analyze a novel gene-environment interaction between obesity predisposing gene variants and alcohol consumption, (4) study the effects of a cumulative environmental risk score on the impact of obesity risk variants, and (5) outline the methodological challenges associated with gene–environment interaction (GEI) studies, and provide potential solutions to these issues based on existing evidence and highlight future directions for gene-environment interaction research.

CHAPTER 2-LITERATURE REVIEW

ABSTRACT

The worldwide obesity epidemic has been mainly attributed to lifestyle changes. However, who becomes obese in an obesity-prone environment is largely determined by genetic factors. In the last twenty years, important progress has been made in the elucidation of the genetic architecture of obesity. In parallel with successful gene identifications, the number of gene-environment interaction studies has grown rapidly. This paper reviews the growing body of evidence supporting gene-environment interactions in the field of obesity. Heritability, monogenic and polygenic obesity studies provide converging evidence that obesity-predisposing genes interact with a variety of environmental exposures including physical activity and diet patterns. Genetic predisposition to obesity-related traits may be amplified in the context of specific disorders such as depression or preexistent overweight. Gene variants also interact with obesity treatments. What follows in this review includes (1) an introduction to the study of gene-environment interactions, (2) the evidence of gene-environment interactions within the field of obesity and (3) an outline of the biological mechanisms that may explain these interaction effects. Thus far, this growing body of evidence has provided a deeper understanding of gene-environment interactions influencing obesity and may have tremendous applications in the emerging field of personalized medicine and individualized lifestyle recommendations.

INTRODUCTION

Obesity has been defined as the accumulation of fat mass that is sufficient to negatively impact health and decrease life expectancy[20]. In epidemiological studies, obesity is typically classified based on the body mass index (BMI) guidelines established by the world health organization (WHO)[21]. BMI is calculated as weight in kilograms (kg) divided by height in meters (m) squared, and those with a BMI greater than 30 are classified into one of the three obesity categories: obese class 1 BMI= 30.00-34.99, obese class 2 BMI= 35.00-39.99, obese class 3 BMI ≥ 40.00[22]. BMI is closely associated with body fat [21,23] and increased BMI is an established risk factor for stroke[24], ischaemic heart disease[25] and cancers of the large intestine, kidney and endometrium[26,27]. Each 5kg/m^2 increase above a BMI of 25kg/m^2 is associated with a 30% higher risk of all-cause mortality[28]. Average BMI is increasing by a few percent per decade in many populations[29], and no specific cause of death is inversely associated with BMI[28]. Previous estimates project a 33% increase in obesity prevalence and a 133% increase in severe obesity prevalence (obese class 3) over the next 20 years[30]. Individuals with a BMI above 40kg/m^2 are at a significantly greater risk for diabetes and other adverse medical conditions than those with a BMI in the range of 30-35[29]. These individuals also have a much shorter life expectancy and incur greater lifetime medical costs[31].

Over the past three decades, the prevalence of obesity has reached epidemic proportions throughout the world[1]. This recent epidemic cannot be explained by sudden changes in the human population gene pool and has been mainly attributed to lifestyle modifications[32]. Over-nutrition and decline in physical activity are the two "usual suspects", but additional factors (reduced gut microflora diversity, sleep debt, endocrine

disruptors, reduction in variability of ambient temperatures) have emerged as significant contributors to the escalating prevalence of obesity[33]. If obesity is a multifactorial disorder that requires environmental influences to manifest, some individuals are more susceptible than others to weight gain in an obesity-prone environment, and who becomes obese at the individual level is largely determined by genetic factors[34]. Technological and methodological breakthroughs in the last twenty years have led to important progress in the elucidation of the genetic architecture of obesity[35]. The first two genes (*LEP* and *MKKS*) associated with a Mendelian non-syndromic or syndromic form of obesity were identified in 1997 and 2000[36,37]. Seven years later, the first common variant (located in the intron 1 of the *FTO* gene) reproducibly associated with polygenic obesity was identified[38,39]. At the time we are writing, over 40 monogenic obesity loci (with or without syndromic features) and 130 polygenic obesity loci have been described, and this list is destined to grow over the coming years[7]. In parallel with successful gene identification efforts, the number of studies on gene-environment interactions has grown rapidly[40]. In this review, we summarize the findings supporting gene-environment interaction in obesity using heritability, monogenic and polygenic studies.

DEFINITIONS

The genetic etiology of obesity can be classified into two categories. First, Mendelian (or monogenic) obesity describes individuals who carry a rare gene variant with a dramatic impact on adiposity [9]. These variants are associated with a high lifetime risk of disease and exhibit a near one-to-one relationship between genotype and phenotype [11,41,42]. Monogenic obesity can be classified as syndromic or non-syndromic. Syndromic obesity refers to Mendelian obesity that co-occurs with a distinct set of

clinical phenotypes, such as mental retardation, dysmorphic features and organ-specific developmental abnormalities [43]. Syndromic forms of obesity result from chromosomal abnormalities or point mutations, which can be autosomal or X-linked disorders [44]. Non-syndromic forms are caused by pathogenic mutations or structural variations in genes involved in the leptin / melanocortin pathway, and are mainly characterized by hyperphagic obesity [35]. While homozygous/compound loss-of-function in monogenic genes from the leptin/melanocortin pathway lead to fully penetrant obesity, a significantly larger proportion of obesity cases occur among individuals carrying heterozygous deleterious coding mutations in these genes[8,45]. These mutations result in non-fully penetrant obesity, yet based on the average penetrance of these genes, they account for a significantly higher proportion of obesity cases than homozygous loss-of-function mutations in obesity genes[46,47]. Second, other cases of obesity can be attributed to the concerted presence of DNA variation in multiple genes, known as polygenic obesity [11]. With respect to body weight regulation, recent simulations estimated that hundreds of variants with small to modest effect may account for the genetic architecture of complex traits such as obesity [48].

The concept of gene-environment interaction in the context of human diseases is not recent and has been discussed since proposed by J.B. Haldane in 1946 [49]. The statistical definition of an interaction between two or more risk factors is simply the coefficient of the product term of the risk factors, also known as effect modification or effect modulation. Interaction is thus measured in terms of departure from a multiplicative or an additive model [50,51]. Alternatively, biological interaction between two factors is defined as their co-participation in the same causal mechanism to disease

development [52]. For statistical evidence of gene-environment interaction to be convincing, it is typically necessary to replicate the findings in additional samples and / or support the evidence with plausible underlying biological mechanisms [50].

Heritability estimates are influenced by the environment

Early indications of the shared influences of genetics and the environment in shaping obesity originated from heritability studies involving environmental exposures among twins. Heritability is the proportion of total phenotypic variability caused by genetic variance in a population. Large pedigree, twin and adoption studies allow the calculation of heritability and they all evidence a strong genetic component in human obesity [53-55]. Before the first obesity gene identification reports, scientists considered the possibility that heritability, a global estimate of genetic predisposition to obesity, may be modulated by specific environments [56]. Specific environmental exposures known to mediate heritability estimates include biological, socio-economic factors and lifestyle factors.

In utero factors have been proposed to modulate offspring's future risk for obesity [57]. The higher estimates of heritability for BMI observed in mother-offspring pairs in comparison with father-offspring pairs suggest a possible modification effect of maternal *in utero* environment on the offspring's genetic predisposition to obesity [58]. Maternal weight gain during pregnancy may interact with genetic factors to render the offspring more susceptible to develop obesity in young adulthood [59].

Genetic influence on obesity may also interact with sex and age. Sex-specific genetic effects on obesity have been observed in adolescents as well as in adults [60,61].

Previous studies of genetic loci linked to fat distribution reported that approximately half of the SNPs identified displayed stronger effects among females than males [62]. Heritability of obesity also varies with age. A previous study of over 12,000 twin pairs reported a heritability of 4-9% for BMI at birth, which increased to more than 50% at 5 months of age [63]. Heritability estimates increase from infancy to childhood [64], from childhood to pre-adolescence [65], from preadolescence to adolescence [66], and reach a plateau during adolescence and adulthood, and then slightly decrease in late adulthood [67]. Longitudinal BMI change from adolescence to young adulthood and from young adulthood to adulthood is a heritable trait, but genetic variants for change in BMI partially overlap with those affecting the level of BMI [68,69]. Moreover, heritability estimates of obesity increase with the severity of obesity status [70].

The investigation of socio-economic factors and lifestyle behaviours has revealed many additional conditions that impact heritability estimates. One may presume that the emergence of a society characterized by food abundance and physical inactivity may increase the impact of environment (and therefore decrease the impact of genes) in the determination of the obese phenotypes. Counter intuitively, the proportion of variability in BMI attributable to genetic variation is increased among people born after the establishment of a modern 'obesogenic' environment [71-73]. These results are congruent with the seminal work by Claude Bouchard and colleagues showing that the BMI response to long-term overfeeding in young adult male twins is mainly influenced by genetic factors [56]. Twin studies have shown that a high level of physical activity can substantially reduce the influence of genetic factors on BMI in both young and older adults [74,75]. PT. Williams studied the parental contribution to offspring's BMI in 47,691

adult runners and showed that vigorous physical activity (running distance ≥ 9 km/day) decreased the parental contribution to BMI, by 48-58 %, in comparison with runners with moderate physical activity (running distance < 3 km/day)[76]. Socio-economic research indicates that higher educational status is associated with decreased risk of obesity [77], but heritability estimates for BMI in late childhood/adolescence are positively correlated with the level of education of parents [78]. Sleep duration is negatively associated with obesity [79]. In a twin study, the heritability of BMI (h^2 = 70%) in short-sleepers (< 7 hours/day) was more than twice the heritability of BMI (h^2 = 32%) when sleep duration was longer (≥ 9 h/day) [80]. Weight gain is a well-known adverse effect of antipsychotic medication [81], but a considerable degree of inter-individual variability has been described in literature [82]. Two pilot twin/sibs comparison studies have reported heritability estimates of 60-80% for body weight gain in response to antipsychotics in adolescents and adults [83,84]. Weight loss in response to vigorous exercise, diet restriction or bariatric surgery is also highly variable which suggests a heritable component [85-87].

A recent analysis of the Framingham Heart Study analysed how the heritability of BMI was influenced by historical period, life course and physical activity [88]. These authors reported that: 1) the heritability estimates of BMI were considerably larger after the mid 1980's compared to the 3 preceding decades; 2) the genetic influence on BMI appears to decrease across the lifespan, with the greatest genetic influence observed during reproductive ages across historical period and 3) the heritability of BMI was considerably smaller among physically active individuals aged 21-50 years, but not among those >50 years old [88].

Obesity predisposing gene variants interact with the environment

Although heritability studies provided early evidence for the genetic contribution to obesity, recent efforts have focused on the identification of specific gene variants that impact obesity risk. Our knowledge about the genetic architecture of Mendelian (syndromic and non-syndromic) and polygenic forms of obesity has greatly expanded in the last 20 years [35]. It is noteworthy that even some forms of Mendelian (syndromic and non-syndromic) obesity can display a somewhat variable phenotype [89-92]. This can be attributed not only to genetic heterogeneity, gene-gene interactions and inheritance model [93,94], but interactions with environmental factors should be considered as one of the causes for the variability in obese phenotypes [51]. Since the rapid increase in obesity prevalence over the last few decades indicates a strong environmental influence on BMI (e.g. physical activity, diet, educational status, age, sex) [33], many researchers have worked on the identification of specific environmental factors that interact with monogenic and polygenic obesity predisposing genes. The existing evidence regarding the study of obesity indicates that lifestyle factors can significantly modify the impact of obesity predisposing gene variants.

Obesity predisposing gene variants interact with non-modifiable biological factors

Obesity predisposing gene variants interact with pregnancy and in utero factors

Pre-pregnancy maternal obesity and excessive weight gain during pregnancy are both associated with increased birth weight, higher rate of macrosomia in the offspring [95,96] and higher risk of adiposity in offspring during childhood, adolescence and adulthood [57,97,98]. Recently, a morbidly obese female patient with a rare homozygous *LEPR* mutation was reported to gain 110 lbs during pregnancy, far beyond the 11-40 lbs

gestational weight gain range recommended by the Institute of Medicine, and gave birth to a baby with macrosomia [99]. These data suggest that a Mendelian predisposition for obesity increases gestational weight gain and offspring's birth weight. However, no such effect on gestational weight gain was observed for a polygenic gene score composed of four common obesity-predisposing common variants in or near *FTO*, *MC4R*, *TMEM18* and *GNPDA2* [100]. Studies with gene scores including more SNPs are needed to further investigate this hypothesis.

Prenatal exposure to maternal cigarette smoking was found to interact with genetic variation in *OPRM1* to modulate fat intake in offspring [101]. Among 956 adolescents, the T allele in *OPRM1* was associated with lower fat intake but only in those without prenatal exposure to cigarette smoke [101]. DNA methylation was significantly reduced within several CpGs across *OPRM1* among adolescents exposed to prenatal maternal cigarette smoking compared to those not exposed [101].

Obesity predisposing gene variants interact with sex

Females are generally more likely to develop morbid obesity than males [102] and these discrepancies may be explained in part by sex-specific genetic effects. In line with this hypothesis, pathogenic monogenic mutations in *MC4R* have an effect on BMI about twice as strong in females as in males [46,103]. Seven out of 14 loci convincingly associated with waist-to-hip ratio displayed sexual dimorphism, all with a stronger effect on the phenotype in women than in men [104,105]. A recent genome-wide interaction meta-analysis did not report any sex-specific for variants associated with BMI, but found 44 loci with

sex-specific effects on waist-to-hip ratio adjusted for BMI (28 of the 44 loci displayed larger effects in women than men) [106].

Obesity predisposing gene variants interact with age

The syndrome of Prader-Willi has two distinct phenotypic stages. In infancy, it is characterized by poor suck, feeding problems and failure to thrive, followed by hyperphagia in later childhood that leads to excessive weight gain [107]. Rare deletions in the region p11.2 of the chromosome 16 have been associated with a highly penetrant mendelian form of obesity with additional developmental features [108]. These individuals generally have early feeding and growth difficulties, and start to gain excessive weight around 5-6 years of age. As a result, an incomplete penetrance for childhood obesity but a complete penetrance for adult obesity has been observed for the carriers of the chromosome 16p11.2 deletion [108,109]. The longitudinal study of adult *MC4R* mutation carriers show an increasing age-dependent penetrance [46]. The life-course analysis of the intronic *FTO* gene variants and BMI in longitudinal studies indicates that this polygenic obesity-predisposing variant is negatively associated with BMI during infancy (age: 0-2.5 years) but positively associated with BMI from the age of 4 years, with an age-dependent increase during childhood, adolescence and young adulthood [65,110-113]. Most of the effect of the *FTO* intron gene variants on BMI gain occurs during this period, and no appreciable effect of *FTO* on BMI increase is observed during adulthood and agedness [112,114-117]. Studying the association of an obesity gene score from multiple markers in longitudinal cohorts provided similar results: the genetic predisposition score displayed a moderately positive association with birth weight, and more strongly associated with

BMI gain during early infancy and childhood, but no association with BMI change during adulthood was observed [118-120]. A negative genotype x age interaction between the *PCSK1* rs6232 SNP and obesity traits was observed in two independent studies, and a recent meta-analysis of up to 331 175 individuals confirmed this result and identified a similar interaction between age and the *PCSK1* rs6235 SNP on obesity [121-123]. A genome-wide interaction meta-analysis also identified 15 BMI loci with age specific effects, 11 of which showed greater effect sizes in younger (<50 years) compared to older (≥ 50 years) adults [106].

Obesity predisposing gene variants interact with lifestyle factors

Obesity predisposing gene variants interact with an obesity-prone environment

The promotion and globalization of societal changes leading to an imbalance between calorie intake and calorie expenditure partly explain the current obesity epidemic, but interactions between genes and this obesity-prone environment also contribute to the development of obesity. Dudley et al. reported a significant cohort effect on the prevalence of obesity in Prader-Willi syndrome [91]. Prevalence of obesity was higher in patients born after 1990 than before [91]. A generation-dependent penetrance of *MC4R* pathogenic monogenic mutations on obesity was also found in multigenerational pedigrees, with the effect of mutations on the obesity phenotype being amplified by the emergence of an "obesogenic" environment [46,124]. This trend is supported by a recent analysis of the Framingham Heart Study (FHS), which demonstrated that risk allele carriage in *FTO* rs9939609 was associated with a greater increase in BMI among individuals born after 1942 compared to those who were born before 1942 [125]. The *FTO*

intron 1 variant is weakly associated with BMI in South Asian Indian populations, but its effect on weight is stronger in urban compared to rural dwellers [126,127]. A lack of association of *FTO* with obesity-related traits was also observed in a Gambian rural population [128]. The authors speculate that the impact of genetic variance in *FTO* rs9939609 on BMI may be marginal in lean populations where excess food is scarce, compared to populations where food is abundant [128]. Lastly, the growing influence of obesity predisposing genes in 'obesogenic' environments has also been supported by the positive interaction between birth year and the impact of 32 obesity predisposing genes [129]. Together, these data suggest a stronger influence of genetic factors on obesity in obesity-prone environments.

Obesity predisposing gene variants interact with physical activity

Recent data indicate that genetic predisposition to obesity can be blunted in part through physical activity. Over twenty independent studies reported an interaction between the *FTO* obesity risk genotype and physical activity on BMI variation or obesity in children, adolescents and adults [16,18,76,130-149]. An interaction between *FTO* intron 1 variant and the level of physical activity on obesity was recently confirmed in a meta-analysis of 218 166 adults where physical activity attenuated the odds of obesity by 27% conferred by the variant [150]. No such interaction was found in 19 268 children and adolescents [150]. We recently studied more accurate surrogates of physical activity and adiposity in a multi-ethnic study of 17 423 participants recruited in 17 low-, middle- and high-income countries and we observed that the effect of *FTO* rs1421085 on the variation of body adiposity index was reduced by 56% in the higher versus lower metabolic

equivalent score tertiles [151]. Similar results were obtained for a genetic predisposition score combining the information of 12 obesity-associated SNPs, and a high level of physical activity was associated with a 40% reduction in the genetic predisposition to obesity in adults (N=20 430), as well as for BMI level and BMI change across time [137]. Physical activity was also found to attenuate the effect of a 28 SNP obesity gene score on BMI among a sample of East Asians and Europeans [152]. We did not evidence any significant interaction between the quantitative level of physical activity and a 14 SNP obesity gene score on BMI or body adiposity index in an international multi-ethnic study of 17 423 participants [151]. Our data, consistent with the conclusions of a recent meta-analysis in participants of European ancestry living in North America and Europe, suggest that the benefits of being physically active may be optimal in genetically predisposed people living in the more sedentary countries [151,153].

A number of recent studies have analysed the interaction between sedentary behaviours and genetic risk for BMI, independent of physical activity level [138,154,155]. The initial report by Qi et al demonstrated that prolonged television watching accentuated the impact of a 32 SNP genetic risk score on BMI [138], and a second study of an adolescent sample reported that screen time increased the impact of two SNPs (*FLJ35779, GNPDA2*), although these interactions were ethnic specific and of nominal significance [154]. The most recent study of this interaction analyzed how total sitting time impacted the association between *FTO* rs9939609 and BMI among the Framingham Heart Study (FHS) and the Women's Health Initiative Study (WHI), but the results were not significant [155].

Obesity predisposing gene variants interact with diet

Rouskas et al. reported that the penetrance of *MC4R* loss-of-function heterozygous mutations on obesity is exceptionally low (6.3 %) in the Greek population, in comparison with those observed in other European countries (60-100%)[156]. A possible explanation of this 'Greek paradox' may be a protective effect of the Mediterranean diet against MC4R deficiency-induced obesity [156].

Several studies have characterized the impact of diet patterns on genetic predisposition to obesity. Independent cross-sectional and longitudinal samples of Caucasians and Latin Americans, suggest that a high daily energy intake, high fat intake or high saturated fat intake can amplify the effect of the *FTO* genotype on obesity risk in children, adolescents and adults [141,157-163]. Higher intake of fried foods has also been shown to increase the impact of a 32 SNP gene score (and an *FTO* variant individually) on BMI over follow-up [164]. These interactions were replicated in an independent cohort of 21 421 women [164]. A recent 25 year follow-up study in Australia reported an interaction between rs9939609 in *FTO* and diet on BMI change [165]. The prudent/healthy diet was associated with a greater BMI change among AA compared to TT genotypes. This interaction was observed at 17 years of follow-up, but was restricted to females [165]. Despite the many studies demonstrating that diet patterns can moderate the genetic risk for obesity, two recent meta-analyses did not detect significant interactions between diet patterns and obesity-associated gene variants [166,167]. Data from 177 330 adults (87% Whites, 10% Asian, 3% African American) did not indicate any significant interactions between the *FTO* variant and dietary intake of total energy, fat, protein or carbohydrate on BMI[167]. A second meta-analysis of 68 317 Europeans did not detect an interaction

between a 32 SNP genetic risk score and a multifactorial diet score on BMI [166]. The diet score moderated the impact of two SNPs on BMI (*LRRN6C* and *MTIF3*), although these effects were nominal and the impact of these risk variants appeared to be greater among those consuming healthier diets [166]. The authors speculate that the broad diet assessment in their analysis may have masked interactions of varying directions and magnitudes that were identified in previous studies [166].

The Apolipoprotein A-II (*APOA2*) -265 T>C promoter functional polymorphism appears to interact with high-saturated fat to increase BMI and obesity risk in several independent populations (Mediterranean, Asian, Caucasian, Hispanic and Carribean) [168,169]. High saturated fat intake was associated with significant increases in the genetic risk for obesity across populations [170]. Specifically, the C allele homozygotes with high saturated fat intake displayed a 1.84 (95% CI, 1.38-2.47) odds of obesity compared to a 0.81 (95% CI, 0.59-1.11) odds in those with low saturated fat intake [170]. A separate analysis of 1 225 obese adults demonstrated the C allele homozygotes with a high saturated fat intake ($\geq$ 20.7 g/day) had higher waist circumference values than individuals with any other genotype in the high saturated fat intake group [171]. While the underlying biological mechanism explaining this association is not fully understood, the *APOA2* -265 T>C SNP has been associated with obesity risk eating behaviours such as meal skipping, and dietary modulation of plasma ghrelin [171].

The Apolipoprotein A5 protein influences plasma triglyceride concentrations in humans and regulators of the *APOA5* gene (peroxisome proliferator-activated receptors, insulin, thyroid hormone) have been implicated in obesity risk [172,173]. In a weight loss study of 606 men with hyperlipidemia, C allele carriers of the -1131 T>C variant in the

APOA5 gene displayed significantly greater BMI reduction while on a fat restriction diet [174]. Additional evidence from the Framingham risk study suggests that carriers of the mutant C allele may have a lower risk of obesity compared to T allele homozygotes when consuming a diet high in monounsaturated fats [175]. This interaction was also tested in a Mediterranean sample and greater fat intake was associated with obesity among T allele homozygotes while no association was observed among carriers of the mutant C allele [176]. These studies suggest that the C allele in *APOA5* may have a protective effect against obesity among individuals consuming a high fat diet.

Several studies have examined the interaction between diet patterns and *PPARG* Pro12Ala polymorphism with regards to obesity [177], [178]. The risk allele (12Ala) has not been consistently shown to increase obesity risk and the heterogenous impact of this mutation indicates the potential for interaction effects [179]. A study of 720 French Canadians found that higher amounts of saturated or total fat consumption were associated with greater waist circumference in Pro allele homozygotes but not in 12Ala carriers [178]. Similar results were observed from studies analysing BMI. An investigation of the Nurse' Health Study demonstrated that high total fat intake was associated with greater BMI among participants homozygous for the Pro allele but not among 12Ala carriers [180]. This study also reported that monounsaturated fat intake was associated with decreased BMI among 12Ala allele carriers and this interaction was replicated in an independent weight loss study [181]. Additional interaction studies of *PPARG* related to body composition have shown sex-specific effects [182,183] and weight change analyses have reported inconsistent results [181,184-187].

A more recent analysis found an interaction between an eight SNP obesity gene score and mono and polyunsaturated fatty acid intake [188]. Among 2 346 children with low unsaturated fat intake, the gene score was associated with increased body fat mass index yet no association was present among unexposed children [188].

Obesity predisposing gene variants interact with psychosocial stress

A recent genome-wide interaction analysis identified a significant interaction between psychosocial stress and five SNPs within the *Early B-cell Factor 1* (*EBF1*) gene and hip circumference [189]. The interaction reached genome-wide significance among the subset of 2460 Whites in the Multi-Ethnic Study of Atherosclerosis (MESA) but was not significant among Chinese Americans, Blacks or Hispanics. This study reported that the impact of risk allele carriage in *EBF1* on hip circumference was greater among participants with a greater chronic stress burden [189]. The authors also replicated the interaction between psychosocial stress and three of the original five SNPs (rs17056278 C>G, rs17056298 C>G, and rs17056318 T>C) in *EBF1* in the Framingham Offspring cohort[189]. A subsequent analysis by the same research group replicated the *EBF1* x psychosocial stress interaction on obesity (waist circumference or BMI) in the Family Heart Study Whites and at trend level in the Duke Caregiver study [190]. The direction of the interaction effect was consistent across each of the studies: chronic psychosocial stress amplified the effect of *EBF1* variation on BMI [190].

Obesity predisposing gene variants interact with educational status

Epidemiological studies have shown an association between a low level of education and higher risk of overweight and obesity [77]. A significant negative association between BMI and educational status was found in non- carriers of *MC4R* mutations but not in *MC4R* loss-of function mutation carriers issued from the same pedigrees [46]. These results show that a high level of education has no protective effect on obesity risk in presence of *MC4R* pathogenic mutations. On the contrary, a significant gene x education interaction has been found in the intron 1 variant in *FTO*, the significant effect of the SNP on BMI and obesity risk restricted to subjects with no university education [191]. This finding is supported by a recent study of European children (N=16 228) indicating that favourable socioeconomic status is protective against obesity, yet this effect was only observed in participants with the low risk genotype TT in *FTO* rs9939609 [192].

Obesity predisposing gene variants interact with smoking status

A meta-analysis of nine European study samples (N=24 198) demonstrated that smoking status moderated the association between genetic variation at the *CHRNA5-CHRNA3-CHRNB4* locus (rs1051730) and BMI [193]. While there was no evidence of association between variation at rs1051730 and BMI in never smokers, each additional risk allele (T) was associated with a BMI decrease of 0.16 and 0.33 kg/m^2 among former and current smokers, respectively [193]. A separate study of 14 131 Pakistani adults reported another gene x smoking interaction: the minor allele (T) in *FLJ33534* was associated with lower BMI in current smokers and positively associated with BMI among adults who had never smoked [194]. A number of gene x smoking interactions were seen

when African Americans and Caucasians were analysed separately, but no significant interactions were observed in the overall sample from the Southern Community Cohort Study [195]. Four nominally significant gene x smoking interactions were reported in a recent study of 16 157 Pakistani adults, with current smoking status amplifying the effect of *PTBP2* rs11165643, *HIP1* rs1167827 and *GRID1* rs7899106 SNPs, and decreasing the effect of *C6orf106* rs205262 SNP [147].

Obesity predisposing gene variants interact with alcohol consumption

Among a sample of 3 522 East Asians, increased alcohol consumption was associated with an increase in the effect of a 29 SNP GRS on BMI [149]. Increased alcohol intake was also reported to increase the impact of *PPARGC1A* rs4619879 among African Americans, but this interaction was not significant in Caucasians or in the combined sample [195].

Obesity predisposing gene variants interact with disease status/response to treatment

Obesity predisposing gene variants interact with specific health conditions

Beyerlein et al suggest that pre-existent overweight may double the effect of an obesity genetic predisposition score on body fat mass in 4 837 European children [196]. This association is supported by an independent study of 7 225 children of European ancestry which found that previously identified obesity predisposing loci had a greater impact on BMI among obese children compared to their non-obese counterparts [197]. Similar results were observed in 1 930 adults of European descent [198]. If true, it signifies

that obesity predisposing genes may have an even more detrimental effect on weight gain once overweight/obesity is established. Depression predicts subsequent development of obesity [199] and depression status has been shown to amplify the effect of *FTO* SNPs on BMI [200]. Moreover, obesity has an important role in the etiology of polycystic ovary syndrome [201] and *FTO* intronic SNP has larger effects on BMI in patients with polycystic ovary syndrome than in subject from the general population [202,203].

Obesity predisposing gene variants interact with lifestyle modifications

A strict, fat-reduced, and carbohydrate-modified diet leads to a long-term marked weight reduction in adolescents with Prader-Willi syndrome who are already overweight [204]. Importantly, if diagnosis is made at an early age and intensive diet management starts early, reasonable weight control is achieved in non-obese patients with Prader-Willi syndrome [205,206]. Regular exercise training has beneficial effects on body composition and weight loss in Prader-Willi syndrome patients [207,208], especially as they tend to be less physically active than obese non-syndromic individuals [209]. *MC4R* or *POMC* monogenic patients respond as well as non-monogenic obese patients to hypocaloric dietary or multidisciplinary (exercise, behavior, nutrition therapy) interventions [210,211] but *MC4R* monogenic patients fail to maintain weight loss after intervention [211].

The obesity risk variant rs9939609, an allele in *FTO,* does not modify the weight loss response to lifestyle intervention [212-214] or caloric restriction [215,216], but is associated with lower additional weight loss and higher risk of weight regain during the weight maintenance phase that follows the caloric restriction program [216]. Carriers of the *FTO* intron 1 obesity risk variant experience a higher rate of dropout when they are submitted

to a high-fat/low carbohydrate (in comparison with a low-fat / high carbohydrate) hypocaloric diet [217], but they achieve better weight maintenance than wild-type individuals during a 3-year intervention with a Mediterranean diet [218]. *FTO* variation has also been shown to interact with protein intake to moderate the response to weight-loss interventions [219,220]. A two-year randomized control trial (RCT) found that higher protein intake was associated with improved weight loss and body composition among carriers of the *FTO* rs1558902 risk allele compared to non-carriers [219]. Another analysis of the same trial (N >700) showed that individuals with the *FTO* rs9939609 risk (A) allele achieved more favourable changes in food cravings and appetite when consuming a higher-protein weight-loss diet [220]. Carriers of the *FTO* obesity risk alleles also lose less weight in response to exercise training [221,222]. Among five childhood obesity susceptibility loci identified in a French-German genome-wide association studies (GWAS) meta-analysis [223], only one (*SDCCAG8*) was associated with differential weight loss after lifestyle intervention in 401 children and adolescents [224]. Eight out of 15 obesity predisposing gene variants recently identified by GWAS showed trends of association with weight loss or weight regain during lifestyle intervention in 3 356 adults of the Diabetes Prevention Program [225].

The *PLIN* gene has received increasing support for its role in obesity risk and insulin resistance and genetic variation at the perilipin locus has been shown to interact with diet behaviours [226-228]. Two separate weight loss studies ranging from 12 weeks [229] to 1 year [227] in length have shown that A carriers of the 11 482 G>A SNP at the perilipin locus lost less weight compared to non-carriers. Other studies report a gene-diet interaction involving the same locus and carbohydrate intake [230]. Among 920 participants

consuming a high amount of complex carbohydrates (≥144 grams/day), minor allele carriers displayed reduced hip and waist circumference compared to major allele homozygotes [230].

Obesity predisposing gene variants interact with therapeutic treatment

As most obese persons are resistant to the weight-reducing effects of leptin, administration of recombinant leptin to obese subjects does not generally result in significant weight loss [231]. However, patients with congenital leptin deficiency markedly reverse obesity and associated phenotypic abnormalities when they are treated with daily injections of recombinant human leptin [232,233]. Leptin administration reduces energy intake, fat mass, hyperinsulinemia, and hyperlipidemia, restores normal pubertal development, endocrine and immune function and improves neurocognitive performances [234]. Although patients with complete leptin deficiency are extremely rare, leptin supplementation may eventually help a far greater number of obese patients with partial leptin deficiency (heterozygous for a loss-of-function mutation in the *LEP* gene) based on the observation that leptin therapy induces more significant weight loss in subjects with low leptin levels [235,236].

The guanine nucleotide binding protein beta polypeptide 3 (*GNB3*) C825T functional gene variant predicts that obese individuals will benefit more from the anti-obesity drug sibutramine treatment. Sibutramine is a serotonin and norepinephrine reuptake inhibitor and given that *GNB3* variance is associated with an altered response to G protein subunit activation (α 2-adrenergic activation) [237], there is biological evidence to support this interaction. Two independent studies showed that the carriers of the 825 T

allele lose more weight in response to sibutramine administration than C allele homozygotes [238,239]. Lastly, obesity predisposing gene variants in *FTO* and *MC4R* are associated with more weight gain in response to antipsychotic treatments [240-244].

Obesity predisposing gene variants interact with bariatric surgery

Bariatric surgery is the most effective long-term treatment for severe obesity [245]. However, this surgery is not recommended for morbidly obese patients with Prader-Willi syndrome due to their inability to understand the necessary operative and follow-up procedures, altered pain threshold, inability to vomit and the potential development of gastric dilation/necrosis. Therefore, an alternative approach including the use of supervised reduced-energy diets with vitamin/mineral supplementation, restricted access to food, and a daily exercise regimen may be more adequate [246]. Laparoscopic adjustable gastric banding did not result in a long-term weight reduction in an 18-year-old patient with complete MC4R deficiency [247], and was associated with a high risk of conversion to bypass operations in individuals with partial MC4R deficiency [248]. On the contrary, three studies confirmed that Roux-en-Y gastric bypass surgery was an efficient strategy to lose weight in *MC4R* mutation carriers [249-251]. These results suggest that diversionary operations, which are more invasive but efficiently improve the neuro-hormonal control of satiety than gastric banding procedures, be recommended in the context of non-syndromic monogenic forms of hyperphagic obesity.

FTO risk allele carriers lose less weight than common allele homozygous individuals after banding surgery [252,253], but experience a similar level or more weight reduction after gastric bypass surgery [252,254].

Biological processes underlying statistical gene-environment interactions

Epigenetic changes are believed to be a primary mechanism explaining interactions between environmental exposures and genetic variation [255,256]. Epigenetics is defined as heritable changes in gene function that cannot be explained by changes in the deoxyribonucleic acid (DNA) sequence, and the three main types of epigenetic modification in mammals include DNA methylation, histone modification and non-coding RNA [256,257]. Of these mechanisms, DNA methylation has received the most attention in human studies, and in mammals this process mainly occurs at CpG dinucleotides [258]. Specifically, covalent bonding of a methyl group to the cytosine base creates a barrier that inhibits transcription factors from binding to the DNA helix [258,259]. CpG DNA methylation at gene promoters is typically associated with gene silencing, whereas CpG methylation in gene bodies is linked to gene activity [258]. Given that epigenetic differences have been linked to obesity status, as well as genetic variation and a variety of pre and postnatal environmental factors these processes likely represent a plausible mechanism of gene-environment interactions.

The emergence of new approaches to study epigenetic variation, such as epigenome-wide association studies (EWAS), has led to the identification of methylation patterns associated with obesity [260]. Increased BMI among adults was found to be positively associated with increased methylation at the hypoxia-inducible factor 3 alpha (*HIF-3α*) locus in blood cells and adipose tissue [261]. This finding was confirmed in two replication cohorts in the initial analysis [261], as well as two additional independent studies [262,263]. A separate EWAS of an African American sample identified an association

between methylation at 37 CpG sites and BMI, which were replicated in two cohorts of European ancestry [262]. Analyzing whole-genome DNA methylation and expression data in human adipose tissue from 96 males and 94 females revealed that DNA methylation and expression of 2 825 genes was correlated with BMI [264].

Existing evidence also supports the association between environmental exposures on DNA methylation patterns. Monozygotic twins, who are epigenetically indistinguishable at birth, exhibited drastically different overall content and genomic distribution of DNA methylation and histone acetylation in later life [265]. Moreover, methylation and expression of 1 050 genes have been found to vary with age [264,266]. The epigenetic divergence that occurs with aging likely reflects the accumulation of environmental exposures that influence methylation patterns. Prenatal factors including maternal BMI and variations in maternal methyl donor intake during pregnancy have been linked to methylation changes in the offspring [267], and multiple studies have shown that maternal vitamin B12, folate and cobalamin levels during pregnancy are associated with offspring adiposity [268,269]. Folate, vitamin B12 and choline are methyl donors and involved in the synthesis of methionine, the precursor of the universal donor of methyl groups needed for DNA methylation (S-adenosylmethionine). As a result, disregulation in any of these components can alter the epigenomic regulation of gene expression [270]. With respect to postnatal determinants of DNA methylation, exercise interventions have been shown to alter the DNA methylation of 2 817 genes in skeletal muscle and 7 663 genes in adipose tissue (18 of which were obesity candidate genes) [271]. The effect of exercise on DNA methylation appeared to be tissue specific, with the majority of genes in skeletal muscle displaying decreased DNA methylation [272], and the majority of genes in

adipose tissue showed increased DNA methylation [273]. These changes mirrored the patterns observed for gene expression: most of the genes showing concurrent changes in DNA methylation and expression displayed increased expression in skeletal muscle and decreased expression in adipose tissue [271]. A recent review of 25 studies (16 observational 9 interventional) found that both acute and chronic exercise significantly influenced DNA methylation, and these changes occurred in a tissue- and gene-specific manner [274]. DNA methylation changes have also been observed in response to high-fat intake [275-278], and after weight loss interventions the methylation profiles of adipose and muscle tissue among those formerly obese became more similar to lean individuals [279-282]. These methylation changes involved a number of known obesity-associated loci, including *LEPR, STAB1, ZNF608, HMGA1, MSRA, TUB, NRXN3, FTO, MC4R* and *BDNF* [280,281].

Although environmental factors have the potential to influence the epigenetic environment, it is estimated that approximately 20-40% of epigenetic variation can be attributed to genetic differences [256,283,284]. Early evidence demonstrated that the risk allele of *FTO* promotes increased methylation of sites within intron 1 of the *FTO* gene, as well as greater methylation of additional genes [285]. Other evidence identified 28 obesity-associated SNPs that were associated with differential methylation at 107 proximal CpG sites [286]. A recent study of Trim28 haploinsufficiency used findings from mice and humans to demonstrate the variation in obesity phenotypes that can be induced through epigenetic changes [287]. These authors also reported that *FTO* expression was decreased among Trim28_Low obese children compared to Trim28_Low lean individuals [287].

These epigenetic findings support an emerging mechanistic model to explain gene-environment interactions in obesity. Existing evidence indicates that obesity, pre and postnatal environmental factors and multiple obesity associated gene variants are linked with epigenetic patterns (Figure 1). Given that gene variants, such as those in *FTO*, and environmental factors both play a role in the methylation of obesity genes, the balance between these two effects likely impact the manifestation of genetic obesity. This biological model provides support for many of the statistical interactions reported to date, and further integration of genomic, epigenomic and transcriptomic data with gene-environment interaction studies will aid in uncovering these biological mechanisms.

Conclusion

Heritability, syndromic, monogenic and polygenic obesity studies provide converging evidence that obesity predisposing genetic factors strongly interact with environment, from birth to agedness and in a wide range of situations. Emerging epigenetic studies have demonstrated that obesity, genetic variants and environmental exposures can influence DNA methylation, which provides a mechanistic model to support the statistical interactions from genetic epidemiology. A comprehensive understanding of gene-environment interactions in obesity may lead to tremendous applications in the emerging field of personalized medicine and individualized lifestyle recommendations. Evidence from interaction studies suggests that specific subgroups of individuals may have an increased risk to develop obesity in specific environments but may also benefit more from lifestyle interventions, a treatment or a surgical procedures

[288]. This information will help determine if population-wide or personalized subgroup interventions are the best suited to fight the worldwide obesity epidemic [289,290].

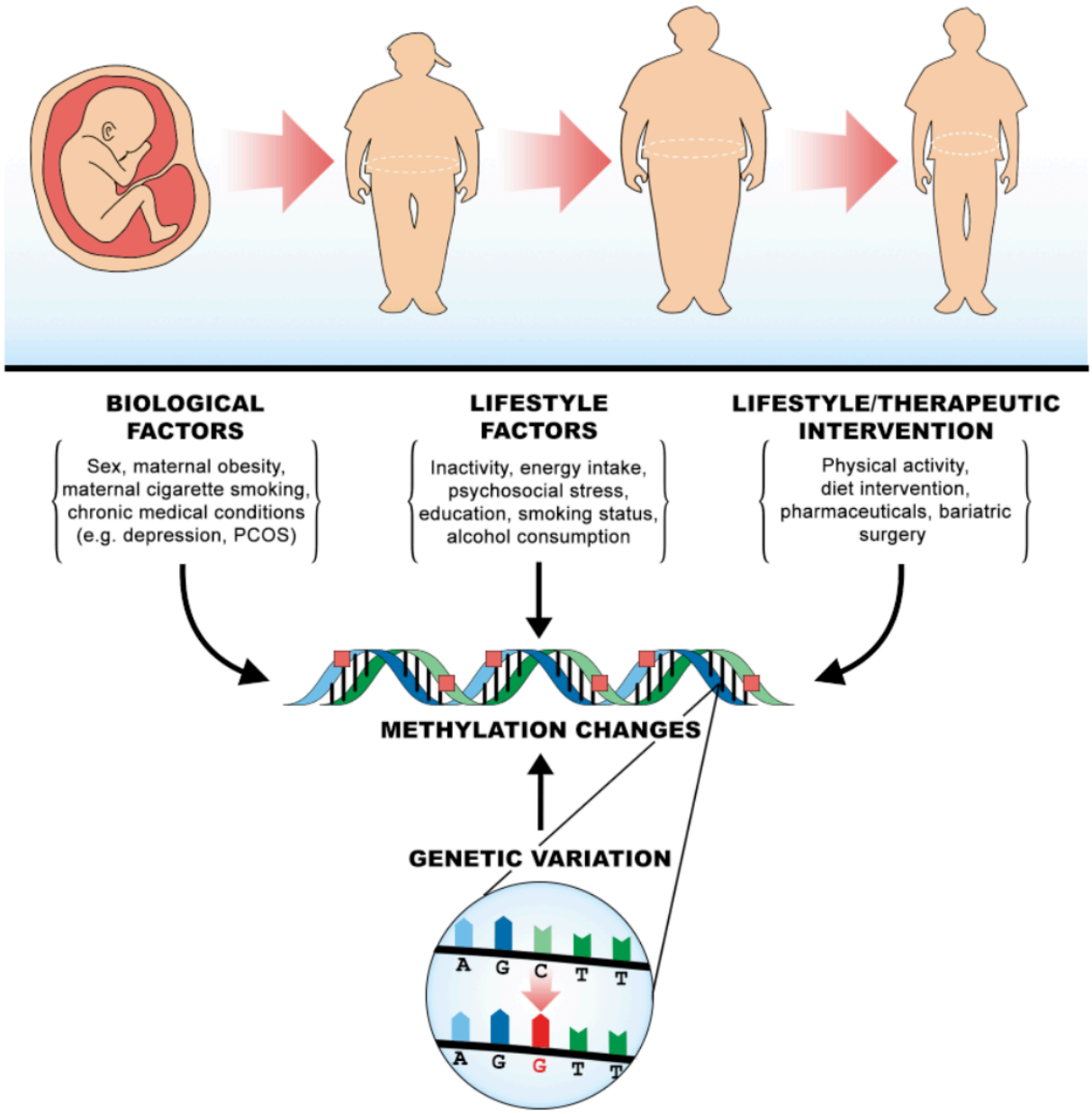


Figure 1. Biological model to explain gene-environment interactions in obesity.

CHAPTER 3- SEDENTARY BEHAVIOUR AND GENETIC PREDISPOSITION TO OBESITY IN A MULTIETHNIC STUDY

ABSTRACT

Background: The objective of this study was to analyze the interaction between sedentary behaviour and 23 obesity predisposing variants (analyzed separately and as a genetic risk score (GRS)) on obesity in an international multi-ethnic cohort.

Methods: The data for this analysis were collected through the multi-ethnic cohort EpiDREAM (16 063 participants from six ethnic groups). Sedentary behaviour was measured as the number of hours per day spent watching television, physical activity was measured as a categorical (low-moderate-high) and quantitative (metabolic equivalent (MET) score) and obesity was measured using the body mass index (BMI).

Results: Increased sedentary behaviour was associated with increased BMI independently of physical activity ($P=1.3 \times 10^{-49}$). *FTO* rs3751812, *CDKAL1* rs2206734, *TNNI3K* rs1514176, *GIPR* rs11671664 *TAL1* rs2984618, *NT5C2* rs3824755, *TCF7L2* rs7903146 and the GRS were associated with BMI ($P=1.5 \times 10^{-12}$). The duration of sedentary time did not moderate the impact of physical activity (measured categorically or quantitatively) on BMI. We did not observe any interaction effects between sedentary time and any of individual SNPs or 23 SNP GRS for BMI.

Conclusion: Our results indicate that increased sedentary behaviour is a distinct risk factor for obesity independent of physical activity level, and increased physical activity does not reduce the impact of sedentary behaviour on BMI. We did not find any evidence of interaction between sedentary behaviour and the 23 SNPs/GRS on BMI.

INTRODUCTION

Obesity has reached epidemic proportions and is estimated to affect over 600 million people worldwide[1] (http://www.who.int/). Average body mass index (BMI) is increasing by a few percent per decade in many populations[29], and greater BMI does not decrease the risk of any specific causes of death[28]. Of particular concern is the projection of a 33% increase in obesity prevalence and a 133% increase in severe obesity prevalence (BMI > 40) over the next 20 years[30]. Obesity is an established risk factor for type 2 diabetes (T2D), cardiovascular disease, certain types of cancer, psychiatric disorders and decreased life expectancy (8-13 year decrease)[2,20]. Above a BMI of 25kg/m^2, each 5kg/m^2 increase is associated with a 30% higher risk of all-cause mortality[28]. The etiology of this disease is complex and involves both genetic and environmental risk factors.

Early evidence from twin and family studies suggested that 40-80% of the inter-individual variation in obesity-related traits observed in the population can be attributed to genetic differences[7]. In addition to Mendelian forms of syndromic and non-syndromic obesity, over 160 loci associated with polygenic obesity have been identified[12]. However, these variants only account for a portion of obesity cases[291,292]. As a result, environmental risk factors are believed to be primarily responsible for the increase in obesity prevalence, with physical inactivity and excessive energy intake being labeled as "the big two"[5]. A 22-year longitudinal study of the National Health and Nutrition Examination Survey (NHANES) cohort reported that physical activity had a greater impact on adiposity measures than calorie intake[293], however the relative contribution of "the big two" is still under debate[294].

The health benefits of physical activity have been recognized for decades[295-297], and exercise is a vital component of weight management[298]. Integrating physical activity as a routine part of daily living is the first recommendation declared by the Institute of Medicine (IOM) to address the obesity epidemic[299]. However, recent epidemiological evidence has focused on the health impacts of sedentary behaviour as a distinct risk factor from inactivity[300]. This distinction has important implications and the Sedentary Behaviour Research Network (SBRN) did not establish consensus regarding the definitions until 2012[301]. Sedentary behaviours are defined as any waking behaviour characterized by an energy expenditure ≤ 1.5 metabolic equivalents (METs) while in a sitting or reclining posture[301]. Alternatively, inactivity refers to adults who fail to engage in at least 150 minutes of moderate to vigorous physical activity (MVPA) (≥ 3 METs) per week, as recommended by the physical activity guidelines issued by the World Health Organization[302] and the U.S. Department of Health and Human Services[301,303]. After adjusting for physical activity, prolonged sitting time is estimated to account for 5.9% of all-cause premature mortality, while physical inactivity accounts for 5.5%[304]. The difference between these exposures is supported by an array of empirical evidence. First, it is possible to accrue large amounts of sedentary time and MVPA in a single day[300,305-307] and there is often modest association between sedentary behaviour and MVPA[308,309]. Second, several studies have highlighted that sedentary behaviour is associated with increased risk of morbidity and mortality regardless of the amount of MVPA[306,310-313]. In conjunction, these results support the distinction between sedentary behaviour and inactivity.

Despite positive trends in leisure time physical activity in some countries, concurrent decreases in incidental, transportation and occupation-related activity have been observed[314-317]. A recent report estimated that 31% of adults and 80% of adolescents worldwide are inactive[318]. With the exception of sleeping and working, TV viewing is the most commonly reported daily activity in many populations throughout the world[311]. Recent estimates suggest that the average number of hours spent viewing TV is 5 hours per day in the United States[319]. Moreover, 6-10% of all deaths from non-communicable diseases worldwide have been attributed to physical inactivity[320]. This corresponds to 5.3 million deaths per year, and one million of these deaths could be averted if the prevalence of inactivity decreased by only 10%[320].

The challenges associated with this issue are magnified by health sequelae of inactivity. A systematic review of prospective cohort studies analyzing inactivity demonstrated a dose-response relationship between TV viewing and increased risk of obesity, T2D, fatal and non-fatal cardiovascular disease and all cause mortality[311]. A large prospective study of 71 363 Danish adults followed for a mean of 5.4 years found similar associations between total daily sitting time and myocardial infarction, coronary heart disease and all-cause mortality[321]. Another recent cross sectional study of 48 882 Norwegian adults found increased sitting to be associated with poorer BMI, waist circumference, total cholesterol, HDL cholesterol, systolic and diastolic blood pressure, non-fasting glucose and gamma glutamyltransferase and triglyceride levels. Importantly, these associations were observed after adjusting for sex, age, education, physical activity, smoking status, fruit and vegetable consumption and general health status[322].

Although many studies have analyzed the impact of sedentary behaviours on metabolic complications, only three studies have investigated the interaction between sedentary behaviours and genetic predisposition to these traits[138,154]. An analysis by Qi et al. indicated that prolonged TV watching may accentuate the impact of a 32 SNP genetic risk score on BMI[138], while a study by Graff et al. reported two nominally significant TV x SNP interactions on BMI[154]. However, the Qi et al. study included two separate cohorts of European men (N=4 564, >40 years old) and women (N=7 740, >30 years old) and did not study interaction effects with individual SNPs, while the second analysis by Graff et al. was restricted to adolescents (N=7 642) and only reported nominal interactions in ethnic subgroups. A recent study of 119 132 participants from the UK biobank analyzed the interaction between *FTO* rs1421085 and several lifestyle factors, yet the interaction with TV watching was not significant[323].

Despite the limited study of gene x sedentary behaviour interactions, significant gene-environment interactions (GEI) between *FTO* intron 1 variation and physical activity have been repeatedly demonstrated in twenty independent studies and a meta-analysis of 218 166 adults[16,18,76,130-146]. A separate meta-analysis of 111 421 found that physical activity also moderates the association between genetic risk scores (GRS) and BMI[324]. These interactions are supported by methylation and expression studies indicating that physical activity can alter the mRNA expression of the *FTO* gene, among others, in muscle and adipose tissue[272,273]. Together, these findings suggest that other activity related behaviours such as TV viewing may interact with genetic variants to influence obesity development. Given the ongoing obesity epidemic and the continual integration of sedentary behaviours into daily living[315,318,319], we undertook the present

study to evaluate 1) the association between sedentary behaviour and physical activity with BMI and 2) the interaction between variation in 23 obesity predisposing variants (analyzed separately and as a GRS), and sedentary time on BMI in the international multi-ethnic cohort EpiDREAM.

METHODS

Study Participants

The EpiDREAM study screened 24 872 individuals aged 18-85 years to be at risk of T2D based on family history, ethnicity, gestational diabetes or abdominal adiposity. Participants were recruited from 17 countries between 1 July 2001 and 2 August 2003 and completed a 75-gram oral glucose tolerance test (OGTT) using a standardized protocol. This analysis focused on 16 063 subjects from six ethnic groups (South Asian, East Asian, European, African, Latin American, Native North American) with both phenotypic and 50K gene-centric array information in the EpiDREAM study (Figure 2). Ethnicity was validated in the 16 063 individuals using the eigensoft software (http://genepath.med.harvard.edu/~reich/Software.htm). If samples failed to cluster with individuals of the same ethnic group they were removed from the analysis. Local ethics committees have approved the EpiDREAM study and informed consent was obtained from each subject before participating in the study, in accordance with the Declaration of Helsinki.

Genotyping

The Gentra System was used to extract the DNA from the buffy coats of 19 498 participants in the EpiDREAM study (Supplementary Figure 1). Genotyping was

performed using the Illumina CVD bead chip microarray ITMAT Broad Care (IBC) array [325] at the McGill University and Genome Quebec Innovation Centre using the Illumina Bead Studio genotyping module, version 3.2. Three resources were used for the key word search (e.g. BMI) to complete the SNP selection procedure: (1) the PubMed database (www.ncbi.nlm.nih.gov/pubmed), (2) the National Human Genome Research Institute (NHGRI) GWAS Catalog (www.genome.gov/gwastudies/) and (3) the HuGE Navigator GWAS Integrator (www.hugenavigator.net/HuGENavigator/gWAHitStartPage.do). This search was performed independently by two individuals (HR and DM) and this strategy resulted in a list of 72 independent SNPs associated with obesity-related traits in GWAS studies. Twenty-three of the 72 SNPs were available on versions 1 and 2 of the IBC 50K SNP array (Supplementary Table 1): rs3751812, rs7203521 in *FTO*, rs1514176 in *TNNI3K*, rs6265 and rs1401635 in *BDNF*, rs1805081 in *NPC1*, rs6232, rs6235 in *PCSK1*, rs2206734 in *CDKAL1*, rs2075650 in *TOMM40/APOE/APOC1*, rs2272903 in *TFAP2B*, rs997295 in *MAP2K5*, rs1211166 in *NTRK2*, rs11671664 in *GIPR*, rs2984618 in *TAL1*, rs1011527 in *LEPR*, rs7605927 in *POMC*, rs611203 in *USP37*, rs2535633 in *ITIH4*, rs3824755 in *NT5C2*, rs7903146 in *TCF7L2*, rs671 in *ADLH2*, rs749767 in *KAT8*. The selected SNPs showed no significant ($P > 10^{-6}$) deviation from Hardy-Weinberg Equilibrium (HWE) in the six ethnic groups. The call rate for the 23 SNPs ranged from 99.8-100% (Supplementary Table 1).

Phenotyping

The screening visit also consisted of a questionnaire that collected information on medical history, physical activity, diet, sedentary time as well as physical measurements

including height, weight, waist and hip circumference using a standardized protocol[326]. Trained medical staff measured standing height to the nearest 0.1 cm and weight was measured to the nearest 0.1 kg in light clothing. Measurements of hip circumference were performed at the level of the greater trochanters and were assessed in duplicate using a non-flexible tape measure with an attached spring balance with a mass of 750g. Averages of the two measures were used in all analyses. Body mass index (BMI) was calculated as weight in kilograms (kg) divided by height in meters (m) squared.

The OGTT data and the 2003 American Diabetes Association criteria were used to categorize participants into one of four categories: (1) normoglycemia= fasting plasma glucose < 5.6 mmol/L, (2) Impaired Fasting Glucose (IFG)= fasting plasma glucose of 5.6 to 6.9 mmol/L, (3) Impaired Glucose Tolerance (IGT)= fasting plasma glucose below 7.0 mmol/L and a 2-h glucose between 7.8 and 11.0 mmol/L, and (4) diabetes was defined if either the fasting plasma glucose was ≥ 7.0 mmol/L or the 2-h glucose was ≥ 11.1 mmol/L [327]. IFG and IGT were collapsed into one category and these three groups (normoglycemia, IFG/IGT, diabetic) comprised the glycemic status variable.

PA was measured using a self-reported rating scale of work-related and leisure time physical activity (1=sedentary, 2=moderately active, 3=very active), hereafter referred to as the basic physical activity score (BPAA). A separate measure of energy expenditure (metabolic equivalent score (MET score)) was also collected using self-reported time of participation in 41 different physical activities (Supplementary Table 2). The MET score was calculated by multiplying the participation time in each activity by the corresponding MET value, and these MET values were summed across all activities

to create a total MET score for each participant. Sedentary time was measured as the self-reported duration of television viewing (hours/day).

Statistical Analyses

All statistical tests were performed using SPSS (version 20) and two-tailed P-values are presented in this manuscript. An additive model was used to perform the single SNP analyses and the obesity risk alleles for each of the 23 SNPs were based on previous literature. The risk alleles of the 23 SNPs were summed to create an un-weighted GRS ranging from 0-46[328]. Ethnic-specific imputations for missing genotype values were performed as previously described[329]. Participants with more than one out of 23 missing genotypes were not included in the GRS calculation.

Prior to conducting the main analyses, two preliminary tests were performed. The association between the SNPs/GRS and BMI were analyzed using linear regression models to restrict the interaction tests to SNPs / GRS with a significant association with the outcome. A P <0.05 was considered significant since existing literature has demonstrated associations between these SNPs and BMI. These tests were adjusted for sex, age, ethnicity (principal components analysis [PCA]), and glycemic status. All SNPs/GRS below the significance threshold were carried forward to the interaction analysis. As a second step, GLM were used to analyze the impact of the 23 obesity predisposing SNPs and GRS on sedentary behaviour, to ensure that the two interacting terms were independent. This test was performed while controlling for sex, age, ethnicity, glycemic status and BMI.

GLM models, with or without the inclusion of a SNP/GRS x sedentary behaviour interaction term were used to analyze our two primary objectives: (1) the association between sedentary behaviour and BMI; (2) the interaction between sedentary behaviour and 23 obesity predisposing gene variants (analyzed separately and as a GRS) and BMI. These analyses were adjusted for covariates including, sex, age, ethnicity, glycemic status, physical activity and a physical activity x SNP/GRS interaction term. A separate Bonferroni correction was applied to objective one (P <0.017 (0.05/3) and objective two (P <6.3 x 10^{-3} (0.05/8) to adjust for multiple testing. The power of this study was calculated using QUANTO (version 1.2.4; University of Southern California, Los Angeles, CA).

RESULTS

Characteristics of the studied cohort

The characteristics of the EpiDREAM cohort are shown in Table 1. The mean BMI was 30.2 (standard deviation [SD]=6.23) and the ethnic distribution of the cohort was 52.4% European, 20.3% Latino, 16.4% South Asian, 6.6% African, 3.0% Native American, 1.3% East Asian. The average age was 52.7 (SD=11.38) and the mean sedentary time was 2.6 hours/day (SD=1.7). Our analysis focused on 16 063 participants at baseline with complete genotype and phenotype data. The power calculations for this study are shown in Supplementary Figures 2-3.

Effect of Sedentary Behaviour on BMI

Increased sedentary time was significantly associated with increased BMI (Table 2). These associations were observed after adjusting for sex, age, ethnicity, glycemic status and physical activity (BPAA). Although sedentary behaviour and physical activity were independently associated with BMI, physical activity measured as a BPAA or MET score did not interact with sedentary time to moderate BMI (Table 2). The interaction analyses were adjusted for sex, age, ethnicity and glycemic status.

Effect of SNPs/GRS on Sedentary behaviour

We did not observe any significant or even nominally significant associations between the 23 SNPs studied and sedentary behaviour (Table 3). The association between the GRS and sedentary behaviour was also not significant. These analyses were adjusted for sex, age, ethnicity, glycemic status and BMI. Since these associations did not reach statistical significance, we are confident that the two interacting variables (obesity predisposing SNPs/GRS and sedentary time) are independent.

Effect of SNPs/GRS on BMI

Of the 23 SNPs analyzed, seven were associated with increased BMI: *TNNI3K* rs1514176, *CDKAL1* rs2206734, *FTO* rs3751812, *GIPR* rs11671664, *TAL1* rs2984618, *NT5C2* rs3824755, *TCF7L2* rs7903146. The GRS was also significantly associated with increased BMI (Table 4). These analyses were adjusted for sex, age, ethnicity and glycemic status.

Interaction Analyses

To reduce the probability of detecting false-positive interaction, only the subset of SNPs/GRS displaying a significant association with BMI were included in the interaction analyses. We did not observe any interaction effects between sedentary time and any of individual SNPs or 23 SNP GRS (Table 5). These analyses were adjusted for sex, age, ethnicity, glycemic status, physical activity and the physical activity x SNP/GRS interaction.

DISCUSSION

In the present study, we observed a significant positive association between TV viewing and BMI, and these associations were independent of physical activity level (measured categorically and quantitatively). Neither physical activity measure interacted with TV viewing to moderate BMI. Our gene-environment interaction analyses revealed that TV watching did not interact with the 23 individual SNPs/GRS to modulate BMI in a multi-ethnic population. Given that other studies of this interaction have provided contrasting results[138], further analysis is needed to clarify this association.

The association between sedentary behaviour and increased BMI (independent of physical activity) is consistent with previous evidence that excessive sitting time is an established risk factor for obesity, as well as type 2 diabetes, cardiovascular disease, cancer (colon, ovarian, endometrial) and premature mortality[304,312,330]. In the EpiDREAM cohort, each additional hour of TV watching was associated with a 0.42 increase in BMI, which is similar to previous reports[322]. Our analysis adds to this literature by confirming this relationship in a multi-ethnic cohort from 21 different countries. Lastly, we evaluated

the interaction between TV viewing and physical activity (measured as the BPAA and a MET score) to determine if the impact of sedentary time on BMI varied based on physical activity level. These analyses were performed to test the "Active Couch Potato" phenomenon, which describes individuals who achieve the public health guidelines for health-enhancing physical activity[331] and still experience the damaging dose-response effects of sedentary time[305]. Neither interaction tested (BPAA x sedentary time nor the MET score x sedentary time) was significant in our analyses and the subgroup analysis indicates that the impact of TV viewing on BMI did not vary by more than more than seven percent across the three physical activity subgroups (Table 2). These results support the "Active Couch Potato" and further emphasize that achieving physical activity guidelines is not sufficient to avoid the harm of sedentary behaviours such as watching TV. This phenomenon highlights the need to integrate recommendations for sedentary behaviours into future physical activity guidelines.

Policy reform in Australia has already incorporated similar evidence in several novel population health initiatives: 1) the Australian National Preventative Health Task Force Report includes specific recommendations to decrease extended periods of sitting in the workplace in order to address obesity, type 2 diabetes and cardiovascular disease; 2) reduced sitting time was introduced as a tenet of a state-wide mass media campaign targeted at obesity prevention; 3) Health Promotion Queensland initiated an evidence-based evaluation of the health impacts of prolonged sitting, and an assessment of interventions to decrease workplace sitting[300]. Another emerging approach to address the harm associated with sedentary time involves avoiding long uninterrupted periods of sitting. Two randomized trials analyzing the impact of uninterrupted versus interrupted

sitting found that two-minute bouts of moderate or even light-intensity walking every 20 minutes improved postprandial glucose and insulin levels, as well as resting systolic and diastolic blood pressure[332,333]. A separate randomized study reported that alternating (every 30 minutes) between sitting and standing using a height-adjustable workstation resulted in less fatigue and reduced musculoskeletal discomfort compared to traditional seated work posture[334]. These improvements were observed in only five days of exposure to the experimental condition among overweight and obese office workers, and productivity did not differ between the two groups[334]. Overall, our results and those from existing studies underscore the harm associated with excessive sedentary time (particularly prolonged sitting) and highlight the need to incorporate sedentary time recommendations into future health and physical activity policy.

Our interaction results indicated that sedentary behaviour did not moderate the impact of the individual SNPs or the GRS on either obesity measure. This is supported by a recent study in the UK Biobank which did not report a significant interaction between TV watching and *FTO* rs3751812[335]. In contrast, a previous study reported that prolonged TV watching accentuated the impact of a 32 SNP GRS on BMI, while a third analysis found that screen time interacted with rs2112347 near *FLJ35779* in European Americans and rs10938357 near *GNPDA2* in African Americans, although these results were nominally significant[138,154]. There are several reasons to explain the negative results in our analysis. First, our analysis included a GRS with only 23 SNPs, which may have decreased our power to detect interaction effects since some genes that may interact individually may have been excluded. Second, our study included individuals ages 18-85 years, while the previous studies by Qi *et al* focused on females 30-55 and males 40-75,

and Graff et al studied adolescents (mean age 16.2 years, SD=1.6). Analyzing samples of different ages may lead to different results due to influences such as cohort effects (e.g. increasing effects of the obesogenic environment). Third, our sample included participants from six ethnic groups recruited from 21 different countries. The ethnic and lifestyle heterogeneity of this sample may have introduced important genetic and environmental differences that influenced our results. Future studies of this interaction should include large samples with accurately measured environmental data to clarify the validity of this interaction.

Although the multi-ethnic context of this study introduces lifestyle and genetic heterogeneity, the large multi-ethnic sample is also a strength in terms of generalizability of the findings. Other strengths include the interaction analyses with individual SNPs, which provide a comprehensive examination of these interaction effects to complement the analysis of the GRS. Limitations of this study include the assessment of environmental exposures such as TV viewing and physical activity, which were measured by self-report. Perhaps most importantly, our measurement of sedentary behaviour only included time watching television and did not include other common forms of sedentary behaviour such as sitting time in the workplace or commuting. Failure to capture this information reduces the accuracy of this measurement and likely decreased the statistical power of the analysis. The 23 SNPs analyzed are only a subset of the obesity predisposing SNPs identified to date. Also, many of these SNPs were identified in European populations and may not be ideal proxies for the causal SNPs in other ethnic groups. Lastly, the EpiDREAM participants were identified as at risk for dysglycemia and are not representative of the general population.

Our results show that prolonged sedentary time is a significant risk factor for obesity,of physical activity level. Moreover, increased physical activity did not diminish the impact of sedentary time on obesity, which supports the "Active Couch Potato" phenomenon and calls for the integration of sedentary behaviour guidelines into future health promotion policy. We did not find any evidence of interaction between sedentary behaviour and any of the 23 obesity predisposing SNPs or the GRS for BMI. Future studies of this interaction should include a more comprehensive assessment of sedentary behaviours using objective instruments such as accelerometers.

Table 1. Baseline characteristics by sedentary level in the EpiDREAM study.

		Sedentary Time (hours/day)				
	Category	**< 2**	**2-3**	**> 3**	**All**	**P-value**
Total at baseline N(%)		4419 (27.5%)	8128 (50.6%)	3516 (21.9%)	16063 (100%)	
Sex N(%)	Male	1803 (28.8%)	3157 (50.4%)	1301 (20.8%)	6261 (39.0%)	2.5×10^{-3}
	Female	2616 (26.7%)	4971 (50.7%)	2215 (22.6%)	9802 (61.0%)	
[a]Age (years)		49.84 ± 10.46	52.90 ± 11.25	54.61 ± 12.17	52.66 ± 11.38	1.4×10^{-82}
Glycemic status N (%)	Normal IFG/IGT Diabetes	2160 (30.9%) 1704 (24.9%) 555 (24.8%)	3466 (49.7%) 3564 (52.1%) 1098 (49.1%)	1354 (19.4%) 1578 (23.0%) 584 (26.1%)	6980 (43.5%) 6846 (42.6%) 2237 (13.9%)	8.7×10^{-21}
Activity	Low Moderate High	1169 (26.5%) 2685 (60.7%) 564 (12.8%)	2046 (25.2%) 5075 (62.4%) 1004 (12.4%)	1302 (37.0%) 1884 (53.6%) 329 (9.4%)	4517 (28.1%) 9644 (60.1%) 1897 (11.8%)	7.4×10^{-39}
[a]METS		5.18 ± 6.55 (2963)	5.46 ± 6.90 (5644)	5.31 ± 7.00 (2287)	5.35 ± 6.83 (10894)	0.14
[a]BMI (kg/m^2)		28.98 ± 5.77 (4417)	30.27 ± 6.05 (8126)	31.58 ± 6.84 (3514)	30.20 ± 6.23 (16057)	5.8×10^{-76}
Ethnic groups N(%)	South Asian	1127 (42.8%)	1131 (43.0%)	375 (14.2%)	2633 (16.4%)	1.4×10^{-12}
	East Asian	77 (35.8%)	92 (42.8%)	46 (21.4%)	215 (1.3%)	
	European	2069 (24.6%)	4439 (52.8%)	1904 (22.6%)	8412 (52.4%)	
	African	228 (21.7%)	434 (41.2%)	390 (37.1%)	1052 (6.6%)	
	Latino American	809 (24.7%)	1801 (55.1%)	660 (20.2%)	3270 (20.3%)	
	Native-North American	109 (22.7%)	231 (48.0%)	141 (29.3%)	481 (3.0%)	

Notes: BMI: body mass index, SD = standard deviation; N = sample size
[a]Data are presented as mean ± S.D. (N).

Table 2. Effect of sedentary behaviour on BMI (adjusted for sex, age, ethnicity and glycemic status).

	Outcome	β (95% CI)	P-value
TV (hours/day)	BMI	0.42 (0.37 to 0.48)	**1.3×10^{-49}**
Interaction between sedentary time and physical activity on BMI (adjusted for sex, age, ethnicity and glycemic status)		**β (95% CI)**	**P-value**
[a]BPAA		-1.49 (-1.75 to 1.24)	**6.5×10^{-30}**
[a]Sedentary behaviour		0.29 (0.15 to 0.44)	**8.9×10^{-5}**
[a]BPAA x Sedentary behaviour		0.01 (-0.07 to 0.09)	0.75
[b]MET score		-0.04 (-0.07 to -0.01)	**3.7×10^{-3}**
[b]Sedentary behaviour		0.45 (0.37 to 0.53)	**8.7×10^{-28}**
[b]MET score x Sedentary behaviour		-2.9×10^{-3} (-0.01 to 6.5×10^{-3})	0.55
Subgroup analysis	<500 MET mins/wk	0.44 (0.37 to 0.51)	1.4×10^{-31}
	500-1000 MET mins/wk	0.41 (0.27 to 0.55)	8.5×10^{-9}
	>1000 MET mins/wk	0.44 (0.17 to 0.71)	1.5×10^{-3}

Notes: [a] Signifies terms included in the first model
[b] Signifies terms included in the second model
BPAA=basic physical activity score

Table 3. Effect of SNPs/GRS on sedentary behaviour (adjusted for sex, age, ethnicity, glycemic status and BMI).

SNP	gene	**β (95% CI)**	**P-value**
rs1514176	*TNNI3K*	0.03 (-4.0×10^{-3} to 0.07)	0.08
rs6235	*PCSK1*	0.01 (-0.03 to 0.05)	0.65
rs6232	*PCSK*	0.07 (-0.02 to 0.16)	0.14
rs2206734	*CDKAL1*	-0.01 (-0.05 to 0.04)	0.82
rs2272903	*TFAP2B*	-0.03 (-0.08 to 0.02)	0.20
rs1211166	*NTRK2*	-0.01 (-0.05 to 0.03)	0.69
rs6265	*BDNF*	0.03 (-0.02 to 0.08)	0.22
rs1401635	*BDNF*	-0.01 (-0.05 to 0.03)	0.69
rs997295	*MAP2K5*	-0.01 (-0.04 to 0.03)	0.69
rs7203521	*FTO*	-0.01 (-0.04 to 0.03)	0.79
rs3751812	*FTO*	-2.0×10^{-4} (-0.04 to 0.04)	0.98
rs1805081	*NPC1*	0.01 (-0.03 to 0.05)	0.64
rs2075650	*APOE*	2×10^{-3} (-0.05 to 0.05)	0.95
rs11671664	*GIPR*	-1.0×10^{-3} (-0.06 to 0.06)	0.96
rs2984618	*TAL1*	4.0×10^{-3} (-0.03 to 0.04)	0.85
rs1011527	*LEPR*	0.04 (-0.02 to 0.10)	0.15
rs7605927	*POMC*	0.02 (-0.02 to 0.06)	0.30
rs611203	*USP37*	-0.01 (-0.05 to 0.02)	0.44
rs2535633	*ITIH4*	0.02 (-0.02 to 0.05)	0.41
rs3824755	*NT5C2*	0.03 (-0.02 to 0.08)	0.19
rs7903146	*TCF7L2*	-0.02 (-0.06 to 0.02)	0.25
rs671	*ALDH2*	0.09 (-0.27 to 0.54)	0.62
rs749767	*KAT8*	-0.01 (-0.04 to 0.03)	0.74
	GRS	-1.6×10^{-3} (-0.01 to 7.3×10^{-3})	0.73

Table 4. Effect of SNPs/GRS on BMI (adjusted for sex, age, ethnicity and glycemic status).

		BMI	
SNP	gene	**β (95% CI)**	**P-value**
rs1514176	*TNNI3K*	0.19 (0.06 to 0.32)	**$3.3x10^{-3}$**
rs6235	*PCSK1*	-0.10 (-0.16 to 0.14)	0.90
rs6232	*PCSK*	0.08 (-0.25 to 0.41)	0.65
rs2206734	*CDKAL1*	0.31 (0.16 to 0.47)	**$9.2x10^{-5}$**
rs2272903	*TFAP2B*	0.17 (-0.01 to 0.35)	0.06
rs1211166	*NTRK2*	0.09 (-0.07 to 0.24)	0.27
rs6265	*BDNF*	0.10 (-0.07 to 0.28)	0.25
rs1401635	*BDNF*	0.10 (-0.05 to 0.24)	0.21
rs997295	*MAP2K5*	0.08 (-0.05 to 0.20)	0.24
rs7203521	*FTO*	-0.08 (-0.21 to 0.05)	0.25
rs3751812	*FTO*	0.51 (0.37 to 0.64)	**$5.7x10^{-14}$**
rs1805081	*NPC1*	0.02 (-0.12 to 0.16)	0.76
rs2075650	*APOE*	0.03 (-0.17 to 0.21)	0.80
rs11671664	*GIPR*	0.31 (0.10 to 0.51)	**$3.2x10^{-3}$**
rs2984618	*TAL1*	0.29 (0.16 to 0.42)	**$1.2x10^{-5}$**
rs1011527	*LEPR*	-0.14 (-0.35 to 0.06)	0.17
rs7605927	*POMC*	0.01 (-0.13 to 0.14)	0.92
rs611203	*USP37*	0.01 (-0.12 to 0.14)	0.84
rs2535633	*ITIH4*	$3.9x10^{-3}$ (-0.13 to 0.13)	0.95
rs3824755	*NT5C2*	0.19 (0.12 to 0.37)	**0.04**
rs7903146	*TCF7L2*	0.44 (0.31 to 0.58)	**$3.3x10^{-10}$**
rs671	*ALDH2*	-0.03 (-1.34 to 1.28)	0.96
rs749767	*KAT8*	0.06 (-0.07 to 0.20)	0.36
	GRS	0.11 (0.08 to 0.14)	**$5.2x10^{-14}$**

Table 5. Interaction analyses between sedentary behaviour and obesity predisposing SNPs/GRS (adjusted for sex, age, ethnicity, relatedness, glycemic status, physical activity and physical activity x SNP/GRS interaction).

Interaction Tests						
Outcome: BMI						
Interaction terms	**SNP Main Effect**			**SNP Interaction**		
	β	**95% CI**	**P-value**	**β**	**95% CI**	**P-value**
TNNI3K rs1514176 x Sedentary behaviour	0.30	0.08 – 0.53	0.01	-0.05	-0.13- 0.02	0.18
CDKAL1 rs2206734 x Sedentary behaviour	0.13	-0.15 – 0.40	0.06	0.07	-0.02 – 0.16	0.13
FTO rs3751812 x Sedentary behaviour	1.27	0.81 – 1.73	$8.1x10^{-8}$	-0.05	-0.12 – 0.03	0.21
GIPR rs11671664 x Sedentary behaviour	0.37	0.02 – 0.73	0.04	-0.02	-0.14 – 0.10	0.76
TAL1 rs2984618 x Sedentary behaviour	0.33	0.11 – 0.56	$3.3x10^{-3}$	-0.02	-0.10 – 0.05	0.53
NT5C2 rs3824755 x Sedentary behaviour	0.41	0.10 – 0.73	0.01	-0.10	$-0.20 – 3.7x10^{-3}$	0.06
TCF7L2 rs7903146 x Sedentary behaviour	0.32	0.08 – 0.57	0.01	0.05	-0.03 – 0.13	0.26
GS x Sedentary behaviour	0.12	0.06 – 0.17	$2.2x10^{-5}$	$-4.2x10^{-5}$	-0.02 – 0.02	0.96

Supplementary Table 1. Genotype distributions, call rate and Hardy-Weinberg Equilibrium values for the 23 SNPs analyzed.

Ethnicity	Genotype Counts			Risk Allele Frequency	Genotype Call Rate (%)	HWE P-Value
rs1514176 in *TNNI3K*	**AA**	**GA**	**GG**	**G**		
European	565	1357	834	0.549	100 %	0.762
South Asian	35	70	120	0.689	100 %	3.90×10^{-5}
East Asian	3173	4555	1652	0.419	100 %	0.804
African	137	551	559	0.669	100 %	0.944
Latinos	805	1564	918	0.517	100 %	0.010
Native American	97	245	157	0.560	100 %	0.935
Total	4812	8342	4240	0.484	100 %	-
rs6235 in *PCSK1*	**CC**	**CG**	**GG**	**C**		
European	262	1096	1397	0.294	99.96 %	0.029
South Asian	20	93	112	0.296	100 %	0.912
East Asian	684	3612	5083	0.265	99.99 %	0.225
African	27	329	891	0.154	100 %	0.600
Latinos	157	1084	2046	0.213	100 %	0.384
Native American	27	188	284	0.242	100 %	0.568
Total	1177	6402	9813	0.252	99.99 %	-
rs6232 in *PCSK1*	**AA**	**GA**	**GG**	**G**		
European	2431	309	16	0.062	100 %	0.074
South Asian	223	2	0	0.004	100 %	0.947
East Asian	8491	866	23	0.049	100 %	0.853
African	1229	18	0	0.007	100 %	0.797
Latinos	3099	185	3	0.029	100 %	0.889
Native American	470	29	0	0.029	100 %	0.504
Total	15943	1409	42	0.043	100 %	-
rs2206734 in *CDKAL1*	**CC**	**TC**	**TT**	**C**		
European	1623	974	159	0.766	100 %	0.421
South Asian	102	94	29	0.662	100 %	0.321
East Asian	6015	3018	346	0.802	99.99 %	0.174
African	715	460	72	0.758	100 %	0.861
Latinos	2104	1047	136	0.799	100 %	0.689
Native American	299	182	18	0.782	100 %	0.128
Total	10858	5775	760	0.790	99.99 %	-
rs2272903 in *TFAP2B*	**AA**	**AG**	**GG**	**G**		
European	132	945	1679	0.781	100 %	0.948
South Asian	8	77	140	0.793	100 %	0.513
East Asian	118	1778	7484	0.893	100 %	0.286

Ethnicity	Genotype Counts			Risk Allele Frequency	Genotype Call Rate (%)	HWE P-Value
African	120	487	640	0.709	100 %	0.054
Latinos	72	783	2432	0.859	100 %	0.338
Native American	4	99	396	0.893	100 %	0.417
Total	454	4169	12771	0.854	100 %	-
rs1211166 in *NTRK2*	**AA**	**GA**	**GG**	**A**		
European	1436	1105	215	0.722	100 %	0.905
South Asian	151	63	11	0.811	100 %	0.196
East Asian	6115	2898	365	0.806	99.98 %	0.351
African	567	540	139	0.671	99.92 %	0.542
Latinos	2114	1019	154	0.798	100 %	0.030
Native American	378	111	10	0.869	100 %	0.582
Total	10761	5736	894	0.784	99.98 %	-
rs6265 in *BDNF*	**AA**	**AG**	**GG**	**G**		
European	146	952	1658	0.774	100 %	0.540
South Asian	46	117	62	0.536	100 %	0.497
East Asian	345	2807	6228	0.814	100 %	0.194
African	4	74	1169	0.967	100 %	0.018
Latinos	94	859	2334	0.841	100 %	0.166
Native American	14	128	357	0.844	100 %	0.539
Total	649	4937	11808	0.821	100 %	-
rs1401635 in *BDNF*	**CC**	**CG**	**GG**	**C**		
European	424	1255	1076	0.382	99.96 %	0.067
South Asian	4	28	193	0.080	100 %	0.020
East Asian	826	3801	4752	0.291	99.99 %	0.094
African	77	478	692	0.253	100 %	0.645
Latinos	156	1119	2012	0.218	100 %	0.979
Native American	28	168	303	0.224	100 %	0.462
Total	1515	6849	9028	0.284	99.99 %	-
rs997295 in *MAP2K5*	**GG**	**GT**	**TT**	**T**		
European	835	1336	585	0.455	100 %	0.239
South Asian	151	63	11	0.189	100 %	0.196
East Asian	1597	4495	3288	0.590	100 %	0.363
African	261	616	370	0.544	100 %	0.876
Latinos	994	1584	708	0.456	99.97 %	0.102
Native American	168	222	109	0.441	100 %	0.029
Total	4006	8316	5071	0.531	99.99 %	-
rs7203521 in *FTO*	**AA**	**GA**	**GG**	**A**		
European	521	1325	909	0.429	99.96%	0.328
South Asian	15	86	124	0.258	100 %	0.986
East Asian	3538	4370	1472	0.610	100 %	0.045
African	503	569	175	0.632	100 %	0.489

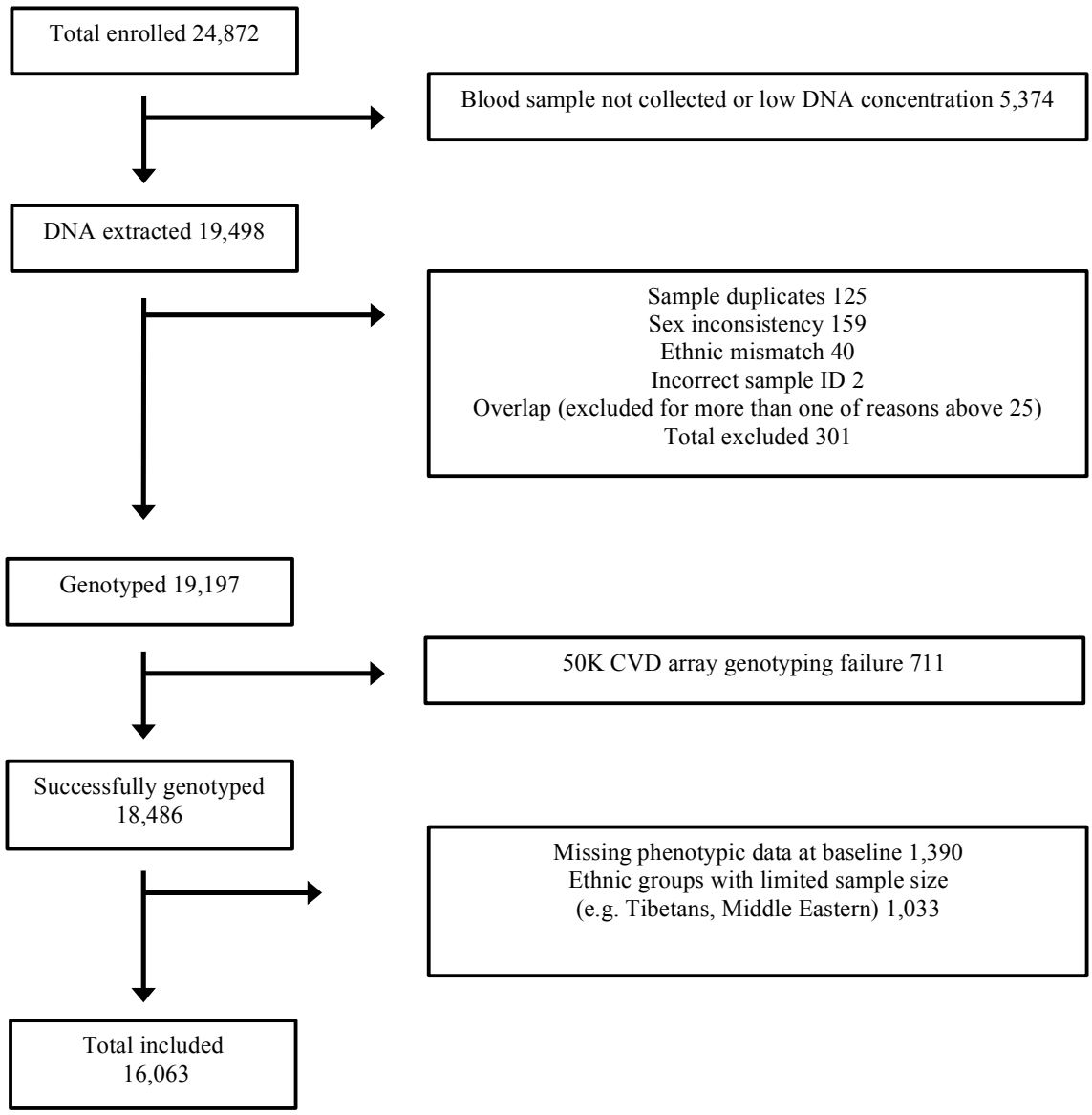
Ethnicity	Genotype Counts			Risk Allele Frequency	Genotype Call Rate (%)	HWE P-Value
Latinos	769	1548	969	0.469	99.97 %	0.002
Native American	93	202	203	0.389	99.80 %	0.001
Total	5439	8100	3852	0.546	99.98 %	-
rs3751812 in *FTO*	**GG**	**GT**	**TT**	**T**		
European	3236	4464	1680	0.417	100	0.039
South Asian	1253	1180	323	0.332	100	0.088
East Asian	154	66	5	0.197	100	0.501
African	991	244	12	0.108	100	0.479
Latinos	1575	1367	345	0.313	100	0.060
Native American	307	160	32	0.224	100	0.078
Total	7516	7481	2397	0.353	100	-
rs1805081 in *NPC1*	**AA**	**GA**	**GG**	**A**		
European	1628	973	155	0.767	100 %	0.545
South Asian	133	77	15	0.762	100 %	0.402
East Asian	3552	4385	1443	0.612	100 %	0.140
African	1083	158	6	0.92	100 %	0.928
Latinos	1588	1372	327	0.692	100 %	0.226
Native American	224	231	44	0.680	100 %	0.151
Total	8208	7196	1990	0.679	100 %	-
rs2075650 in *TOMM40-APOE-APOC1*	**AA**	**GA**	**GG**	**A**		
European	2084	621	51	0.869	100 %	0.549
South Asian	173	47	5	0.873	100 %	0.402
East Asian	6951	2248	181	0.861	100 %	0.962
African	956	278	13	0.878	100 %	0.144
Latinos	2593	647	47	0.887	100 %	0.361
Native American	398	96	5	0.894	100 %	0.767
Total	13155	3937	302	0.870	100 %	-
rs11671664 in *GIPR*	**AA**	**AG**	**GG**	**G**		
European	33	521	2202	0.894	100 %	0.727
South Asian	41	89	95	0.620	100 %	0.016
East Asian	122	1767	7488	0.892	99.97 %	0.126
African	15	256	976	0.885	100 %	0.696
Latinos	29	531	2726	0.910	99.97 %	0.577
Native American	3	102	394	0.892	100 %	0.187
Total	243	3266	13881	0.892	99.98 %	-
rs2984618 in *TAL1*	**GG**	**TG**	**TT**	**T**		
European	554	1313	889	0.561	100 %	0.084
South Asian	3	19	203	0.944	100 %	0.003
East Asian	3414	4445	1521	0.399	100 %	0.245

Ethnicity	**Genotype Counts**			**Risk Allele Frequency**	**Genotype Call Rate (%)**	**HWE P-Value**
African	87	484	676	0.736	100 %	0.977
Latinos	766	1513	1008	0.537	100 %	$2.0x10^{-5}$
Native American	115	248	136	0.521	100 %	0.924
Total	4939	8022	4433	0.485	100 %	-
rs1011527 in *LEPR*	**GG**	**AG**	**AA**	**A**		
European	1362	1144	250	0.299	100 %	0.660
South Asian	184	40	1	0.093	100 %	0.449
East Asian	8038	1284	56	0.074	99.98 %	0.544
African	928	293	26	0.138	100 %	0.612
Latinos	2845	428	13	0.069	99.97 %	0.467
Native American	450	47	2	0.051	100 %	0.520
Total	13807	3236	348	0.113	99.98 %	-
rs7605927 in *POMC*	**GG**	**GC**	**CC**	**G**		
European	586	1364	794	0.536	99.56 %	0.996
South Asian	73	113	39	0.424	100 %	0.675
East Asian	577	3386	5413	0.758	99.96 %	0.123
African	263	578	406	0.557	100 %	0.032
Latinos	508	1491	1282	0.617	99.82 %	0.031
Native American	77	215	207	0.630	100 %	0.092
Total	2084	7147	8141	0.673	99.87 %	-
rs611203 in *USP37*	**AA**	**GA**	**GG**	**G**		
European	1351	1101	303	0.310	99.96 %	0.001
South Asian	149	66	10	0.191	100 %	0.442
East Asian	3264	4517	1599	0.411	100 %	0.591
African	452	616	179	0.391	100 %	0.183
Latinos	1094	1630	563	0.420	100 %	0.292
Native American	218	219	62	0.344	100 %	0.544
Total	6528	8149	2716	0.390	99.99 %	-
rs2535633 in *ITIH4*	**GG**	**GC**	**CC**	**G**		
European	456	1287	1013	0.399	100 %	0.168
South Asian	27	111	87	0.367	100 %	0.351
East Asian	1565	4503	3309	0.407	99.97 %	0.618
African	584	538	125	0.684	100 %	0.946
Latinos	774	1618	894	0.482	99.97 %	0.426
Native American	123	245	131	0.492	100 %	0.691
Total	3529	8302	5559	0.442	99.98 %	-
rs3824755 in *NT5C2*	**GG**	**CG**	**CC**	**C**		
European	178	961	1617	0.761	100 %	0.030
South Asian	15	100	110	0.711	100 %	0.220
East Asian	103	1671	7606	0.900	100 %	0.297

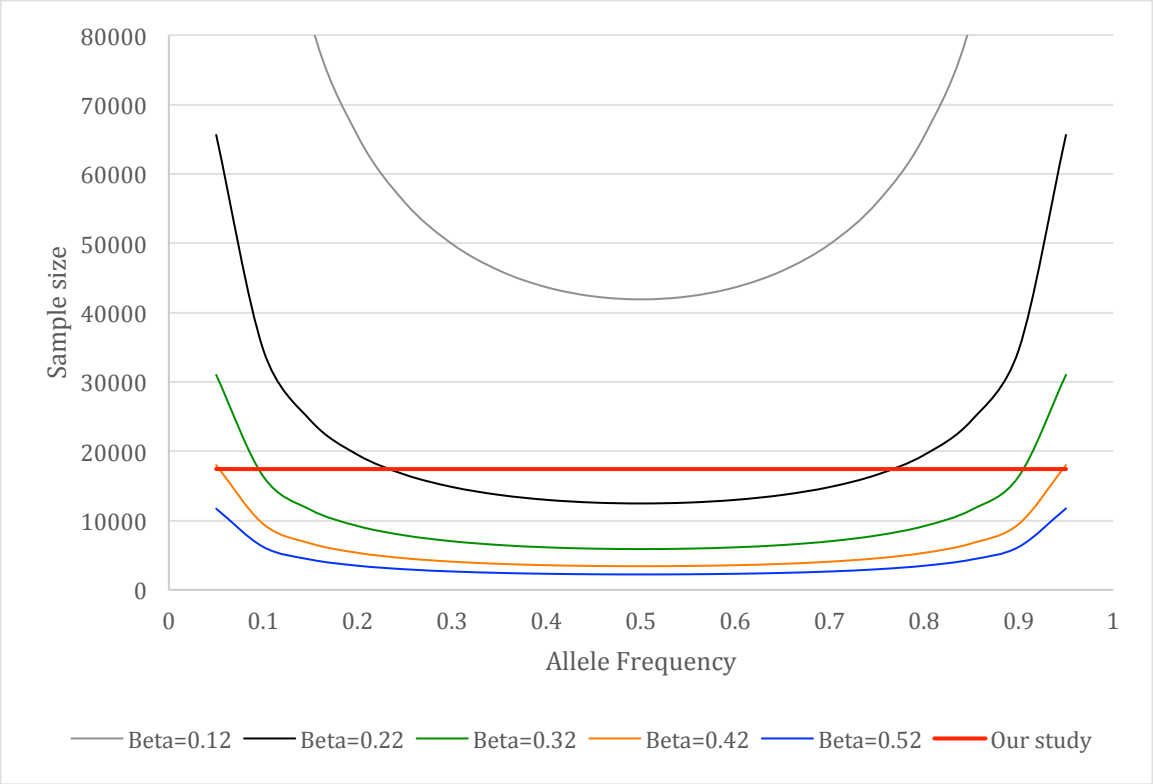
Ethnicity	Genotype Counts			Risk Allele Frequency	Genotype Call Rate (%)	HWE P-Value
African	45	395	807	0.806	100 %	0.696
Latinos	93	903	2291	0.834	100 %	0.723
Native American	8	122	369	0.862	100 %	0.562
Total	442	4152	12800	0.855	100 %	-
rs7903146 in *TCF7L2*	**TT**	**TC**	**CC**	**C**		
European	272	1170	1314	0.689	100 %	0.624
South Asian	2	22	201	0.942	100 %	0.126
East Asian	956	3903	4521	0.690	100 %	0.01
African	112	520	615	0.702	100 %	0.890
Latinos	280	1304	1703	0.716	100 %	0.177
Native American	10	145	344	0.835	100 %	0.238
Total	1632	7064	8698	0.703	100 %	-
rs749767 in *KAT8*	**GG**	**GA**	**AA**	**A**		
European	52	620	2082	0.868	99.93 %	0.461
South Asian	154	60	11	0.182	100 %	0.114
East Asian	1437	4477	3466	0.608	100 %	0.889
African	111	496	640	0.712	100 %	0.291
Latinos	715	1559	1012	0.545	99.97 %	0.013
Native American	109	232	158	0.549	100 %	0.172
Total	2578	7444	7369	0.638	99.98 %	-
rs12617233 in *FANCL*	**TT**	**CT**	**CC**	**T**		
European	352	1301	1103	0.636	100 %	0.297
South Asian	67	119	39	0.438	100 %	0.264
East Asian	1537	4480	3363	0.597	100%	0.489
African	152	579	516	0.646	100 %	0.593
Latinos	638	1583	1066	0.565	100 %	0.247
Native American	76	247	176	0.600	100 %	0.483
Total	2822	8309	6263	0.599	100 %	-

Supplementary Table 2. List of 39 self-reported physical activities.

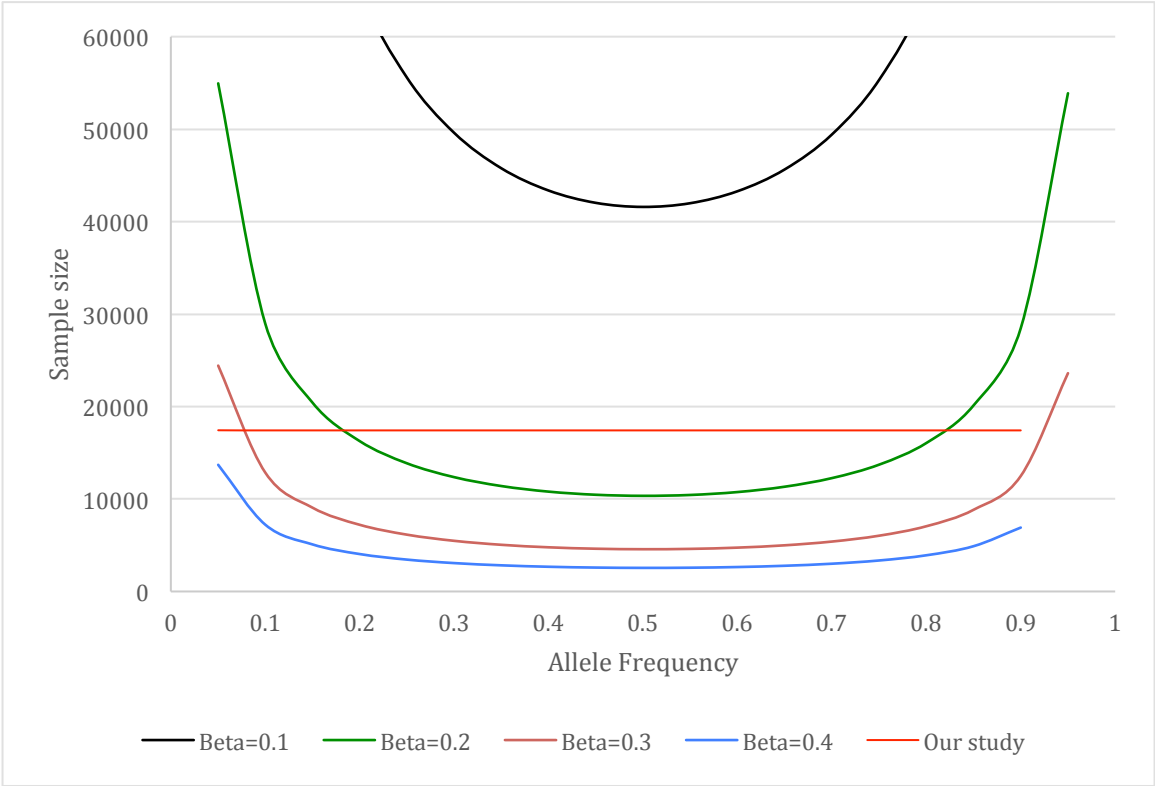
1. Aerobics/Calisthenics
2. Badminton
3. Basketball
4. Bicycling
5. Bowling
6. Pilates
7. Dance
8. Fishing
9. Soccer
10. Yard work
11. Golfing
12. Hiking
13. Hockey
14. Horseback riding
15. Jogging
16. Jump rope
17. Martial arts
18. Squash
19. Mountain climbing
20. Rugby
21. Scuba diving
22. Skating
23. Snowshoeing
24. Downhill skiing
25. Cross country skiing
26. Softball
27. Stairs
28. Weightlifting
29. Swimming
30. Ping pong
31. Tai chi
32. Tennis
33. Volleyball
34. Walking
35. Aquatics exercise
36. Water skiing
37. Wood chopping
38. Yoga
39. Rowing



Supplementary Figure 1. Flow chart of EpiDREAM study (sedentary x gene interaction).



Supplementary Figure 2. Power calculation (80%) for the main effect of obesity predisposing SNPs for a P-value=0.05.



Supplementary Figure 3. Power calculation (80%) for the interaction effect between obesity predisposing SNPs and sedentary behaviour for a P-value=0.05.

CHAPTER 4-SLEEP DURATION AND BODY MASS INDEX IN A MULTIETHNIC STUDY: EVIDENCE FROM OBSERVATIONAL AND GENETIC EPIDEMIOLOGY

ABSTRACT

Background: Obesity is a global epidemic that is caused by a range of environmental and genetic influences. We assessed the interaction between sleep duration and 23 single nucleotide polymorphisms (SNPs) on obesity.

Methods: The study included 17 377 participants from six ethnic groups with phenotypic and genotypic information available. We analysed the impact of sleep duration on the association between 23 SNPs (analyzed individually and as a genetic risk score (GRS)) and body mass index (BMI).

Results: We observed a significant U-shaped (quadratic) relationship between sleep duration and BMI ($P=6.5x10^{-4}$). *FTO* rs3751812, *CDKAL1* rs2206734, *TNNl3K* rs1514176, *GIPR* rs11671664, *TAL1* rs2984618, *NT5C2* rs3824755, *TCF7L2* rs7903146 and the GRS were associated with BMI ($4.7x10^{-14} \leq P \leq 0.01$). We did not observe any significant interaction effects between sleep duration and any of the 23 individual SNPs or GRS on BMI.

Conclusion: Our results indicate a U-shaped association between sleep duration and BMI, whereby moderate sleepers display lower BMI values than short or long sleepers. We did not detect an interaction effect between sleep duration and obesity predisposing SNPs in a multi-ethnic international sample.

INTRODUCTION

Obesity is the outcome of an imbalance between energy intake and expenditure[336]. Currently affecting more than 600 million adults worldwide according to the World Health Organization, obesity has reached epidemic proportions and shows no sign of slowing down[337]. Obesity is a risk factor for several adverse health outcomes, including type 2 diabetes, hypertension, cardiovascular disease, osteoarthritis, certain types of cancers and mental health disorders[2]. In its more severe forms, obesity is associated with a 8-13 year decrease in life expectancy[3].

Obesity and obesity-related health problems place a significant financial burden on the healthcare infrastructures of countries[338]. Part of the explanation is that obesity is difficult to treat. Traditional treatments include diet restriction and physical exercise programs, yet the effects of lifestyle interventions on long-term weight reduction are limited[339]. Weight loss medications including orlistat, lorcaserin, and phentermine-topiramate, result in 3-9% more weight loss than lifestyle modifications alone, yet carry the potential for adverse side effects[340]. Bariatric surgery is the most efficient method of treating severe obesity yet it is highly invasive and associated with post-operative complications and death in 1% of cases[341]. Additionally, only a small fraction of eligible patients can benefit from a surgical treatment due to the increasing divide between the severe obesity epidemic and the evolution of health infrastructures[341]. There is a growing consensus among the scientific and medical communities that the key to curbing the obesity epidemic lies in prevention rather than treatment[342].

The environmental causes of obesity are multiple and accumulate in a variable fashion among individuals. The two most commonly cited causes of obesity are excess

food consumption and lack of physical activity, yet an increasing amount of research has highlighted the role of sleep patterns in the development of obesity[343]. Variation in sleep duration influences the neuroendocrine control of appetite by modulating circulating levels of ghrelin and leptin which can alter food intake and obesity risk[344]. Although several studies have reported that only decreased sleep duration is associated with increased obesity risk[345,346], another substantial body of evidence, including long-term follow-up studies[347] and large-scale cross-sectional analyses[348,349], support a U-shaped relationship between sleep duration and obesity. These inconsistent findings create uncertainty regarding the nature of this relationship and this association deserves further investigation.

Although the obesity epidemic is driven by environmental and social changes, inter-individual variations in BMI involve biological roots. For instance, age and sex are strong predictors of obesity[33]. The existence of significant ethnic disparities in the prevalence of obesity suggests that specific genes and lifestyles play a role in causing the disorder[350]. Admixture studies confirm that ethnic specific genetic variations account for BMI differences[351]. Data from twin and family studies provide heritability estimates for BMI ranging between 0.24-0.90[352]. Twelve genes involved in the neuronal differentiation of the paraventricular nucleus and in the leptin-melanocortin system lead to Mendelian forms of hyperphagic obesity[353]. Currently, more than 160 common gene variants have been associated with obesity phenotypes at a genome-wide level of significance[12,45]. The recent progress in obesity gene identification has been followed by gene-environment interaction (GEI) studies. Obesity predisposing genes have been shown to interact with obesogenic environments, physical activity, TV watching, and diet[354]. A recent twin

study indicates that habitual sleep duration had a significant effect on heritability estimates among 1,088 twin pairs[355]. The heritability of BMI when sleep duration was < 7 hr (h^2 = 70%) was more than twice as large as the heritability of BMI when sleep duration was ≥ 9 hr (h^2 = 32%) in this study[355]. A recent analysis of the UK Biobank found that the effect of a genetic risk score on BMI was greater among participants who slept less than seven hours per day or more than nine hours per day compared to normal length sleepers (7-9 hours per day).

In this study, we assessed: (1) the association between sleep duration and obesity; and (2) the interaction between obesity predisposing SNPs and sleep behaviors on obesity-related traits. These objectives were analyzed using 23 obesity predisposing SNPs (analyzed individually and as a genetic risk score (GRS)) in the multi-ethnic EpiDREAM study.

MATERIALS AND METHODS

Study Participants

The data for this study were assembled through a multi-ethnic study of participants at risk for type 2 diabetes, which has been described in detail previously[326,356]. In brief, 24 872 individuals from 17 countries who were enrolled in EpiDREAM were assessed for eligibility to enter the DREAM clinical trial[326]. Study participants completed a survey that collects information including physical activity, demographic data, medical history, and sleep behaviors. A 75-gram oral glucose tolerance test (OGTT) was administered to those individuals who were identified as at risk for type 2 diabetes based on ethnicity, abdominal adiposity, and family history. Participants were screened between July 2001 and August 2003, and were between 18-85

years of age. This study focused on 17 337 subjects having both phenotypic and gene-centric 50 K single nucleotide polymorphism (SNP) array information in the EpiDREAM study (Supplementary Figure 4). Overall, the participants included in this study represented six ethnic groups (South Asian, East Asian, European, African, Latin American, Native North American). Self-reported ethnicity of the participants was validated using Eigensoft software[357] and the first 10 principal components were included in the analyses to adjust for population stratification. The EpiDREAM study has been approved by local ethics committee and in accordance with the Declaration of Helsinki, each participant provided informed consent before participating in the study.

Genotyping

DNA was extracted from buffy coats using the Gentra System (Supplementary Figure 1). Illumina CVD bead chip microarray ITMAT Broad Care (IBC) array was genotyped using the Bead Studio genotyping module, version 3.2 at the McGill University and Genome Quebec Innovation Centre[325]. We established a list of SNPs that reached genome-wide significance ($P<5x10^{-8}$) with BMI or binary obesity status in populations of European ancestry. Three separate approaches were used to select SNPs using a key word search (e.g. BMI) on i) the National Human Genome Research Institute (NHGRI) GWAS Catalog (www.genome.gov/gwastudies/), ii) the HuGE Navigator GWAS Integrator (www.hugenavigator.net/HuGENavigator/gWAHitStartPage.do), iii) the PubMed database (www.ncbi.nlm.nih.gov/pubmed). This search was performed independently by two individuals (HR and DM) and in October 2015, this approach yielded 136 independent SNPs that were found to be associated with BMI or binary obesity status. If lead SNPs were not available, proxy SNPs were selected if they

displayed an r^2 >0.90 with the lead SNP in a population of similar ancestry from the 1000 Genomes Project, and were included in the Illumina cardiovascular gene-centric array. The SNAP (SNP Annotation and Proxy Search) tool from the Broad Institute website was coupled with an independent method we developed to identify proxy SNPs. Twenty-three lead or proxy SNPs from the list of 136 were available on version 1 and 2 of the IBC 50K SNP array (Supplementary Table 1). All SNPs selected were in Hardy-Weinberg Equilibrium in the six ethnic groups ($P > 10^{-6}$) and the call rate for each of the 23 SNPs was comprised between 99.8-100% (Supplementary Table 1).

Phenotyping

A standardized protocol was used to assess anthropometric measurements. Weight (kg) and height (m) measurements were assessed by trained medical staff. Weight was measured to the nearest 0.1 kg in light clothing and standing height was measured to the nearest 0.1 cm. Body mass index (BMI) was calculated as weight in kilograms (kg) divided by height in meters (m) squared. The oral glucose tolerance test was used to categorize participants into four groups: normal glucose tolerant (NGT), impaired fasting glucose (IFG), impaired glucose tolerance (IGT), or T2D at baseline, all based on the 2003 ADA criteria[327]. The IFG and IGT categories were combined to form a three-level categorical variable referred to as glycemic status (normoglycemia, IFG/IGT, diabetic).

Measures of sleep duration were based on self-reported information from the participant questionnaire. Sleep duration was measured by average number of hours spent sleeping per day. Participants were placed into one of three sleep categories that reflect current practice in the sleep literature: sleep-low (<6), sleep-moderate (6-8), sleep-high (>8)[358].

Statistical Analysis

Single SNPs analyses were performed under the additive model, and the previously identified obesity risk alleles for each of the 23 SNPs were used as the risk allele for the analyses. The GRS was calculated by summing the alleles of the 23 obesity predisposing SNPs so that the GRS ranged from 0 to 46. An un-weighted GRS was used for these analyses[359]. We performed ethnic-specific imputations for the missing genotypic values as previously described[329].

Before conducting the main analyses, two preliminary verification tests were performed. First, the associations between the SNPs / GRS and BMI were analyzed using linear regression models. Since the associations between the SNPs and obesity measures have been shown in previous studies, a $P < 0.05$ was considered significant. These tests were adjusted for sex, age, ethnicity and glycemic status and all SNPs / GRS below the significance threshold were carried forward to the interaction analysis. Second, linear/ordinal regression were used to analyze the impact of the 23 obesity predisposing SNPs and GRS on sleep duration. These tests were performed to ensure that the two interacting terms were independent. A Bonferroni adjusted level of significance was used for this analysis $P < 2.1 \times 10^{-3}$ (0.05/24 (23 SNPs + GRS)). This analysis was performed while controlling for sex, age, ethnicity, glycemic status and BMI.

Linear regression models, with or without the inclusion of a SNP / GRS x sleep interaction term were used to analyze (1) the association between sleep parameters and BMI; (2) the interaction between sleep duration and the 23 obesity predisposing gene variants (analyzed independently and as a GRS) on BMI. These tests were adjusted for

covariates including, sex, age, ethnicity and glycemic status. A Bonferroni correction was applied to the interaction analysis $P < 6.3 \times 10^{-3}$ (0.05/8 (7 SNPs + GRS)).

Given the conflicting evidence regarding the nature of the association between sleep duration and obesity (inverse linear[346,360-363] *versus* U-shaped[347,348,364]), two linear regression models were generated with and without the inclusion of a quadratic term to compare the two associations. A partial F-test was used to determine if the addition of the quadratic term significantly improved the fit of the model. These tests were adjusted for covariates including, sex, age, ethnicity and glycemic status. Two-tailed P-values are presented in this manuscript and SPSS (version 20, New York, USA, IBM Corporation) was used to perform all statistical analyses.

RESULTS

Characteristics of the studied cohort

The EpiDREAM cohort enrolled 24 872 people from 17 countries and 6 different ethnic groups. Of the original sample, 17 337 had complete genotype and phenotype data for this analysis. The baseline characteristics of the study sample are outlined in Table 6. The majority of the sample was European (53.89%), followed by Latino (18.93%), South Asian (15.88%), African (7.15%), Native North American (2.87%), and East Asian (1.29%). Participants had a mean age of 52.65 years, the average sleep duration was 6.95 hours/day (Standard deviation [SD]= 1.54) and the mean BMI was 30.16 (SD= 6.22) kg/m^2.

Effect of sleep duration on BMI

Comparing the linear and U-shaped relationship between sleep duration and BMI revealed that the quadratic term (sleep duration variable (hours/day) squared) was

significantly associated with BMI (β= 0.04, 95% CI= 0.02 to 0.06, P= 6.5 x 10^{-4}) (Table 7). In addition, the addition of this term significantly increased the explanatory ability of the model (F (1,17336)=10.15, P=0.002).

Effect of SNPs/GS on sleep duration

The effect of the 23 obesity predisposing SNPs / GRS on the sleep duration is presented in Table 8. The 23 SNPs and the GRS were not associated with sleep parameters after appropriate correction for multiple tests ($P < 2.1 \times 10^{-3}$). Therefore, the interaction terms (sleep and SNPs) were assumed to be independent.

Effect of SNPs/GRS on BMI

Seven SNPs displayed significant associations with BMI: *TNNI3K* rs1514176 (β= 0.20, 95% confidence interval [CI]= 0.07 to 0.32, P= 2.0 x 10^{-3}), *CDKAL1* rs2206734 (β= 0.28, 95% CI= 0.13 to 0.43, P= 2.6 x 10^{-4}), *FTO* rs3751812 (β= 0.51, 95% CI= 0.38 to 0.64, P= 4.7 x 10^{-14}), *GIPR* rs11671664 (β= 0.27, 95% CI= 0.08 to 0.47, P= 6.7 x 10^{-3}), *TAL1* rs2984618 (β= 0.28, 95% CI= 0.16 to 0.41, P= 8.5 x 10^{-6}), *NT5C2* rs3824755 (β= 0.24, 95% CI= 0.05 to 0.43, P= 0.01), *TCF7L2* rs7903146 (β= 0.43, 95% CI= 0.30 to 0.57, P= 2.3 x 10^{-10}). The GRS was also significantly associated with greater BMI (β= 0.11, 95% CI= 0.09 to 0.14, P= 4.2 x 10^{-14}).

Sleep duration x SNP/GRS interaction analysis

The interaction analysis was restricted to SNPs/GRS that were significantly associated with BMI. Sleep duration did not moderate the impact of the seven individual SNPs of the GRS on BMI (Table 10). These analyses were adjusted for sex, age, ethnicity and glycemic status.

DISCUSSION

In this study, we observed a significant U-shaped association between sleep duration and BMI, whereby those who slept a moderate amount (6-8 hours per night) had lower BMI values than participants who slept less than six hours per night and those who slept more than eight hours per night. We did not identify any significant associations between the 23 obesity predisposing SNPs or the GRS with sleep duration. Our interaction analyses did not indicate that sleep duration moderated the effect of any of the 23 obesity predisposing SNPs or the GRS.

Even though the association between sleep duration and BMI has been studied extensively, there have been mixed findings regarding the nature of this relationship (inverse linear vs. U-shaped) and this is the first study to analyze this association in a multi-ethnic international sample. Our analyses showed a significant U-shaped relationship and inclusion of a U-shaped (quadratic) coefficient improved the overall fit of the model. This indicates that both short and long sleep duration is associated with increased BMI compared to moderate sleep duration. This pattern of association is consistent with the relationship between sleep duration and mortality[365,366]. Sleep reduction is known to decrease insulin sensitivity, a consequence of both increased growth hormone secretion during sleep and an amplification of sympathetic nervous system activity[367,368]. The changes in endocrine and sympathoadrenal functioning are also established risk factors for the development of obesity[367], and short sleep duration is associated with a number of pathological metabolic traits such as dyslipidemia, type 2 diabetes, and hypertension[2]. Experimental sleep restriction studies have demonstrated an increase in calorie consumption and a preference for higher glycemic index foods in

individuals with restricted sleep[369]. Sleep loss is also associated with an increase in ghrelin and a decrease in leptin, which creates synergistic changes that favor weight gain[344]. The specific mechanisms linking long sleep duration to obesity are not as well understood, although disrupted eating patterns (e.g. excessive snacking), high fat intake and decreased fruit and vegetable consumption have been associated with long sleep duration[370]. Together, this evidence and our findings herein support the U-shaped relationship between sleep duration and obesity. Future studies analyzing this association could advance our understanding by using objectively measured sleep data. Additional avenues for future research include the analysis of sleep quality measures such as snoring, since these measures have been associated with BMI and DNA methylation in existing studies [371,372].

Our interaction analysis did not reveal any significant interactions between sleep duration and the 23 obesity predisposing SNPs or GRS. However, two recent studies with larger sample sizes from the UK Biobank have found that sleep duration moderated the impact of variation in *FTO* rs1421085 and a 93 SNP GRS on BMI [335,373]. In these studies, short and long sleep duration increased the effect of *FTO* rs1421085 and the 93 SNP GRS on BMI compared to the effect among moderate sleepers. This association is consistent with the main effect of sleep duration on BMI that we observed in our study although we did not observe significant interaction effects. Our results may differ from these studies for several reasons. The EpiDREAM sample included six ethnic groups living in 17 different countries, which include both high- and low-income countries, while the UK Biobank analyzed white Europeans in the UK. These differences likely introduced significant heterogeneity into the analysis in terms of genetic differences and

unmeasured lifestyle behaviours that may have influenced the sleep x gene interactions tested (residual confounding)[373]. The sleep duration categories used in our analysis (<6, 6-8, >8 hours per day) also differ from that of the UK Biobank analyses (<7, 7-9, >9 hours per day) and this could also potentially account for differences in the effect of sleep between studies. Future studies using objectively measured sleep data will be useful to further clarify the pattern of interaction between sleep duration and genetic susceptibility to obesity.

Strengths of our study include the multi-ethnic nature of our sample, which provides greater ethnic generalizability of the results. Our study was limited in that the 23 SNPs that were analyzed represent only a subset of currently identified obesity predisposing SNPs[12]. Furthermore, most of the analyzed SNPs were identified in Europeans and may exert different effects in other ethnicities. We also acknowledge that the power of our study is modest (Supplementary Figure 5). Additionally, study participants were selected based on being at high-risk for dysglycemia, and are not representative of the general population. All sleep phenotype information was self-reported and likely contained a degree of error or bias.

In summary, we observed a U-shaped relationship between sleep duration and BMI, whereby those who slept a moderate amount displayed lower BMI values than participants with shorter and longer sleep durations. We did not detect any associations between the obesity predisposing SNPs analyzed and sleep duration, and sleep duration did not moderate the impact of the obesity predisposing SNPs or the GRS. Since other larger scale studies have provided evidence of gene-sleep interactions, future large

studies with objectively measured sleep information and more SNPs may provide more insight into this interaction.

Table 6. Baseline characteristics of the EpiDREAM study stratified by sleep duration.

	Category	Sleep-Low <6 hrs/day	Sleep-Moderate 6-8 hrs/day	Sleep-High >8 hrs/day	All	P-value
Total at baseline N(%)		5927 (32.1%)	9864 (53.4%)	1586 (8.6%)	17337 (100%)	
Gender N(%)	Male	2292 (38.7%)	1970 (43.0%)	2529 (36.8%)	6791 (39.1%)	1.51×10^{-10}
	Female	3635 (61.3%)	2609 (57.0%)	4342 (63.2%)	10586 (60.9%)	
[a]Age (years)		54.28 ± 11.14	51.80 ± 11.30	51.87 ± 12.02	52.65 ± 11.37	2.83×10^{-41}
Glycemic status N (%)	Normal	2370 (40.0%)	2057 (44.9%)	3003 (43.7%)	7430 (42.8%)	3.24×10^{-7}
	IFG/IGT	2646 (44.6%)	1861 (40.6%)	2878 (41.9%)	7385 (42.5%)	8.84×10^{-5}
	Diabetes	911 (15.4%)	661 (14.4%)	990 (14.4%)	2562 (14.7%)	0.245
[a]Sleep (hrs/day)		4.23 ± 0.97	7.13 ± 0.80	9.54 ± 0.80	6.95 ± 1.54	1.00×10^{-36}
[a]BMI at baseline (kg/m^2)		30.67 ± 6.41 (5924)	29.72 ± 6.10 (4584)	30.00 ± 6.10 (6869)	30.16 ± 6.22 (17377)	3.78×10^{-15}
Ethnic groups N(%)	South Asian	797 (13.4%)	933 (20.4%)	1029 (15.0%)	2759 (15.9%)	3.80×10^{-83}
	East Asian	81 (1.4%)	60 (1.3%)	83 (1.2%)	224 (1.3%)	
	European	2900 (48.9%)	2473 (4.0%)	3991 (58.1%)	9364 (53.9%)	
	African	658 (11.1%)	227 (5.0%)	357 (5.2%)	1242 (7.1%)	
	Latino American	1354 (22.8%)	753 (16.4%)	1183 (17.2%)	3290 (18.9%)	
	Native-North American	137 (2.3%)	133 (2.9%)	228 (3.3%)	498 (2.9%)	

IFG: impaired fasting glucose, IGT: impaired glucose tolerance, BMI: body mass index, SD = standard deviation; N = sample size

[a]Data are presented as mean ± S.D. (N).

Table 7. Linear vs. U-shaped models of the association between sleep duration and BMI.

Variable	Linear Model		U-shaped Model	
	β (95% CI)	**P- value**	**β (95% CI)**	**P- value**
Sex (female vs. male)	1.35 (1.17 – 1.53)	4.1×10^{-49}	1.34 (1.16 – 1.52)	6.7×10^{-48}
Age (per years older)	-0.07 (-0.08 – -0.06)	1.0×10^{-56}	-0.07 (-0.08 – -0.06)	4.7×10^{-57}
South Asian (vs. European)	-4.39 (-4.65 – -4.12)	1.5×10^{-229}	-4.35 (-4.61 – -4.09)	5.4×10^{-225}
East Asian (vs. European)	-4.83 (-5.60 – -4.06)	9.4×10^{-35}	-4.81 (-5.58 – -4.04)	2.1×10^{-34}
African (vs. European)	1.37 (1.02 – 1.72)	8.9×10^{-15}	1.35 (1.01 – 1.70)	1.8×10^{-14}
South American (vs. European)	0.24 (0.01 – 0.47)	0.044	0.25 (0.02 – 0.48)	0.034
Native North American (yes vs. no)	1.46 (0.93 – 1.98)	5.3×10^{-8}	1.46 (0.93 – 1.98)	5.6×10^{-8}
Glycemic status (IGT/IFG vs. normal)	1.97 (1.78 – 2.16)	5.5×10^{-88}	1.97 (1.78 – 2.16)	6.5×10^{-88}
Glycemic status (Diabetic vs. normal)	2.75 (2.48 – 3.01)	8.8×10^{-90}	2.75 (2.48 – 3.01)	9.4×10^{-90}
Sleep duration (per hour increase)	-0.18 (-0.24 – -0.12)	8.1×10^{-10}	-0.66 (-0.94 – -0.38)	4.4×10^{-6}
Sleep duration squared (per hour increase)	NA	NA	0.04 (0.02 – 0.06)	6.5×10^{-4}

Notes: NA= not applicable since not included in the model.

Table 8. Effect of SNPs/GRS on sleep duration.

(adjusted for gender, age, ethnicity, glycemic status and BMI)

		Sleep Duration		
SNP	Gene	β	95% CI	P-value
rs1514176	*TNNI3K*	-0.02	0.05 - 0.02	0.29
rs6235	*PCSK1*	0.03	-0.01 - 0.06	0.21
rs6232	*PCSK*	-0.10	-0.18 - -0.01	**0.02**
rs2206734	*CDKAL1*	0.02	-0.03 - 0.05	0.47
rs2272903	*TFAP2B*	-0.01	-0.05 - 0.04	0.71
rs1211166	*NTRK2*	0.01	-0.03 - 0.05	0.58
rs6265	*BDNF*	0.05	$1.3x10^{-3}$ - 0.09	**0.04**
rs1401635	*BDNF*	-0.01	-0.04 - 0.03	0.80
rs997295	*MAP2K5*	$1.8x10^{-3}$	-0.03 - 0.03	0.92
rs7203521	*FTO*	$-2.6x10^{-3}$	-0.04 - 0.03	0.88
rs3751812	*FTO*	-0.04	-0.08 - -0.01	**0.02**
rs1805081	*NPC1*	-0.01	-0.04 - 0.03	0.77
rs2075650	*APOE*	-0.01	-0.06 - 0.03	0.56
rs11671664	*GIPR*	-0.04	-0.09 - 0.01	0.12
rs2984618	*TAL1*	-0.01	-0.05 - 0.02	0.47
rs1011527	*LEPR*	0.05	$2.0x10^{-3}$ - 0.11	**0.04**
rs7605927	*POMC*	0.02	-0.02 - 0.05	0.32
rs611203	*USP37*	-0.01	-0.04 - 0.03	0.71
rs2535633	*ITIH4*	-0.01	-0.04 - 0.03	0.67
rs3824755	*NT5C2*	-0.02	-0.06 - 0.03	0.50
rs7903146	*TCF7L2*	0.02	-0.02 - 0.05	0.34
rs671	*ALDH2*	0.09	-0.03 - 0.05	0.62
rs749767	*KAT8*	-0.01	-0.05 - 0.02	0.49
	GRS	-2.1x10-3	-0.01 - 0.01	0.59

Notes: GRS=genetic risk score, bold text indicates $p<0.05$.

Table 9. Effect of SNPs/GRS on BMI.

(BMI analysis adjusted for sex, age, ethnicity and glycemic status.)

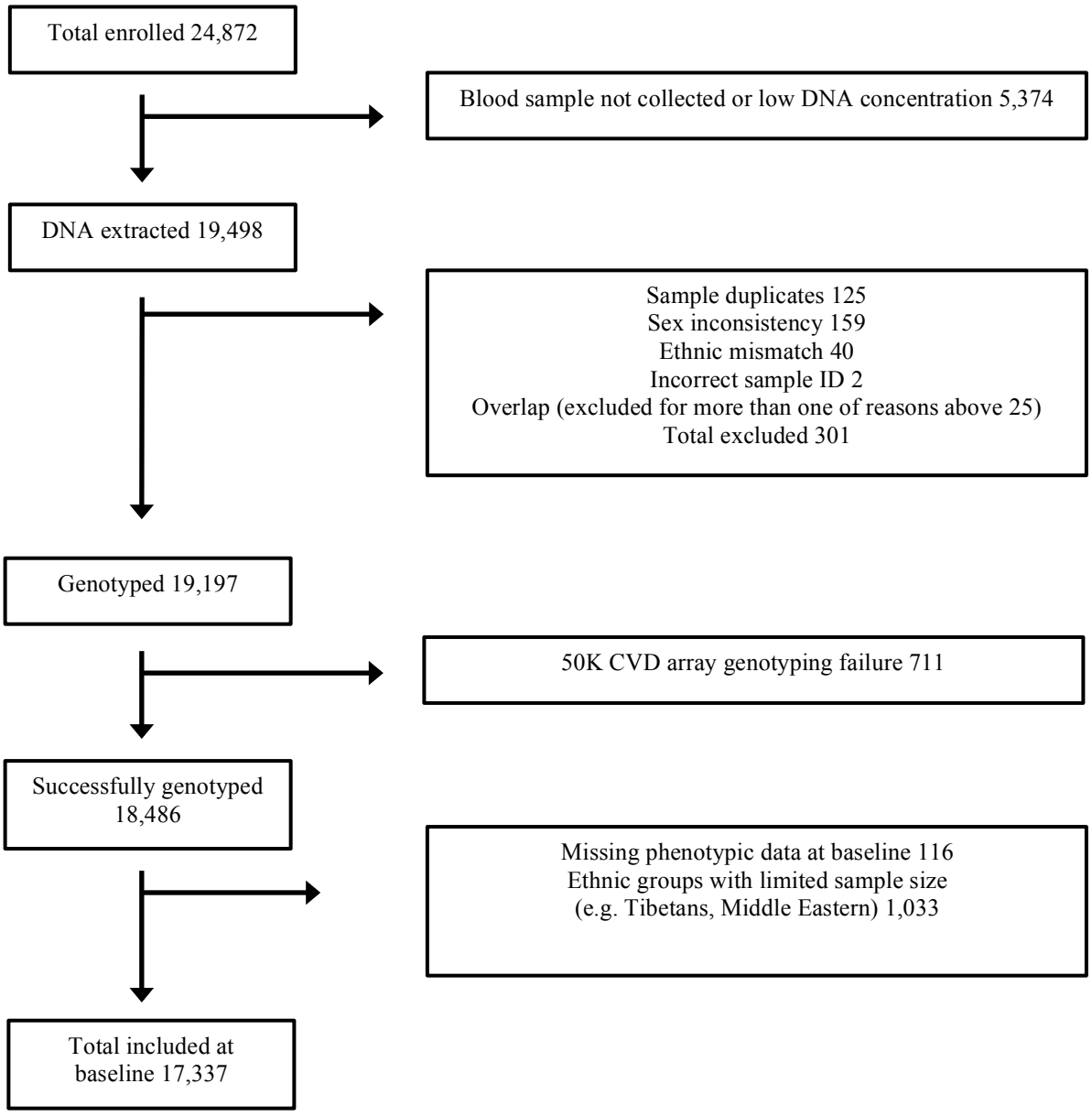
		BMI	
SNP	gene	**β (95% CI)**	**P-value**
rs1514176	*TNNI3K*	0.20 (0.07 to 0.32)	**$2.0x10^{-3}$**
rs6235	*PCSK1*	0.03 (-0.12 to 0.18)	0.68
rs6232	*PCSK*	0.05 (-0.27 to 0.37)	0.75
rs2206734	*CDKAL1*	0.28 (0.13 to 0.43)	**$2.6x10^{-4}$**
rs2272903	*TFAP2B*	0.12 (-0.06 to 0.29)	0.19
rs1211166	*NTRK2*	0.06 (-0.09 to 0.20)	0.46
rs6265	*BDNF*	0.10 (-0.07 to 0.26)	0.26
rs1401635	*BDNF*	0.07 (-0.07 to 0.22)	0.32
rs997295	*MAP2K5*	0.05 (-0.08 to 0.17)	0.45
rs7203521	*FTO*	-0.05 (-0.18 to 0.08)	0.47
rs3751812	*FTO*	0.51 (0.38 to 0.64)	**$4.7x10^{-14}$**
rs1805081	*NPC1*	0.06 (-0.07 to 0.20)	0.36
rs2075650	*APOE*	0.08 (-0.11 to 0.26)	0.41
rs11671664	*GIPR*	0.27 (0.08 to 0.47)	**$6.7x10^{-3}$**
rs2984618	*TAL1*	0.28 (0.16 to 0.41)	**$8.5x10^{-6}$**
rs1011527	*LEPR*	-0.11 (-0.31 to 0.09)	0.30
rs7605927	*POMC*	-0.02 (-0.16 to 0.11)	0.73
rs611203	*USP37*	-0.03 (-0.16 to 0.09)	0.59
rs2535633	*ITIH4*	$-2.7x10^{-3}$ (-0.13 to 0.12)	0.97
rs3824755	*NT5C2*	0.24 (0.05 to 0.43)	**0.01**
rs7903146	*TCF7L2*	0.43 (0.30 to 0.57)	**$2.3x10^{-10}$**
rs671	*ALDH2*	-0.02 (-1.29 to 1.31)	0.96
rs749767	*KAT8*	0.08 (-0.05 to 0.21)	0.25
	GRS	0.11 (0.09 to 0.14)	**$4.2x10^{-14}$**

Notes: GRS=genetic risk score, bold text indicates $p<0.05$.

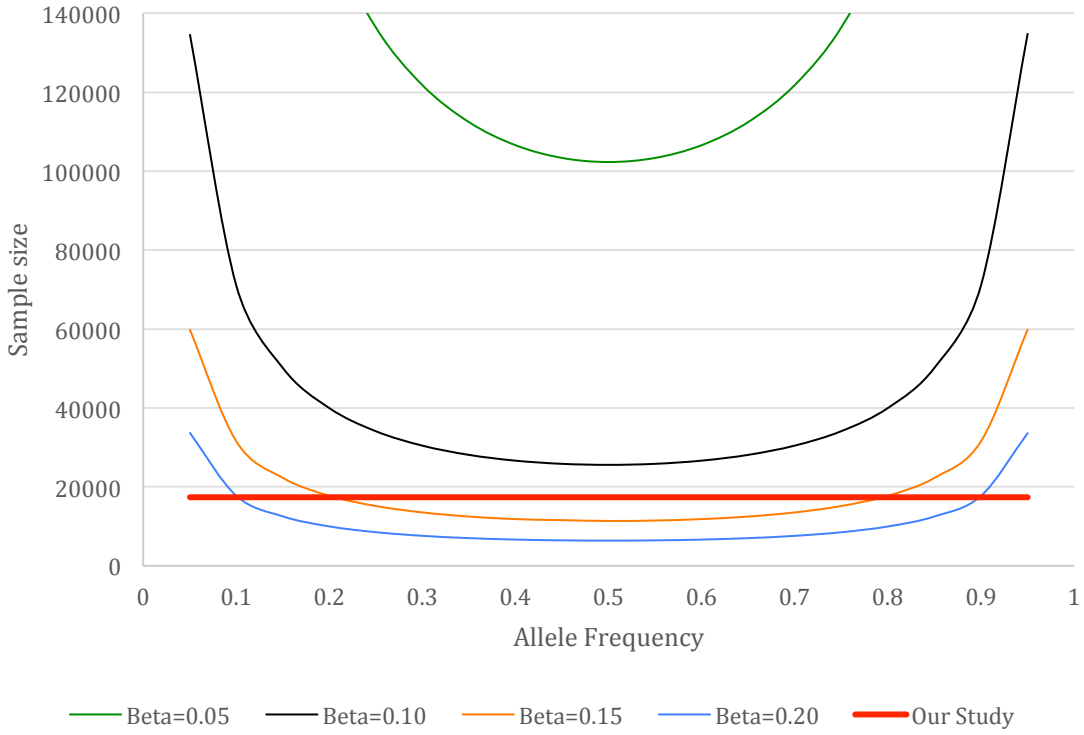
Table 10. Interaction analysis between SNPs/GRS and sleep duration on BMI.

(adjusted for gender, age, ethnicity, rosiglitazone use, and glycemic status)

Sleep Continuous (hrs squared)							
		Main effect of SNP			**Interaction**		
SNP	Gene	**β**	**95% CI**	**P-value**	**β**	**95% CI**	**P-value**
rs1514176	*TNNI3K*	-0.19	-0.51 – 0.13	0.26	$1.4x10^{-3}$	-0.01 – 0.01	0.49
rs2206734	*CDKAL1*	0.11	-0.30 – 0.52	0.60	$3.3x10^{-3}$	$-4.2x10^{-3}$ – 0.01	0.39
rs3751812	*FTO*	0.72	0.39 – 1.05	$2.3x10^{-5}$	$-4.5x10^{-3}$	-0.01 – $1.6x10^{-3}$	0.15
rs11671664	*GIPR*	0.14	-0.38 – 0.65	0.61	$2.7x10^{-3}$	-0.01 – 0.01	0.58
rs2984618	*TAL1*	0.24	-0.08 – 0.55	0.14	$8.8x10^{-4}$	-0.01 – 0.01	0.77
rs3824755	*NT5C2*	0.06	-0.44 – 0.57	0.80	$3.2x10^{-3}$	-0.01 – 0.01	0.50
rs7903146	*TCF7L2*	0.20	-0.16 – 0.55	0.28	0.01	$-1.6x10^{-3}$ - 0.01	0.14
	GRS	0.15	0.07 – 0.22	$1.3x10^{-4}$	$-6.7x10^{-4}$	$-2.3x10^{-3}$ – $7.0x10^{-4}$	0.34



Supplementary Figure 4. Flow chart of EpiDREAM study (sleep x gene interaction)



Supplementary Figure 5. Power calculation for the interaction between obesity predisposing SNPs and sleep duration for a P-value=0.05.

CHAPTER 5-ALCOHOL INTAKE AND OBESITY: EVIDENCE OF SEX AND GENE INTERACTIONS IN A MULTIETHNIC STUDY

ABSTRACT

Background: A number of studies have analyzed the association between alcohol intake and obesity with inconsistent results. This study analyzed the impact of alcohol consumption on obesity in interaction with obesity predisposing genetic variants.

Methods: We investigated these associations using 23 obesity predisposing variants (analyzed independently and as a genetic risk score (GRS)) in the multiethnic study EpiDREAM (14828 participants from six ethnic groups). Alcohol intake was measured as self-reported drinks/week (non-drinkers, 1-3 drinks/week, >3 drinks/week) and the body mass index (BMI) was used to measure obesity.

Results: Increased alcohol intake was associated with decreased BMI and the magnitude of this association was four-fold greater among females compared to males. The risk alleles of seven SNPs and the GRS were significantly associated with increased BMI (*TNNI3K* rs1514176, *CDKAL1* rs2206734, *FTO* rs3751812, *GIPR* rs11671664, *TAL1* rs2984618, *NT5C2* rs3824755, *TCF7L2* rs7903146). We did not observe any significant associations between the obesity risk variants and alcohol intake. An alcohol intake x *FTO* rs3751812 was also observed, whereby increased alcohol intake decreased the impact of *FTO* rs3751812 on BMI by 68%.

Conclusion: Our results indicate that alcohol consumption significantly reduced the impact of *FTO* rs3751812 variation on BMI in a multi-ethnic sample. Further study of this association is needed to determine the implications for public health intervention and prevention strategies.

INTRODUCTION

According to World Health Organization (http://www.who.int/), the prevalence of obesity has more than doubled since 1980 and over 600 million adults are currently defined as obese. Obesity is a risk factor for osteoarthritis, type 2 diabetes, dyslipidemia, hypertension, fatty liver, cardiovascular disease, cancer and ultimately decreased life expectancy[3,374,375]. The rise in obesity has also created substantial economic concerns. This financial burden is attributed to direct medical expenditures and indirect productivity costs, transportation costs and human capital costs[375,376]. The traditional treatments for obese individuals include diet restriction augmented by physical exercise[377]. Individuals who cannot improve their health through lifestyle changes alone often use weight loss medication in conjunction[340,378]. A subset of individuals with severe obesity resort to bariatric surgery as a method of treatment[341]. Based on the limited efficacy or accessibility of current treatments, prevention through modifiable environmental risk factors has become an important focus of research[379].

While obesity remains a multifactorial condition, it has largely been attributed to environmental risk factors, particularly 'the big two': physical inactivity and excessive energy intake[33,380]. Additional environmental factors such as smoking cessation, sleep duration/quality, psychosocial stress or exposure to pollutants / endocrine disruptors have also been identified as contributors to obesity[33,380-382]. While many of these risk factors have shown clear relationships with obesity development, research pertaining to the effects of alcohol intake has been less conclusive[6,383]. Considering its high caloric value (7 kilocalories/gram), one may expect energy consumed through alcohol to have an additive effect on other dietary sources of energy and increase obesity risk[384,385].

However, findings remain controversial and paradoxical[385,386]. Alcohol has been shown to interfere with metabolic and cognitive functions[387], and has been suggested to enhance food consumption after an alcoholic preload condition, and promote short-term over-consumption[385]. However, epidemiological studies over longer periods of time showed no consistent evidence of increased food intake on days when alcohol was consumed[385,388]. Current evidence has not established a consistent relationship between alcohol intake and obesity measures[6]. Several studies indicate a negative or no association, while some report sex-specific effects or a J-shaped relationship between alcohol intake and obesity[6]. The J-shaped association suggests that moderate alcohol intake is associated with lower risks of obesity compared to heavy drinkers or abstainers[385,386,388].

Despite the impact of obesogenic environmental risk factors, biological factors such as sex, age and ethnicity are important determinants of variation in body weight[33,72]. Admixture, twin, adoption and family studies suggest a strong genetic influence despite their varying environmental influences[352,389,390]. At this time, 20 Mendelian obesity syndromes (e.g. Bardet-Biedl) have been fully genetically elucidated[391]. Defects in twelve genes involved in the neuronal differentiation of the paraventicular nucleus and leptin/melanocortin pathway have been linked to monogenic non-syndromic obesity[353]. Candidate gene and genome-wide association studies (GWAS) have identified over 160 single-nucleotide polymorphisms (SNPs) associated with polygenic obesity[353] and emerging gene-environment interaction studies (GEI) have identified lifestyle factors (socioeconomic status, physical activity, diet) that can moderate the impact of a subset of genetic variants (e.g. *FTO*) on BMI[392]. However, only one study to date has analyzed the interaction between alcohol intake and genetic predisposition to obesity. This analysis

found that increased alcohol intake decreased the impact of *FTO* rs1421085 on BMI among 119 132 individuals from the UK Biobank. Genome wide- analysis confirmed that the majority of these participants were of northern-European ancestry, although genetic diversity was evident in a subset of this sample[323].

Existing studies have also revealed that alcohol use and abuse display a heritable component, with 36-40% of the variance in drinking measures attributable to genetic factors[393]. Similarities in personality traits, neural mechanisms, functional brain abnormalities and a likelihood of disruptive behaviour syndromes have been reported to be shared between obesity and addictive disorders, including alcohol dependence[394,395]. However, the overlap between genetic predisposition to obesity and alcohol overconsumption has been poorly investigated. Several studies explored the association of *FTO* intron 1 variation with alcohol consumption or dependence and found conflicting results[144,386,396,397]. The rs671 major allele (G) in the *ALDH2* gene has been associated with increased alcohol consumption and increased BMI in East Asian population through GWAS[398,399].

In this study, we aimed to assess: (1) the relationship between alcohol intake and BMI (2) the interaction between alcohol intake and sex on BMI; and (3) the interaction between alcohol intake and obesity predisposing SNPs on BMI. We used 23 obesity predisposing SNPs (analyzed independently and as a genetic risk score (GRS)) in the multi-ethnic study EpiDREAM.

MATERIALS AND METHODS

Study Participants

The EpiDREAM study is an international multi-ethnic cohort of participants at risk for type 2 diabetes (T2D)[400,401]. This cohort consists of 24 872 individuals from 17 countries and includes those who participated in the DREAM trial[401]. Between July 2001 and August 2003 individuals between the ages of 18 – 85 years, who were identified as being at risk for T2D based on abdominal adiposity, ethnicity and family history were screened using a 75-gram oral glucose tolerance test (OGTT). A detailed description of the cohort and methodology has been published previously[401].

This analysis included 14 828 individuals with both phenotypic and gene-centric 50 K single nucleotide polymorphism (SNP) array information (Supplementary Figure 6). The participants represented six ethnic groups (South Asian, East Asian, European, African, Latin American and Native North American). Self-reported ethnicity of all 14 828 individuals were validated using Eigensoft software[357] and the first 10 principal components were included in the analyses to adjust for population stratification. In accordance with the Declaration of Helsinki, informed consent was obtained from each participant prior to participation in the study. The EpiDREAM study has been approved by local ethics committees and all experiments were performed in accordance with relevant guidelines and regulations.

Genotyping

The Qiagen Gentra System was used to extract the DNA from the buffy coats of 19 498 participants in the EpiDREAM study (Supplementary Figure 4). Genotyping was

performed on the Illumina CVD bead chip microarray ITMAT Broad Care (IBC) array through the Illumina Bead Studio genotyping module, version 3.2 at McGill University and Genome Quebec Innovation Centre[325]. We established a list of SNPs that reached genome-wide significance ($P < 5x10^{-8}$) with BMI or binary obesity status in literature. Three different approaches were implemented for the SNP selection process using a key word search (e.g. BMI) on i) the National Human Genome Research Institute (NHGRI) GWAS Catalog (www.genome.gov/gwastudies/) ii) the HuGE Navigator GWAS Integrator (www.hugenavigator.net/HuGENavigator/gWAHitStartPage.do) and iii) the PubMed database (www.ncbi.nlm.nih.gov/pubmed). This search was performed independently by two individuals (HR and DM) and this strategy generated a list of 136 independent SNPs associated with BMI and/or binary obesity status in October 2015. When a lead SNP was not present, we used the following criteria to select proxy SNPs: 1) SNPs were included in the Illumina cardiovascular gene-centric array; 2) $r^2 >0.90$ in a population of similar ancestry in which the lead SNP was identified from the 1000 Genomes Project. We used the Broad Institute website tool SNAP (SNP Annotation and Proxy Search) paired with an independent method we developed to identify proxy SNPs[402]. Of these 136 SNPs, 23 lead or proxy SNPs were available on version 1 and 2 of the IBC 50K SNP array (Supplementary Table 3). Supplementary Table 1 illustrates the genotype distributions of the 23 obesity-associated gene variants stratified by ethnicity. The SNPs showed no significant ($P < 10^{-6}$) deviation from Hardy-Weinberg Equilibrium (HWE) within the six ethnic groups. The call rate for each of the 23 SNPs was between 99.8 – 100% (Supplementary Table 1).

Phenotyping

A standardized protocol was implemented to assess demographic data and direct anthropometric measurements including height (m) and weight (kg). Trained staff at clinical centers measured standing height to the nearest 0.1 cm and weight to the nearest 0.1 kg in light clothing. Standing height measurements required the individual to stand with his/her back against a wall, look straight ahead standing with bare feet on the ground. Averages of the two measures were used in all analyses. BMI was calculated as total weight in kilograms (kg) divided by height in meters (m) squared.

Based on the 2003 ADA criteria, the oral glucose test was used to classify participants into four distinct groups: normal glucose tolerant (NGT), impaired fasting glucose (IFG), impaired glucose tolerance (IGT) and T2D[403]. IFG and IGT categories were collapsed to form the three groups (NGT, IFG/IGT, T2D) which comprised the glycemic status variable.

Average alcohol consumption (drinks/week) was measured based on self-reported information from the participant questionnaire as a six-level categorical variable: 1 = non-drinkers, 2 = 1-3 drinks/week, 3 = 4-6 drinks/week, 4 = 7-10 drinks/week, 5 = 11-14 drinks/week, 6 = $\geq$ 15 drinks/week. Based on the limited sample size of the groups who reported drinking more than three drinks/week, these six categories were collapsed to form a three-category alcohol consumption variable (non-drinkers, 1-3 drinks/week, >3 drinks/week).

Statistical Analysis

All statistical analyses were performed using SPSS (version 20, New York, USA, IBM Corporation). Statistical power calculations were performed using QUANTO

(version 1,2,4, University of Southern California, Los Angeles, CA, USA). Risk alleles for the 23 SNPs were based on existing literature and single SNP analyses were performed under the additive inheritance model.

The genetic risk score (GRS) was determined by summing the alleles of the 23 obesity risk SNPs. Based on previous work by Dudbridge, an unweighted GRS was used for these analyses to generate a range from 0 to 46[404]. Ethnic-specific imputations were assigned using the mean number of obesity predisposing alleles in successfully genotyped individuals, as previously described[405]. Before conducting the primary analyses, two preliminary verification steps were performed. First, the linear regression models were applied to test association between the 23 SNPs/GRS and BMI, and the SNPs/GRS significantly associated with BMI were carried forward to the SNP/GS x alcohol consumption interaction analysis. A significance threshold of $P<0.05$ was selected for this step based on the strong prior evidence of association between the SNPs and BMI in literature (Supplementary Table 3). This test was adjusted for sex, age, ethnicity and glycemic status. Second, the association between the SNPs/GRS and alcohol consumption was tested using ordinal logistic regression to ensure that the environmental confounding was not influencing the interaction results[406]. A Bonferroni adjusted *P* value $<2.1 \times 10^{-3}$ (0.05/24 (23 SNPs + GRS) was used as a significance threshold for this step. This analysis was adjusted for sex, age, ethnicity, glycemic status and BMI. As the primary analysis, linear regression models were used to examine (1) the association between alcohol consumption and BMI (with inclusion of an alcohol x sex interaction term) and (2) the interaction between alcohol consumption and the SNPs/GRS on BMI. These tests were adjusted for covariates including, sex, age, ethnicity, and

glycemic status. A Bonferroni correction for multiple testing was applied to the alcohol x SNP/GRS interaction analysis yielding a significance threshold of $P < 6.3 \times 10^{-3}$ (0.05/8).

RESULTS

Characteristics of the studied cohort

The baseline clinical and anthropometric characteristics of the EpiDREAM study stratified by alcohol consumption are summarized in Table 11. The mean age of the 14 828 participants in the study was 52.1 years and 62.1% of the sample was female. Participants in this analysis represented 17 of the 21 countries from which recruitment originally took place. Of the 14 828 participants, the ethnic distribution of the cohort was 54.0% European, 19.3% Latino American, 15.6% South Asian, 6.9% African, 3.0% Native North American and 1.2% East Asian. The mean BMI was 30.0 (standard deviation [SD]=6.17) kg/m^2. The majority of the sample was non-drinkers (N=9 641, 65.0%) while 15.3% (N=2 267) reported 1-3 drinks/week and 19.7% (N=2 920) reported >3 drinks/week. The unadjusted analysis suggests that alcohol intake varies by sex, age and ethnicity (Table 11).

Effect of Alcohol Intake on BMI

Increased alcohol intake was associated with decreased BMI (β= -0.25, 95% confidence interval [CI]= -0.42 to -0.07, P= 6.3×10^{-3}) and the impact of alcohol intake on BMI varied significantly by sex ($\beta_{interaction}$= -0.82, 95% CI= -1.06 to -0.58, P= 1.7×10^{-11}) (Table 12). The inverse relationship between alcohol intake and BMI was consistent among both males (β= -0.26, 95% CI= -0.42 to -0.10, P= 1.2×10^{-3}) and

females (β= -1.05, 95% CI= -1.24 to -0.86, P= 2.2 x 10^{-26}), although the magnitude of the effect was four-fold greater among females. These tests were adjusted for age, ethnicity and glycemic status.

Effect of SNPs/GS on Alcohol Intake

The association between the 23 obesity predisposing SNPs and the corresponding GRS on alcohol intake was analyzed adjusting for sex, age, ethnicity, glycemic status and BMI (Table 13). Of the 23 SNPs analyzed, none surpassed the Bonferroni adjusted threshold of significance. The association between the obesity risk GRS and alcohol intake was also not significant (Table 13).

Effect of SNPs/GRS on BMI

At baseline, significant associations were observed between the obesity risk alleles of seven SNPs and BMI: *TNNI3K* rs1514176 (β= 0.23, 95% CI= 0.09 to 0.36, P= 1.0 x 10^{-3}), *CDKAL1* rs2206734 (β= 0.25, CI= 0.09 to 0.42, P= 2.6 x 10^{-3}), *FTO* rs3751812 (β= 0.49, 95% CI= 0.35 to 0.63, P= 2.6 x 10^{-11}), *GIPR* rs11671664 (β= 0.23, 95% CI= 0.01 to 0.44, P= 0.04), *TAL1* rs2984618 (β= 0.31, 95% CI= 0.17 to 0.44, P= 7.5 x 10^{-6}), *NT5C2* rs3824755 (β= 0.24, 95% CI= 0.05 to 0.43, P= 0.01) and *TCF7L2* rs7903146 (β= 0.45, 95% CI= 0.30 to 0.59, P= 2.1 x 10^{-9}). The GRS was also significantly associated with greater BMI (β= 0.12, 95% CI= 0.09 to 0.15, P= 2.1 x 10^{-14}) (Table 14).

Alcohol Intake x SNP/GRS Interaction Analysis

The interaction analysis was restricted to the seven SNPs and GRS, which displayed an association (P<0.05) with BMI. A significant interaction was observed between the *FTO* rs3751812 and alcohol intake on BMI (β= -0.23, 95% CI= -0.40 to -0.06, P= 7.0 x 10^{-3}) (Table 14). This suggests that increased alcohol intake decreased the impact of the *FTO* rs3751812 risk allele (T) on BMI. The subgroup analysis demonstrated that each additional *FTO* rs3751812 risk allele (T) was associated with a 0.53 increase in BMI among non-drinkers (CI= 0.35 to 0.71, P= 1.1 x 10^{-8}), a 0.44 increase in BMI among those consuming 1-3 drinks/week (CI= 0.12 to 0.76, P= 6.5 x 10^{-3}) and a 0.17 increase in BMI among those consuming over 3 drinks/week (CI= -0.11 to 0.44, P= 0.23). This suggests that the impact of *FTO* rs3751812 can change by up to 68% depending on the frequency of alcohol intake (Figure 2).

The interactions between alcohol consumption and the six remaining SNPs (*TNNI3K* rs1514176, *CDKAL1* rs2206734, *GIPR* rs11671664, *TAL1* rs2984618, *NT5C2* rs3824755, *TCF7L2* rs7903146) were not significant. The interaction between alcohol consumption and the GRS was also not significant for BMI (Table 14).

DISCUSSION

In the present study, we analyzed the complex relationship between alcohol consumption and BMI from an observational and genetic epidemiology perspective using a multi-ethnic study. Our results indicated that greater alcohol consumption was associated with decreased BMI, and the magnitude of this effect was four-fold greater among females compared to males. Increased alcohol consumption also attenuated the

impact of *FTO* rs3751812 variation on BMI, suggesting that the association between alcohol consumption and BMI may be partially explained by gene-environment interactions. We did not observe any significant evidence of direct association between the 23 SNPs or GRS and alcohol consumption.

The negative association between alcohol consumption and BMI has been observed in several existing studies[407-413]. These investigations include large-scale longitudinal studies in the United States (N=7 230 - 19 220, follow-up 7 - 12.9 years), and have reported larger inverse associations among women compared to men[407,408,410-412]. This finding is consistent with the alcohol intake x sex interaction identified in our study and is also supported by a cohort study of 43 543 Danish men and women followed prospectively for five years[409]. Previous evidence indicates that men add alcohol to their daily dietary intake whereas alcohol consumption tends to replace the intake of other macronutrients, particularly carbohydrates among women[408], without increasing total energy intake[414-417]. Women also display reduced alcohol dehydrogenase activity compared to men and are therefore more likely to metabolize ethanol via the hepatic microsomal ethanol-oxidizing system (MEOS)[414,418,419]. The MEOS system is stimulated by chronic alcohol intake and produces ATP less efficiently than alcohol dehydrogenase oxidation since most of the energy derived from MEOS metabolism of alcohol increases heat production[412,414]. Another explanation supported by metabolic studies suggests that the acute thermogenic effect induced by low doses of alcohol consumption varies by sex[420]. In men, the change in energy expenditure associated with alcohol consumption is modest, while in women the energy expenditure increased substantially beyond the energy content of the alcohol consumed[421,422].

Although three previous studies have reported that risk allele carriage in *FTO* is associated with decreased alcohol intake, we did not find a significant association in our study following Bonferroni correction[144,386,423]. However, the relationship between *FTO* rs3751812 and alcohol consumption displayed nominal evidence of association and the direction of effect was consistent with the existing studies (Table 3)[144,386,423]. Challenges with replicating this finding were also reported in a previous replication effort that analyzed three independent cohorts[397]. These results reinforce the difficulty associated with identifying gene pleiotropy between obesity and drinking patterns, with the exception of *ALDH2*[398,399]. It is important to note that we did not analyze extreme phenotypes (e.g. alcohol addiction), and we cannot exclude association with specific alcohol consumption patterns, such as binge drinking. Finally, ethnic or country specific trends in drinking patterns and ethnic variations in the alcohol associated variants may explain the lack of association between *FTO* rs3751812 and alcohol consumption in this study.

The only gene-environment interaction observed in the present study involved alcohol consumption and *FTO* rs3751812, whereby increased alcohol intake appeared to attenuate the impact of the *FTO* rs3751812 risk allele (T) on BMI by 68%. Although this interaction was nominally significant, a large-scale study (N=119 132) from the UK biobank also identified an alcohol x *FTO* interaction with a similar direction of effect, suggesting that this interaction may be valid and the nominal significance level could be attributed to a lack of power. There is also additional evidence to support this association. First, variation in *FTO* has been shown to interact with a variety of lifestyle factors including physical activity, diet patterns and socioeconomic status[18,164,424]. Second, *FTO*

functions as a demethylase and differential methylation and BMI variance have been linked to variation in intron 1 of *FTO*[285,425]. Third, reduction in S-adenosylmethionine levels associated with ethanol consumption inhibits methyl group transfer and contributes to DNA hypomethylation[426-428]. In addition, alcohol consumption alters the NAD^+/NADH ratio and produces reactive oxygen species and acetate, both of which can influence epigenetic regulatory mechanisms[428]. Disrupting DNA methylation through these pathways may compromise gene regulation at the transcription level[428-430]. Since DNA methylation is influenced by alcohol consumption, DNA methylation may represent the biological mechanism that is influenced by both alcohol consumption and *FTO* variation[431,432].

Strengths of this study include the large sample size, the multiethnic sample and the international recruitment. This study was limited by the lack of detail involved in the assessment of alcohol consumption. The type of alcohol consumed (e.g. beer, wine, spirits), patterns of consumption (e.g. binge drinking) and presence of disorders (e.g. alcoholism) were not assessed, which limits the complexity of the analysis and creates the potential for residual confounding. Assessing alcohol intake by self-report also creates the possibility for recall bias, such as socially desirable reporting. Another limitation is that is that since a portion of the study sample is dysglycemic, it is possible that some of the participants may have been taking medications that restricted them from consuming alcohol, or perhaps interfered with the metabolism of alcohol. The SNPs selected for this study may not be ideal proxies in each ethnic group since most were originally identified in European populations. Finally, the 23 SNPs analyzed do not represent the complete list of currently identified obesity risk variants.

In summary, we observed an inverse relationship between alcohol consumption and BMI in a multiethnic cohort recruited from 17 countries. Although this effect was identified in both men and women, the size of the effect was four-fold greater among females. We were not able to replicate the association between *FTO* variation and alcohol consumption, yet the direction of effect was consistent with previous studies. An alcohol x *FTO* interaction was observed whereby increased alcohol consumption appeared to attenuate the impact of *FTO* rs3751812 on BMI by 68%. While this interaction has been identified in previous studies, this is the first report of this interaction in a multi-ethnic sample and further study of this association is needed to determine the impact for public health intervention and prevention strategies.

Table 11. Baseline characteristics stratified by alcohol consumption in the EpiDREAM study.

	Category	Non drinkers	1-3 drinks/week	>3 drinks/week	All	P-value
Total at baseline N(%)		9641 (65.0%)	2267 (15.3%)	2920 (19.7%)	**14828 (100%)**	
Gender N(%)	Male	2786 (28.9%)	1032 (45.5%)	1803 (61.7%)	5621 (37.9%)	3.8×10^{-24}
	Female	6855 (71.1%)	1235 (54.5%)	1117 (38.3%)	9207 (62.1%)	
[a]Age (years)		51.13 ±11.52	52.43 ± 11.02	55.10 ± 10.59	52.11 ± 11.37	1.9×10^{-60}
Glycemic status N (%)	Normal IFG/IGT	5067 (52.6%) 4574 (47.4%)	1169 (51.6%) 1098 (48.4%)	1200 (41.1%) 1720 (58.9%)	7436 (50.1%) 7392 (49.9%)	9.2×10^{-27}
[a]BMI at baseline (kg/m^2)		30.23 ± 6.53	29.57 ± 5.59	29.42 ± 5.26	29.97 ± 6.17	2.2×10^{-11}
Ethnic groups N(%)	South Asian	2129 (22.1%)	111 (4.9%)	74 (2.5%)	2314 (15.6%)	5.8×10^{-48}
	East Asian	147 (1.5%)	12 (0.5%)	15 (0.5%)	174 (1.2%)	
	European	4214 (43.7%)	1614 (71.2%)	2185 (74.8%)	8013 (54.0%)	
	African	841 (8.7%)	105 (4.6%)	77 (2.7%)	1023 (6.9%)	
	Latino American	2035 (21.1%)	359 (15.9%)	470 (16.1%)	2864 (19.3%)	
	Native-North American	275 (2.9%)	66 (2.9%)	99 (3.4%)	440 (3.0%)	

Table 12. Effect of alcohol consumption on BMI.

Effect of alcohol consumption on BMI (adjusted for age, ethnicity, and glycemic status)			
	β	**95% CI**	**P-value**
Sex	1.10	0.85 to 1.35	1.1×10^{-17}
Alcohol	-0.25	-0.42 to -0.07	6.3×10^{-3}
Alcohol x sex	-0.82	-1.06 to -0.58	1.7×10^{-11}
Subgroup analysis			
Females	-1.05	-1.24 to -0.86	2.2×10^{-26}
Males	-0.26	-0.42 to -0.10	1.2×10^{-3}

Table 13. Effect of SNPs/GRS on alcohol consumption (adjusted for sex, age, ethnicity, glycemic status and BMI).

SNP	gene	OR (95% CI)		P-value
rs1514176	*TNNI3K*	1.00	(0.95 to 1.05)	0.88
rs6235	*PCSK1*	0.98	(0.92 to 1.05)	0.61
rs6232	*PCSK*	0.96	(0.84 to 1.10)	0.55
rs2206734	*CDKAL1*	0.98	(0.92 to 1.04)	0.54
rs2272903	*TFAP2B*	1.02	(0.95 to 1.11)	0.54
rs1211166	*NTRK2*	0.97	(0.91 to 1.03)	0.38
rs6265	*BDNF*	1.09	(1.02 to 1.17)	0.01
rs1401635	*BDNF*	1.01	(0.95 to 1.07)	0.87
rs997295	*MAP2K5*	1.00	(0.95 to 1.05)	0.93
rs7203521	*FTO*	1.05	(1.00 to 1.11)	0.06
rs3751812	*FTO*	0.94	(0.89 to 0.99)	0.02
rs1805081	*NPC1*	1.02	(0.97 to 1.08)	0.47
rs2075650	*APOE*	1.09	(1.01 to 1.17)	0.03
rs11671664	*GIPR*	0.96	(0.88 to 1.04)	0.29
rs2984618	*TAL1*	1.01	(0.96 to 1.07)	0.67
rs1011527	*LEPR*	1.00	(0.91 to 1.09)	0.97
rs7605927	*POMC*	0.98	(0.93 to 1.04)	0.49
rs611203	*USP37*	1.04	(0.99 to 1.10)	0.12
rs2535633	*ITIH4*	0.98	(0.93 to 1.03)	0.44
rs3824755	*NT5C2*	0.96	(0.89 to 1.04)	0.31
rs7903146	*TCF7L2*	0.98	(0.92 to 1.03)	0.42
rs671	*ALDH2*	1.04	(0.98 to 1.08)	0.62
rs749767	*KAT8*	0.97	(0.92 to 1.02)	0.19
rs12617233	*FANCL*	0.97	(0.92 to 1.02)	0.24
	GRS	1.00	(0.99 to 1.01)	0.82

Table 14. Main association between SNPs/GRS and BMI, and interaction effects with alcohol consumption (adjusted for sex, age, ethnicity and glycemic status).

		Main SNP effect			Interaction analysis (SNP/GRS x alcohol)		
SNP	gene	β (95% CI)		P-value	β (95% CI)		P-value
rs1514176	*TNNI3K*	0.23	(0.09 to 0.36)	**$1.0x10^{-3}$**	0.13	(-0.03 to 0.30)	0.11
rs6235	*PCSK1*	0.06	(-0.11 to 0.22)	0.47		-	
rs6232	*PCSK*	0.01	(-0.34 to 0.35)	0.96		-	
rs2206734	*CDKAL1*	0.25	(0.09 to 0.42)	**$2.6x10^{-3}$**	-0.20	(-0.40 to 0.01)	0.06
rs2272903	*TFAP2B*	0.18	(-0.01 to 0.37)	0.07		-	
rs1211166	*NTRK2*	0.08	(-0.08 to 0.24)	0.30		-	
rs6265	*BDNF*	0.09	(-0.10 to 0.27)	0.35		-	
rs1401635	*BDNF*	0.13	(-0.03 to 0.28)	0.11		-	
rs997295	*MAP2K5*	0.05	(-0.08 to 0.18)	0.46		-	
rs7203521	*FTO*	-0.04	(-0.17 to 0.11)	0.63		-	
rs3751812	*FTO*	0.49	(0.35 to 0.63)	**$2.6x10^{-11}$**	-0.23	(-0.40 to -0.06)	**$7.0x10^{-3}$**
		Subgroup analysis	Non-drinkers		0.53	(0.35 to 0.71)	$1.1x10^{-8}$
			1-3 drinks/week		0.44	(0.12 to 0.76)	$6.5x10^{-3}$
			>3 drinks/week		0.17	(-0.11 to 0.44)	0.23
rs1805081	*NPC1*	0.05	(-0.10 to 0.19)	0.54		-	
rs2075650	*APOE*	0.09	(-0.11 to 0.28)	0.39		-	
rs11671664	*GIPR*	0.23	(0.01 to 0.44)	**0.04**	-0.12	(-0.38 to 0.15)	0.40
rs2984618	*TAL1*	0.31	(0.17 to 0.44)	**$7.5x10^{-6}$**	-0.02	(-0.18 to 0.15)	0.86
rs1011527	*LEPR*	-0.13	(-0.35 to 0.09)	0.23		-	
rs7605927	*POMC*	-0.05	(-0.19 to 0.10)	0.54		-	
rs611203	*USP37*	-0.06	(-0.20 to 0.07)	0.38		-	
rs2535633	*ITIH4*	0.00	(-0.14 to 0.13)	0.95		-	
rs3824755	*NT5C2*	0.24	(0.05 to 0.43)	**0.01**	0.12	(-0.12 to 0.37)	0.32
rs7903146	*TCF7L2*	0.45	(0.30 to 0.59)	**$2.1x10^{-9}$**	-0.01	(-0.19 to 0.16)	0.88
rs749767	*KAT8*	0.06	(-0.08 to 0.20)	0.40		-	
rs12617233	*FANCL*	0.04	(-0.09 to 0.18)	0.53		-	
	GRS	0.12	(0.09 to 0.15)	**$2.1x10^{-14}$**	-0.03	(-0.07 to 0.01)	0.18

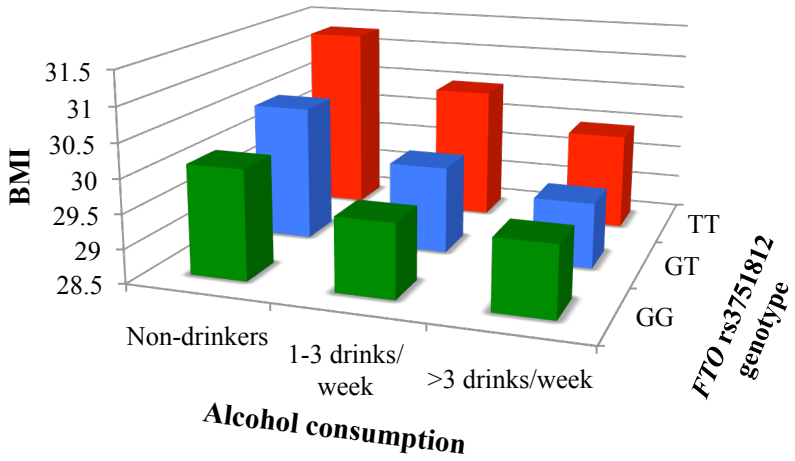
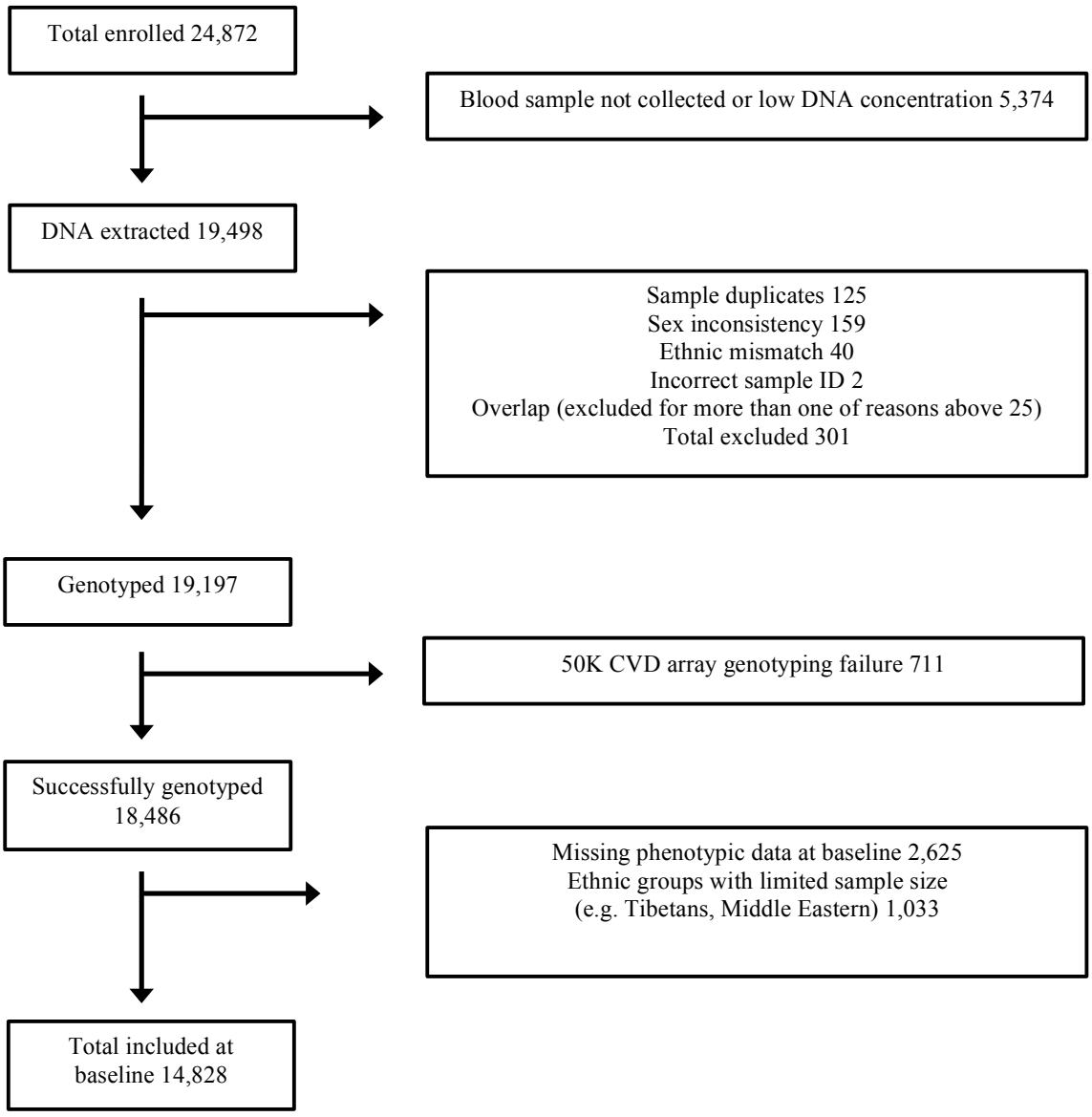


Figure 2. Mean BMI values stratified by alcohol consumption and *FTO* rs3751812 genotype.

Supplementary Table 3. Description of the 23 SNPs selected in EpiDREAM.

Gene	SNP	Risk Allele	Other Allele	Traits	Chromosome	Publication
TNNI3K	rs1514176	G	A	BMI	1	Speliotes, Nat Genet 2010
PCSK1	rs6235	C	G	Obesity	5	Benzinou, Nat Genet 2008
	rs6232	G	A			
CDKAL1	rs2206734	C	T	BMI	6	Okada, Nat Genet 2012
BDNF	rs6265	G	A	BMI	11	Thorleifsson, Nat Genet 2009
	rs1401635	C	G			
FTO	rs3751812	T	G	BMI	16	Frayling, Science 2007
	rs7203521	A	G			Thorleifsson Nat Genet 2009
NPC1	rs1805081	A	G	Obesity	18	Meyre Nat Genet 2009
TFAP2B	rs2272903	G	A	BMI	6	Speliotes, Nat Genet 2010
NTRK2	rs1211166	A	G	BMI	9	Guo, Hum Mol Genet 2013
MAP2K5	rs997295	T	G	BMI	15	Speliotes, Nat Genet 2010
APOE	rs2075650	A	G	BMI	19	Guo, Hum Mol Genet 2013
GIPR	rs11671664	G	A	BMI	19	Speliotes, Nat Genet 2010
TAL1	rs2984618	T	G	BMI	1	Locke et al., Nat Genet 2015
LEPR	rs1011527	A	G	Obesity	1	Wheeler Nat Genet 2013
POMC	rs7605927	G	C	BMI	2	Graff Hum Mol Genet 2013
USP37	rs611203	G	A	BMI	2	Locke et al., Nat Genet 2015
ITIH4	rs2535633	G	C	BMI	3	Wen Hum Mol Genet 2014
NT5C2	rs3824755	C	G	BMI	10	Wen Hum Mol Genet 2014
TCF7L2	rs7903146	C	T	BMI	10	Locke et al., Nat Genet 2015
KAT8	rs749767	A	G	BMI	16	Locke et al., Nat Genet 2015
FANCL	rs12617233	C	T	BMI	2	Guo et al., HMG 2012



Supplementary Figure 6. Flow chart of EpiDREAM study (alcohol x gene interaction).

CHAPTER 6-CUMULATIVE ENVIRONMENTAL RISK AND GENETIC PREDISPOSITION TO OBESITY

ABSTRACT

Background: Several gene-environment interaction studies have identified lifestyle factors that can moderate the impact of obesity predisposing gene variants. However, the extent to which the cumulative effect of multiple lifestyle factors influences the genetic risk to obesity is uncertain.

Methods: The data for this analysis were collected from the EpiDREAM study (17,373 participants from six ethnic groups). Exposure to 12 lifestyle factors was measured and combined into a cumulative environmental score. We analyzed the interaction between this score and variation in *FTO* rs3751812 and a genetic risk score (GRS) on body mass index (BMI).

Results: The genetic risk score, *FTO* rs3751812 and the cumulative environmental score were significantly associated with BMI. We also observed a significant interaction whereby the accumulation of unhealthy lifestyle exposures increased the impact of *FTO* rs3751812 variation on BMI. The interaction between the GRS and the cumulative environmental score was not statistically significant.

Conclusion: In the present study, genetic and lifestyle were independently associated with BMI in a large multi-ethnic sample. These findings also demonstrate that the accumulation of unhealthy lifestyle factors can create a two-fold increase in the impact of *FTO* rs1751812 on BMI.

INTRODUCTION

Obesity has become a worldwide epidemic, with 600 million adults classified as obese in 2014[433]. It is associated with several comorbidities including type 2 diabetes mellitus, cardiovascular disease, certain cancers and psychological distress[434]. Furthermore, obesity can reduce life expectancy by approximately 8-13 years[3]. Although treatments such as lifestyle modifications, pharmacotherapy and bariatric surgery are available, they offer modest long-term benefit to the patient[435-437]. While obesity is challenging to treat, it is also preventable, making effective prevention a critical aspect of the solution to this epidemic[438]. Therefore, a comprehensive understanding of the genetic and environmental determinants is crucial to develop effective prevention strategies.

The recent development of the obesity epidemic has been attributed to an 'obesogenic' environment resulting from lifestyle and societal changes[439]. It constitutes obesity-promoting factors such as a high calorie diet, physical inactivity, lack of sleep, and increased stress[439-441]. While this environment creates favourable conditions for the development of obesity, it is not solely responsible for this epidemic. Biological factors such as ethnicity, age, and sex determine which individuals develop obesity in unhealthy environments[47]. If ethnic-dependent prevalence of obesity can be explained in part by lifestyle differences, admixture studies elegantly demonstrated that the genetic landscape was an important contributor as well[442]. The genetic basis of obesity has also been established through twin and family studies that report heritability estimates of 40-70% [352]. The genetic determinants of obesity can be classified into two categories: monogenic and polygenic[443]. Monogenic forms of obesity can be syndromic or non-syndromic, with more than 70 syndromic forms identified and 11 human genes associated with non-

syndromic forms [353,444,445]. Polygenic forms of obesity result from the accumulation of multiple common gene variants with modest effects and genome-wide association studies have identified over 160 independent loci associated with body mass index (BMI) and obesity status[353,446].

Obesity is also influenced by the interaction of predisposing genes and environmental exposures, likely through epigenetic regulation[447,448]. Interaction between genetic and environmental factors may increase the prevalence of obesity synergistically. Gene-environment interaction studies (GEI) have demonstrated that heritability estimates for BMI are affected by the environment (e.g. physical activity)[72,449]. Even highly penetrant forms of monogenic obesity are sensitive to the obesogenic environment[46]. Several studies have demonstrated interaction between an *FTO* variant or genotype score and environmental factors such as age, diet, physical activity and socioeconomic status [47,450]. Ahmad *et al.* investigated the cumulative effects of physical activity and caloric intake on the *FTO* rs8050136 variant and demonstrated that the risk allele had the greatest effect in the low physical activity-high calorie intake subgroup[141].

While these studies demonstrate that there is an interaction between obesity genes and environmental factors, there are several limitations in current literature. It is unclear whether the interaction between an obesogenic environment and genotype scores is driven by *FTO* variant, considering its major genetic contribution in comparison with other variants[148]. In addition, most studies have been conducted in people of European ancestry and in high-income populations, which limits the generalizability of results on a global scale[451,452]. Lastly, the vast majority of GEI studies have analyzed the interaction between genetic risk factors and individual environmental exposures, yet the cumulative

effect of only two of these factors on genetic predisposition has only been explored in one previous study[141]. This prompted us to investigate the effect of the *FTO* rs3751812 variant and a genetic risk score (GRS) on adiposity in interaction with cumulative risk environments. We used data from a multiethnic, international study that includes measurements of multiple environmental exposures and adiposity. Our specific objective was to investigate the interaction between this gene variant and the GRS, and a cumulative measure of environmental risk on obesity.

METHODS

Study Participants

Data from the EpiDREAM cohort was used for this investigation. This study enrolled 24 872 individuals from 17 countries and from six major ethnic groups (South Asian, East Asian, European, African, Latin American, Native North American). Participants were screened for eligibility to enter the DREAM clinical trial. Eligibility was based on being at risk for type 2 diabetes (T2D) based on family history, ethnicity and abdominal adiposity, and participants were screened using a 75-gram oral glucose tolerance test (OGTT). All participants were screened between July 2001 and August 2003 and were between the ages of 18-85 years. Our analysis focused on 17 373 individuals with complete phenotypic and 50K gene-centric array information. The eigensoft software (http://genepath.med.harvard.edu/~reich/Software.htm) was used to validate self-reported ethnicity and samples that did not group with participants of the same self-reported ethnicity were excluded from the analysis. In accordance with the

Declaration of Helsinki, all individuals provided informed consent prior to participating and local ethics committees have approved the EpiDREAM study.

Genotyping

DNA was extracted from the buffy coats of 19 498 participants using the Gentra System (Supplementary Figure 1). Genotyping was performed at the McGill University and Genome Quebec Innovation Centre using the Illumina Bead Studio genotyping module, version 3.2 and the Illumina CVD bead chip microarray ITMAT Broad Care (IBC) array [325]. Three strategies were employed to identify SNPs associated with BMI or binary obesity status that reached genome-wide significance ($P<5x10^{-8}$) in European populations: i) the National Human Genome Research Institute (NHGRI) GWAS Catalog (www.genome.gov/gwastudies/) ii) the HuGE Navigator GWAS Integrator (www.hugenavigator.net/HuGENavigator/gWAHitStartPage.do) iii) the PubMed database (www.ncbi.nlm.nih.gov/pubmed). This strategy produced a list of 72 independent SNPs, 23 of which were available on versions 1 and 2 of the IBC 50K SNP array (Supplementary Table 1): rs3751812, rs7203521 in *FTO*, rs1514176 in *TNNI3K*, rs6265 and rs1401635 in *BDNF*, rs1805081 in *NPC1*, rs6232, rs6235 in *PCSK1*, rs2206734 in *CDKAL1*, rs2075650 in *TOMM40/APOE/APOC1*, rs2272903 in *TFAP2B*, rs997295 in *MAP2K5*, rs1211166 in *NTRK2*, rs11671664 in *GIPR*, rs2984618 in *TAL1*, rs1011527 in *LEPR*, rs7605927 in *POMC*, rs611203 in *USP37*, rs2535633 in *ITIH4*, rs3824755 in *NT5C2*, rs7903146 in *TCF7L2*, rs671 in *ADLH2*, rs749767 in *KAT8*. None of these SNPs showed significant deviation from Hardy-Weinberg Equilibrium (HWE) in the six ethnic groups ($P<0.001$). The call rate for each of the 23 SNPs was between 99.8-

100% (Supplementary Table 1). For the purpose of this analysis we focused the interaction tests to a genetic risk score (GRS) comprised of these 23 SNPs and *FTO* rs3751812 since variation in this SNP has been shown to interact with multiple lifestyle factors and there is strong biological evidence to support gene-environment interactions involving this SNP[150,164,453].

Phenotyping

Participants also completed a questionnaire that included demographic data, medical history. Lifestyle information was collected by self-report and included physical activity, alcohol consumption, depression status, employment status, stress at home, financial stress, stressful life event, income, marital status, sleep duration, smoking status and TV watching (Table 15). Trained medical staff measured weight and standing height to the nearest 0.1 kg and 0.1 cm using a standardized protocol. Hip circumference was measured in duplicate at the level of the greater trochanters using a non-flexible tape measure with an attached spring balance with a mass of 750g and averages of the two measures were used in all analyses. Body mass index (BMI) was calculated as weight in kilograms (kg) divided by height in meters (m) squared.

The OGTT results and the 2003 ADA criteria were used to classify participants as having normal glucose tolerance (NGT), impaired fasting glucose (IFG), impaired glucose tolerance (IGT), or T2D at baseline: normoglycemia=fasting plasma glucose < 5.6 mmol/L, IFG=fasting plasma glucose of 5.6 to 6.9 mmol/L, IGT=fasting plasma glucose < 7.0 mmol/L and a 2-h glucose between 7.8 and 11.0 mmol/L, and diabetes=fasting plasma glucose ≥ 7.0 mmol/L or the 2-h glucose ≥ 11.1 mmol/L [327].

After combining the IFG and IGT groups, the glycemic status variable included three categories (normoglycemia, IFG/IGT, diabetic).

Statistical Analyses

Statistical analyses were performed using SPSS (version 20, New York, IBM Corporation). Single SNP analyses were performed under the additive model, and the obesity risk alleles previously identified for each of the 23 SNPs in literature were used as the risk allele. Linear regression models were used to examine (1) the association between each of the lifestyle factors (individually in separate models) and BMI (2) the association between *FTO* rs3751812 and the GRS on BMI. These tests were adjusted for covariates including, sex, age, ethnicity and glycemic status. To determine if there was multicollinearity between the lifestyle factors, the variance inflation factor was calculated for each of the 12 lifestyle factors. Since the variance inflation factor did not exceed a value of two for any of the lifestyle factors, they were all retained in the analysis [454]. We expect that the lack of correlation among the variables could be attributed to the multi-ethnic international recruitment, whereby these lifestyle factors may cluster differently among participants in different ethnic groups or living in different countries. To combine the information from each of the 12 lifestyle factors, we created a cumulative environmental score. This score was generated by summing the values for each of the lifestyle factors that displayed a significant association (P<0.05) with BMI. The score was weighted based on the regression coefficients between each environmental factor and their association with BMI. Participants missing data for more than five of the lifestyle variables were excluded from the analysis. The genetic risk score was calculated by summing the alleles of the 23 obesity predisposing SNPs so that the score could range

from 0 to 46. Since weighting has been shown to have no major impact on the effect of a GRS, [328] an unweighted GRS was used. We performed imputations for missing genotypic values as previously described [329] using the mean number of predisposing obesity alleles in successfully genotyped individuals. This procedure was performed separately in each ethnic group. Individuals with more than one missing genotype out of 23 were not included in the genetic risk score calculation. The association between *FTO* rs3751812 and the GRS with the cumulative environmental score was analyzed to ensure that the interacting terms were independent. Linear regression models were used to test the gene-environment interactions with the inclusion of a cumulative environmental score x *FTO* rs3751812 interaction term, and a cumulative environmental score x GRS interaction term. The interaction between the cumulative environmental score and *FTO* rs3751812, and the cumulative environmental score and the GRS were analyzed in separate models. These tests were also adjusted for covariates including, sex, age, ethnicity and glycemic status. Two-tailed P-values are presented in this manuscript and $P<0.05$ were considered as statistically significant.

RESULTS

Characteristics of the studied cohort

Table 15 summarizes the baseline clinical and anthropometric characteristics of participants included in the EpiDREAM study. The average BMI was 30.16 (SD = 6.22), the mean age was 52.66 years (SD = 11.38), and the majority of individuals included in the study were female (61.0%). The sample represented six ethnic groups, including South Asian (15.8%), East Asian (1.3%), European (53.9%), African (7.2%), Latino

(18.9%) and Native North American (2.9%). Participants in this study had both phenotypic and 50K gene-centric array information available.

Effect of Lifestyle Factors on BMI

The effect of the 12 lifestyle factors on BMI is summarized in Table 16. Of the 12 lifestyle factors, 10 were significantly associated with BMI. Seven lifestyle factors were associated with increased BMI: depression (β= 0.87, SE= 0.12, P= 1.41 x 10^{-13}), employment status (β= 0.08, SE= 0.04, P= 0.04), stress at home (β= 0.52, SE= 0.11, P= 1.36 x 10^{-6}), financial stress (β= 0.53, SE= 0.12, P= 1.11 x 10^{-5}), stressful life event (β= 0.49, SE= 0.09, P= 1.04 x 10^{-7}), smoking status (β= 0.42, SE= 0.05, P= 8.96 x 10^{-16}) and TV watching (β= 0.44, SE= 0.03, P= 3.31 x 10^{-53}). In contrast, three lifestyle factors displayed negative associations with BMI: physical activity (β= -1.59, SE= 0.08, P= 5.38 x 10^{-95}), alcohol consumption (β= -0.40, SE= 0.04, P= 4.78 x 10^{-25}) and sleep duration (β= -0.34, SE= 0.07, P= 3.40 x 10^{-6}). These ten environmental factors were combined into a weighted cumulative environmental score with a mean of 4.77 (SD = 1.45).

Effect of FTO rs3751812 and GRS on BMI

The effect of the *FTO* rs3751812 risk allele on BMI was statistically significant (β= 0.51, SE= 0.07, P= 5.93 x 10^{-14}). The GRS was also significantly associated with increased BMI (β= 0.12, SE= 0.02, P= 3.26 x 10^{-16}). This analysis was adjusted for sex, age, ethnicity and glycemic status (Table 17).

Interaction Analyses

The results of the interaction analyses between the *FTO* rs3751812 SNP, the GRS and the cumulative environmental score are presented Table 17. We observed a statistically significant interaction whereby increases in the cumulative environmental score increased the impact of *FTO* rs3751812 on BMI (β= 0.17, SE= 0.04, P= 9.8 x 10^{-5}). The subgroup analysis revealed that the effect of the risk allele in *FTO* rs3751812 increased substantially across each quartile of the cumulative environmental score: first quartile (β= 0.36, SE= 0.11, P= 1.3 x 10^{-3}), second quartile (β= 0.30, SE= 0.12, P= 0.02), third quartile (β= 0.42, SE= 0.13, P= 9.5 x 10^{-4}), fourth quartile (β= 0.83, SE= 0.15, P= 4.1 x 10^{-8}) (Figure 3). This indicates that the effect of the *FTO* rs3751812 risk allele on BMI among those with an unhealthy lifestyle can be more than twice as large as among those with a healthy lifestyle. The accumulation of unhealthy lifestyle factors increased BMI across all *FTO* rs3751812 subgroups, although the effect of the unhealthy lifestyle factors was more pronounced among the high-risk *FTO* rs3751812 subgroup (Figure 4). The interaction between the GRS and the cumulative environmental risk score was not statistically significant (β= 0.33, SE= 0.27, P= 0.22)

DISCUSSION

In the present study, we identified a significant statistical interaction between the *FTO* rs3751812 risk allele and a score reflecting the accumulation of lifestyle behaviours that increase the risk of obesity. Specifically, we found that the accumulation of unhealthy lifestyle factors significantly increased the effect of the *FTO* rs3751812 risk allele on BMI in a large multiethnic sample recruited from 17 different countries.

Specifically, the effect of the *FTO* rs3751812 risk allele among participants with a less healthy lifestyle was 2.3 times greater than the effect among those with a more healthy lifestyle. These results also indicated a threshold effect whereby the effect of *FTO* rs3751812 variation on BMI was relatively stable among the first three quartiles of the cumulative environmental score (β= 0.30-0.42), and then increased significantly among those in the fourth quartile (most unhealthy) of the cumulative environmental score (β= 0.83). While this study is the first to assess the cumulative impact of lifestyle behaviours on genetic risk to obesity, these findings are supported by previous studies demonstrating gene-environment interactions with individual lifestyle factors such as physical activity, diet patterns and alcohol intake.

As demonstrated in previous studies, there is strong biological rationale to support interaction effects involving variation in the *FTO* gene. *FTO* functions as a nucleic acid demethylase and variation in *FTO* has been linked to distinct methylation profiles and is associated with BMI [455-457]. Since lifestyle behaviours such as physical activity and diet are also known to influence DNA methylation, methylation may be a biological mechanism that "bridges" the impact of *FTO* variation and lifestyle factors on BMI [272,458]. In addition, lifestyle behaviours and variation in *FTO* influence energy expenditure through their impact on adipocyte browning in visceral and subcutaneous adipose tissue [459,460]. Recent evidence also indicates that both sedentary behaviour and physical activity were significantly associated with methylation of the obesity-related gene *FAIM2,* which suggests that multiple lifestyle factors could impact the effect of a specific obesity gene[461].

This study also demonstrates that the impact of certain obesity genes is highly malleable depending on lifestyle behaviours. The accumulation of healthy lifestyle behaviours was associated with a dose-response decrease in BMI in all genetic risk groups (Figure 4). From a public health perspective, these results support the benefits of a healthy lifestyle for all individuals[462]. The cumulative effect of the interactions is also noteworthy since the presence of many risk factors amplified the effect of *FTO* variation more than one individual risk factor (environmental dosage effect). Personalized lifestyle recommendations for obesity prevention and treatment may benefit from including information from gene-environment interaction studies, in addition to genetic information. This would allow patients to identify their specific genetic risk factors for obesity, as well as the lifestyle behaviours that can moderate the effect of the genes that they carry. It should be noted that in terms of disease prediction, *FTO* variation alone has little predictive value for obesity and genetic risk scores are more informative for this application. However, we did not observe a significant interaction between the genetic risk score and the cumulative environmental score in this study. This may be due to the limited number of obesity SNPs included in this score: only 23 of the 160 identified obesity SNPs were available in this study. It is also likely that since only certain SNPs, such as *FTO* rs3751812, interact with environmental exposures, that including SNPs with a stable impact on BMI regardless of environmental exposure will decrease the probability of detecting an interaction with a genetic risk score. As an alternative, including only SNPs that display a variable magnitude of effect with BMI across environmental risk may increase the value of gene-environment interaction tests for disease prediction and prevention. Providing this form personalized information to inform

behavioural change may provide a message of hope for individuals with an increased genetic risk for disease. Patients often perceive DNA-based risk estimates as deterministic, yet our findings, and those from other gene-environment interaction studies, provide evidence that modifiable lifestyle behaviours can substantially influence disease risk regardless of genetic predisposition [450,463]. The expectation is that disclosure of genetic information may provide a greater incentive for behavioural change among participants with a greater genetic risk for disease [462]. Although the evidence for the effectiveness of this approach is mixed, a randomized controlled trial demonstrated that disclosure of genetic information for personalized nutrition resulted in greater changes in some dietary intake patterns compared to general population-based dietary advice after a 12-month follow-up period [464].

The strengths of this study include the multiethnic sample, international recruitment and the measurement of several different environmental factors associated with obesity facilitated the calculation of a comprehensive environmental risk score. However, a major limitation of this score was the unavailability of diet information. Dietary intake is one of the two largest environmental contributors to obesity and the validity of the cumulative environmental score is threatened by the absence of this information. The self-report data collection method for the remaining lifestyle factors also creates the possibility for recall bias and socially desirable responding. Additional limitations of this study involve the genetic data. The SNPs selected may not be ideal proxies in each ethnic group since they were originally identified in samples of European ancestry and the 23 SNPs included in the GRS only represent a subset of the obesity predisposing SNPs currently identified in the literature.

In conclusion, we identified a statistical interaction between a cumulative environmental risk score and *FTO* rs3751812 on BMI in a multi-ethnic international sample. The impact of the *FTO* rs3751812 was more than twice as large among those who accumulate unhealthy lifestyle behaviours compared to those with a healthier lifestyle. The presence of multiple environmental risk factors was also cumulative and amplified the effect of *FTO* variation on BMI significantly more than one individual risk factor. This finding contributes to the evidence demonstrating that a variety of lifestyle factors can significantly moderate the impact of obesity risk genes. Integrating this message into obesity prevention initiatives may be valuable strategy to improve the effectiveness of obesity interventions.

Table 15. Baseline characteristics of the EpiDREAM study.

Variable	Frequency N (%)*
BMI (kg/m^2)	30.16 ± 6.22*
Ethnicity	
South Asian (Indian)	2762 (15.9)
East Asian (Chinese/Japanese)	225 (1.3)
European (Caucasian)	9395 (53.9)
African (Black)	1249 (7.2)
Latino (South American)	3292 (18.9)
Native (North American)	500 (2.9)
Age (years)	52.66 ± 11.38*
Gender	
Male	6805 (39.0)
Female	10618 (61.0)
Income	
$0 – $29,999	5199 (31.7)
$30,000 - $39,999	3177 (19.4)
$40,000 - $49,999	2515 (15.4)
$50,000 - $69,000	2304 (14.1)
> $70,000	3179 (19.4)
Physical Activity level	
Sedentary	4727 (27.2)
Moderately active	10529 (60.5)
Active	2151 (12.4)
Marital status	
Never married	1490 (8.6)

Common law/married	12949 (74.4)
Divorced/separated	1814 (10.4)
Widowed	1148 (6.6)
Smoking status	
Never smoker	9505 (54.6)
Current smoker	2577 (14.8)
Former smoker	5317 (30.6)
Alcohol consumption	
0 drinks/week	11329 (65.1)
1-3 drinks/week	2642 (15.2)
4-6 drinks/week	1503 (8.6)
7-10 drinks/week	987 (5.7)
11-14 drinks/week	520 (3.0)
>14 drinks/week	420 (2.4)
Depression status	
Depressed	3203 (18.4)
Not depressed	14172 (81.6)
Employment status	
Retired	53.2
Not retired	46.8
Stress	
Money stress	
Yes	3049 (17.5)
No	14374 (82.5)
Home stress	
Yes	4063 (23.3)
No	13360 (76.7)
Stressful life event	
Yes	8646 (49.6)
No	8777 (50.4)
TV (hours/day)	
0	785 (4.9)
1	3634 (22.6)
2	4889 (30.4)
3	3239 (20.2)
4	1924 (12.0)
5	738 (4.6)
6+	854 (5.3)
Sleep (hours/day)	
< 6	5927 (34.1)
6-8	9864 (56.6)
> 8	1586 (9.1)

*Denotes standard deviation

Table 16. Effect of lifestyle factors on BMI (adjusted for sex, age, ethnicity and glycemic status).

Environmental Factor	β (SE)	P-value
Physical activity	- 1.59 (0.08)	**5.38×10^{-95}**
Alcohol consumption	- 0.40 (0.04)	**4.78×10^{-25}**
Depression status	0.87 (0.12)	**1.41×10^{-13}**
Employment status	0.08 (0.04)	**0.04**
Stress at home	0.52 (0.11)	**1.36×10^{-6}**
Income	0.03 (0.03)	0.33
Marital status	0.11 (0.07)	0.13
Financial stress	0.53 (0.12)	**1.11×10^{-5}**
Sleep duration	- 0.34 (0.07)	**3.40×10^{-6}**
Stressful life event	0.49 (0.09)	**1.04×10^{-7}**
Smoking status	0.42 (0.05)	**8.96×10^{-16}**
TV watching	0.44 (0.03)	**3.31×10^{-53}**

Table 17. Main effect and interaction between cumulative environmental score and *FTO* rs3751812/GRS and BMI (adjusted for sex, age, ethnicity and glycemic status).

Variable	**β (SE)**	**P-value**
FTO rs3751812	0.51 (0.07)	**5.93×10^{-14}**
GRS	0.12 (0.02)	**3.26×10^{-16}**
Cumulative environmental score	0.86 (0.04)	**3.0×10^{-87}**
GRS x Cumulative environmental score	0.33 (0.27)	**0.22**
FTO rs3751812 x Cumulative environmental score	0.17 (0.04)	**9.8×10^{-5}**
Subgroup analysis	**β (SE)**	**P-value**
1st quartile of cumulative environmental score	0.36 (0.11)	**1.3×10^{-3}**
2nd quartile of cumulative environmental score	0.30 (0.12)	**0.02**
3rd quartile of cumulative environmental score	0.42 (0.13)	**9.5×10^{-4}**
4th quartile of cumulative environmental score	0.83 (0.15)	**4.1×10^{-8}**

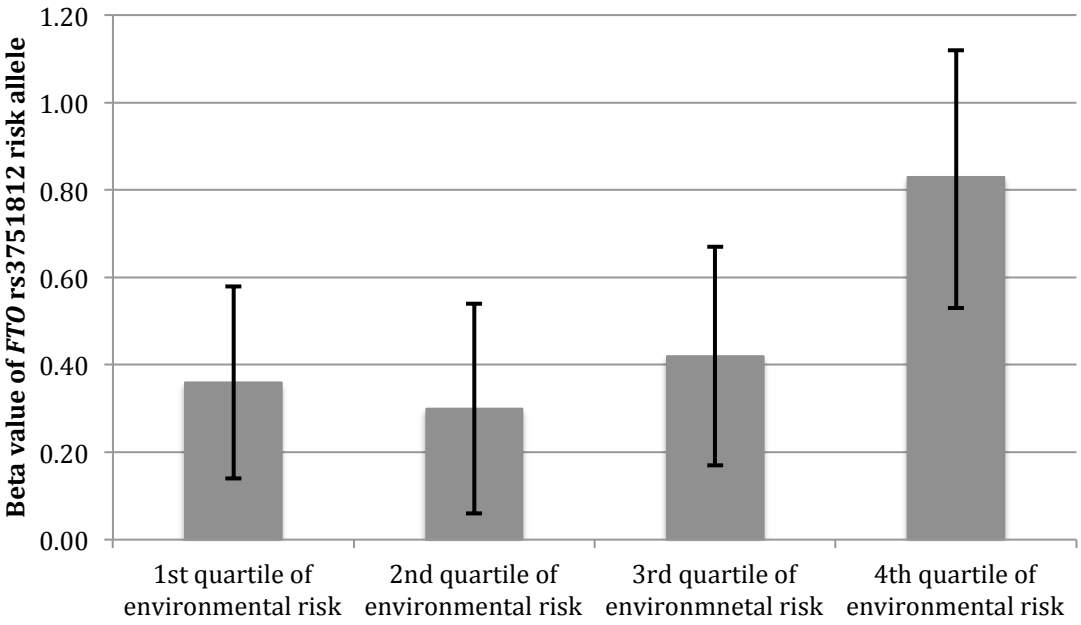


Figure 3. Beta value for the effect of the *FTO* rs3751812 risk allele on BMI in each quartile of environmental risk.

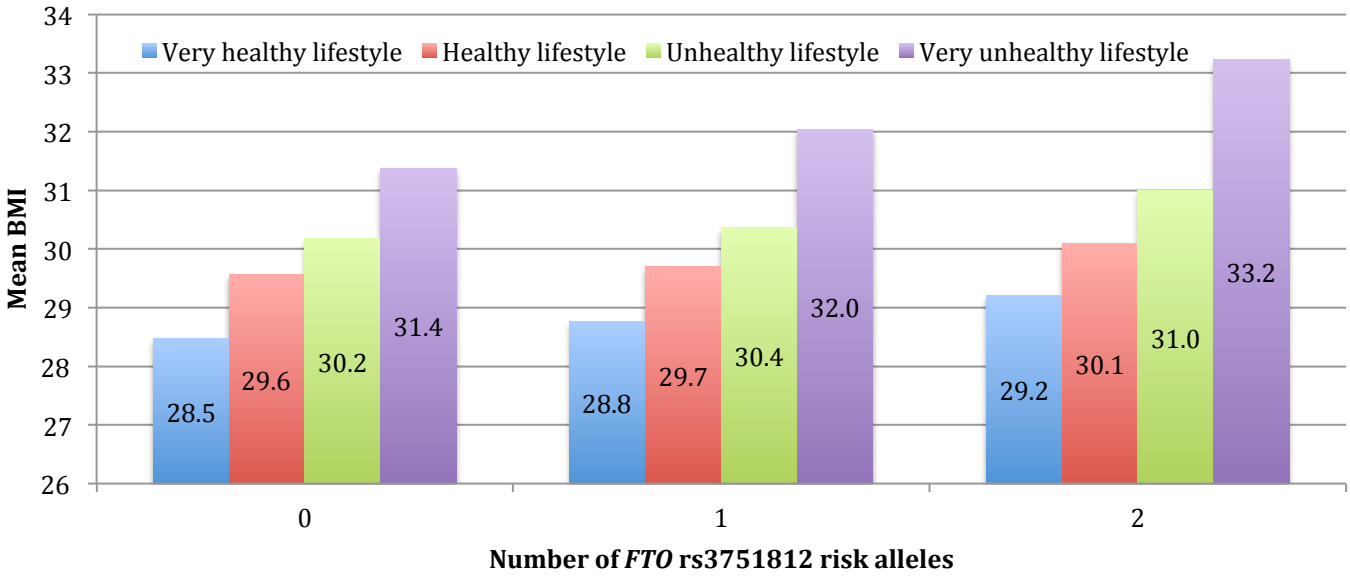


Figure 4. Mean BMI values stratified by *FTO* rs3751812 genotype and lifestyle quartile.

CHAPTER 7-METHODOLOGICAL CONSIDERATIONS IN GENE-ENVIRONMENT INTERACTION STUDIES OF OBESITY: RECENT DEVELOPMENTS AND FUTURE OPTIONS

ABSTRACT

Although the number of gene-environment interaction (GEI) analyses has increased rapidly in recent years, particularly in the obesity field, some skepticism remains for the validity of these studies based on several issues, which include statistical modelling, confounding, low replication rate, underpowered analyses, biological assumptions and measurement precision. In this review, we outline these issues and provide potential solutions to these problems based on existing evidence. Suggestions for future research of GEI are provided. These future directions include 1) studying disorders in the context of different genetic etiologies (syndromic, polygenic, monogenic) and diverse experimental designs (observational, intervention), 2) gene-environment-wide interaction studies (GEWIS) that apply a joint test for the main genotype effect and the GEI, 3) variance prioritization and 4) combining statistical tests of interaction effects with biological data to identify variants that 'redundantly' interact with multiple environments. Applying these methods will help to exploit the potential of GEI studies and translate this information to public health policy and personalized medicine for obesity and other human complex diseases.

INTRODUCTION

Obesity is now considered to be a global epidemic that is caused by genetic and environmental risk factors [1,5]. Twin and family studies have indicated that 40-80% of the inter-individual variation in obesity-related traits can be attributed to genetic differences [291]. Despite technological innovations such as genome-wide association studies (GWAS) or next-generation sequencing (NGS), the gene variants identified to date only explain a modest proportion of this variance [291]. Interactions between gene variants and environmental exposures may modulate obesity risk and account for part of this 'missing heritability' [465]. For instance, obesity predisposing gene variants in the *FTO* gene have been found to interact with physical activity [16,466] (among other exposures), and these findings have been successfully replicated in diverse ethnic groups [151].

Interest in the area of gene-environment research is growing across multiple disciplines although there has been increasing scrutiny regarding the validity of these studies based on several issues including statistical modeling [467-469], confounding [406,467], a low replication rate [470-472], underpowered analyzes [473], lack of biological assumptions [474,475] and measurement precision [476]. The relevance of testing interactions between individual genetic variants and specific environmental exposures has also been questioned [467]. Based on these concerns, some leaders of opinion have suggested that a large proportion of significant gene x environment interaction (GEI) findings are in fact false positives [51,467]. This skepticism has been adopted by multiple journals, which have implemented stringent criteria for candidate gene and interaction studies considered for review [477,478].

This review will focus on issues in GEI studies related to (1) statistical modelling of interaction terms, (2) modelling of confounding variables, (3) timing of environmental exposure across the life span and (4) measurement of predictor and outcome variables. The final section will provide suggestions to address these issues based on existing evidence and will outline future directions of GEI research.

Statistical modeling issues in gene-environment interaction research

Although the study of gene-environment interactions is an active area of research, there has been considerable scrutiny regarding the appropriate statistical method to study these effects. Using a multiple linear regression model with the inclusion of a cross-product term signifying the product of environmental (E) and genetic (G) variables is the most common method to assess interactions [16,141,144,324]. However, the specific terms included in the model are a point of contention. Coding genetic polymorphisms is either performed to create a binary variable (under a recessive or dominant model) or a three-category variable based on an additive model, with the latter often used when the true functional model of a given marker is not known [468]. A recent analysis demonstrated that modelling gene-environment interactions with a simple cross-product term (G x E) often produces misleading results when assuming an additive model [468]. This limitation has important implications for interpreting the results of interaction effects, which were likely not considered in many existing interaction studies. First, the simple cross-product model always forces the regression lines to be ordered (0, 1, 2 and never 0, 2, 1). While this assumption may be intuitive from a biological perspective, this approach will always predict an ordered effect of genotypic differences even when the data do not reflect this

assumption [472]. Second, the differences in slopes between the adjacent regression lines are always assumed to be the same [468]. There is no rationale for this assumption and in practice, sampling error alone would be expected to create uneven differences between regression slopes [479]. Alternatively, non-linear gene-environment interaction effects may be present, which could not be estimated accurately with only a cross-product interaction term [468,480]. Third, this model constrains all three regression lines to cross at the same point when interaction effects are present [468]. This implies that there is a certain level of environmental exposure that confers the same level of risk/disease for all three different genotypes. There is no statistical or biological evidence to justify this assumption, especially since the specific genetic model is often not established for the genetic marker of interest [467]. These simulations indicate that the models using only the cross-product term are more vulnerable to Type 1 and Type 2 errors [468]. In all cases, including two additional coefficients, one to model non-linear genetic effects ($\beta_4 G^2$) and another to account for non-linear interaction effects ($\beta_5 G^2$ x E), represent the interaction (or lack of interaction) more accurately. Many authors recommend this model for genetic variants following an additive or unknown genetic model, and emphasize that failure of an interaction to match a plausible biological interaction likely indicates a false positive result [468,481]. In summary, re-conceptualizing interaction models to account for non-linear effects removes the constraints of traditional regression techniques and provides a more accurate representations of gene-environment interaction effects [468]. A complete discussion of this approach is provided by Aliev *et al* [468].

Other authors contend that traditional GEI analyses neglect to test the *a priori* hypotheses that form the basis of these studies [482]. The implicit framework adopted in

traditional GEI analyses is the diathesis-stress model of environmental action [483], which specifies that certain individuals are more vulnerable to the adverse consequences of some exposures than others [484]. As an extension of this limitation, exploratory approaches also fail to test or compare the competing predictions from alternative theoretical frameworks such as the more recent differential-susceptibility framework [485]. This theory posits that some individuals are more susceptible to not only negative exposures, but to positive environmental influences as well [486]. Based on this characterization some authors have proposed that gene variants classically referred to as 'vulnerability genes' be reclassified as 'plasticity genes' to correspond with the differential susceptibility framework [487]. This theory has been proposed by several authors [485,488], and has been applied to the study of GEI [487,489,490]. In response to these competing frameworks, many statistical criteria were developed to distinguish differential-susceptibility interactions from those representing diathesis-stress theory [486,491], and Widamen's confirmatory method appears to be the most efficient [492]. This technique directly evaluates alternative theoretical frameworks by aligning the analyses with each hypothesis [492]. Specifically, this method systematically adjusts the parameters included in the regression equation to compare different theoretical frameworks, and specifies where the regression lines (representing each genetic subgroup) will cross relative to the value of the environmental exposure [492]. With respect to the frameworks discussed above, the diathesis-stress theory models an ordinal interaction whereby the predicted outcome value for the genetically vulnerable subgroup is always less than that of the genetically low-risk group. The differential-susceptibility framework predicts that the risk of the outcome for the genetically malleable group can be higher or lower than the genetically non-malleable

group depending on the level of the environmental exposure [482]. A graphical representation of these interaction effects is provided by Bakermans-Kranenburg and van IJzendoorn [493]. It is important to note that Widaman's confirmatory approach can be used for dominant/recessive or additive genetic models, and can be applied to other forms of statistical interaction involving competing hypotheses about the nature of the interaction [482,494]. Integration of this technique into interaction studies where the theoretical framework is uncertain may help to improve the accuracy and replication rates of interaction studies. This model is supported by biological evidence from a recent study, which demonstrated that TRIM28 knockout mice are alternatively lean or obese depending on subtle environmental changes [287].

Another consideration when analyzing interaction effects is the selection of either an additive or multiplicative interaction scale [495-497]. This decision has important implications given that different scales can lead to different conclusions and consequently, different public health recommendations [498]. An additive interaction exists when the combined genetic and environmental risk is significantly greater than would be expected if their effects were additive, whereas a multiplicative interaction describes a joint genetic and environmental risk that is greater than expected from multiplying their effects [51,499]. Some authors contend that the selection of measurement scale is less crucial when the underlying biological processes are not known, and both scales can be appropriate in certain situations [500-502]. If the pathophysiology consists of a multistage process, such as cancer initiation and promotion stages, two factors that act at the same stage will generally fit an additive model and those acting at different stages will typically fit a multiplicative model [503,504]. It has also been suggested that if the main

objective is to study public health impact, an additive scale is better suited to identify heterogeneous effects across subgroups, while the multiplicative scale is more appropriate for studying disease etiology [503].

Selecting the measurement scale of the outcome is another issue that has the potential to alter the conclusions of interaction analyses [505]. As an example, a multiplicative interaction between two predictors on an outcome will interact additively if the outcome is log-transformed [481]. In situations where transformation of the outcome is necessary, this implication must be taken into account.

Confounding issues in gene-environment interaction research

Several modeling strategies have been proposed to address the impact of confounding in gene-environment interaction studies [406,467]. Variables with the potential to offer alternative explanations of an interaction are typically entered into the regression equation as covariates to control for their potential confounding effects [472]. While this method controls the influence of confounding on the main effect of the genotype and environment, it does not adjust for potential confounding of the interaction term [467]. In response, an alternative method has been proposed whereby all covariate x gene and covariate x environment interaction terms are included in the model that tests the gene x environment interaction of interest [467]. If significant covariate interactions are observed, the validity of any gene x environment interactions that are found may be compromised by the covariate and warrant additional analysis. Although there are potential objections to this modelling technique, the justifications of this approach are outlined by Keller [467]. First, there is a potential that over-fitting the model will preclude accurate measure of the

parameter estimates for the many covariate interactions. While this appears problematic, the purpose of including covariate interactions is to control for their effects on the gene x environment interaction rather than producing accurate parameter estimates. Second, multicolinearity between the many interaction terms may diminish the strength of the main gene x environment interaction. This however is the purpose of this procedure and if inclusion of the covariates weakens the main interaction, then the covariates may be significantly influencing the interaction. Lastly, it is reassuring to recognize that the gene x environment interaction term is only marginally affected if there is no 'true' relationship between the covariate and the gene x environment interaction [467]. One caveat to this approach is that shared heritability between the covariates and the outcome can introduce bias and increase the risk of false-positive results [506]. Therefore, including heritable covariates in the model should be avoided if they are associated with gene variant being tested [506].

Confounding issues can be further complicated if the interacting genetic variant and environmental exposure of interest are correlated [507]. Under these circumstances, simulations have demonstrated that uncontrolled confounding will bias the estimates of the main genetic effect and the gene-environment interaction even if the genetic and environmental factor are not associated with the outcome [508]. If the genetic variant and environmental factors are independent, this is no longer an issue as long as unmeasured environmental confounders are not related to genetic factor. The issue of gene-environment dependence has been highlighted in extreme cases where the genetic variants are associated with both the environmental factor and the outcome. For example, variants on 15q25 have been linked to both smoking behaviours and lung cancer [509-511].

As a result, some authors suggest directly analyzing the relationship between the interacting genetic variant and environmental exposure [508].

Considering time of exposure in gene-environment interaction research

Given that gene expression and silencing varies significantly throughout development, it may be important to consider time of exposure when modelling exposures that can have differential effects throughout the life cycle [512]. Evidence from toxicology research indicates that many environmental exposures display distinct dose response curves that vary based on the developmental stage at which exposure occurs [513,514]. The identification of these developmental windows suggests a need to include time of exposure as a third interacting factor when analysing gene-environment interactions [475]. However, the inclusion of a three-way interaction term dramatically increases the necessary sample size [467,515] and this information is rarely available. Simulation studies have shown that the sample size required to detect three-way interactions is four-fold the sample size necessary to detect a two-way interaction of the same magnitude [516]. Although these estimates were derived from a clinical trial design with balanced group sizes, the four-fold increase was virtually identical with unbalanced group sizes [516]. Another statistical method to address this issue involves considering environmental exposure as a time-varying factor to analyze the lag effects of gene x time-varying environment interactions [475]. Yet, the repeated measurements needed to measure lag effects are often not feasible due to the cost of repeated measurement in large studies. This constraint explains the high prevalence of cross-sectional case-control designs to study gene-environment interactions [475]. The challenge of measuring variations in the

impact of environmental exposures is compounded by changes in the heritability of the outcome across time. A meta-regression of heritability studies of BMI found that the genetic contribution to BMI varies with age: heritability was positively associated with age among child studies and negatively associated with age among studies of adults [517]. A recent genome-wide interaction meta-analysis identified 15 BMI loci that interacted with age, 11 of which had a greater effect impact in younger (<50 years) compared to older (≥ 50 years) adults [106]. Failure to address the time-varying effects of environmental exposures and heritability may account for some of the challenges with replicating gene-environment interactions [51,467,518].

Measurement issues in gene-environment interactions research

Aside from issues related to statistical modeling, the measurement of the exposure and outcome represent important considerations for gene-environment interaction research. Major determinants of power include allele frequency, genetic effect size, the magnitude of interaction effect, risk allele frequency, degree of genetic misclassification and measurement error associated with the exposure and outcome [476,515,519]. Although the trade-off between precision and feasibility is common to most study designs, the large samples required to study interaction effects make this balance particularly important. Currently, the most notable gene-environment interactions in obesity have measured diet patterns or physical activity as environmental exposures [16-19]. The gold standard criterion measure for these exposures are a 7-day weighed diary and doubly labelled water, respectively. Unfortunately, the large number of participants required for these studies have restricted the measurement of these exposures to less precise instruments. The error

associated with exposure measurement generally attenuates the estimate of the true effect size [520,521]. Similar problems occur when the outcome used is an indirect measure for the true outcome of interest. In gene-environment analyses of obesity, BMI is commonly used as the outcome [18,21,28,324], which further contributes to this error given that BMI fails to distinguish between fat and fat-free mass [522].

Previous simulations have characterized how varying different determinants of power can impact the required sample size of gene-environment interaction studies [476,515]. As an example of these analyses [476], genetic misclassification was fixed at 2.5% to be consistent with prior empirical studies [523,524] and the magnitude of effect for the common allele was also constant. With a correlation between the true and observed exposure and outcome of 0.6 and 0.7, respectively, a sample size of just over 9 500 is needed to detect an interaction at a significance of 10^{-4} with 95% power (17). However, the correlations of 0.6 and 0.7 between the true and observed exposure and outcome are unusually high for gene-environment interactions in obesity due to the cost of precise measurement tools [324,519,525]. With more typical correlations of 0.3 and 0.4, the required sample size can increase to over 100 000 participants with all other variables held fixed [476]. If precise instruments are not available to mitigate this error, performing repeated measurements is a useful strategy on condition that the error in repeated measures is not correlated [520]. As an example, performing two independent repeated measures using a tool with a validity coefficient of 0.6 increases the overall validity coefficient to almost 0.8. With all other variables being fixed, this reduces the necessary sample size by more than a factor of six [476].

Although all gene-environment interaction studies must balance the practical trade-off between sample size and measurement precision, the evidence addressed above suggest that this balance should favour better measurement [476]. The value of this approach can be reinforced with the example of physical activity measurement, a common exposure analyzed in gene-environment interactions of obesity [144,324]. Physical activity is usually assessed by questionnaire, and even comprehensive instruments that address occupational and leisure activity rarely correlate with objective measures of energy expenditure above 0.3 [526]. Specific examples include the ARIC/Baeke questionnaire with a correlation of 0.24 in men and 0.19 in women when compared to estimates of energy expenditure derived from objective accelerometer data [527]. The physical activity assessment used in the EPIC-Norfolk study displayed an overall correlation of 0.44 with objective measures, although this fell to 0.28 after adjustment for age and sex [528]. The error associated with measuring this exposure is compounded by the moderate correlation (0.5) of BMI with body fat percentage as measured by dual-energy x-ray absorptiometry [529]. Using the EPIC-Norfolk questionnaire with BMI as an outcome would require almost 90 000 participants to detect an interaction that doubled the effect of a genetic variant, when the variant is present in 20% of the population [476]. Since a doubling of genetic risk from an environmental exposure is at the upper limit of interaction effect estimates reported for common variants and exposures [530-532], some authors speculate that the majority of published interaction studies are underpowered and report "lucky" true-positives or false-positive results [51]. A recent study by our group provides an empirical example of how measurement precision can influence statistical power. We analyzed physical activity x *FTO* interactions on BMI using two measures of

physical activity: a three-level categorical variable and a quantitative estimated measure of energy expenditure. The categorical data was available in 99% of the sample while the quantitative energy expenditure data was only available in 63% of the sample. Despite this disparity, similar interactions were detected using both measures, which may suggest that the added precision of the energy expenditure data compensated for the decrease in sample size [148].

Given the sample size requirements imposed by this type of data, more direct measurement techniques have been proposed. Objective measures such as heart rate monitors carry increased precision while maintaining feasibility in moderately sized epidemiological studies [533]. Heart rate monitor data have demonstrated a correlation with the gold standard of energy expenditure methods (doubly labelled water) of 0.73 [534]. Two repeated measures can increase this correlation to over 0.88. Substituting this method of exposure measurement for questionnaire methods would decrease the necessary sample size to 9453, a decrease by a factor of 10 [476]. Therefore, the gain in precision associated with more accurate measurements of exposure appear to be less resource intensive than accruing large sample sizes. The power implications of using precise measurement techniques suggest that smaller studies with more accurate measures of exposure and outcome may be better suited to detect gene-environment interactions than large sample sizes with imprecise measurement [476]. The issue of measurement imprecision has long been debated in the nutrition field and 'deep phenotyping' strategies (measuring metabolic markers such as circulating plasma lipids as a surrogate of a high-fat diet) may be worthwhile alternatives to traditional self-report measures [535-537]. Other assessments that may mitigate the concerns associated with traditional diet measures include ad

libitum energy intake tests or analyzing the dietary information of food consumed in cafeterias or restaurants [538].

The issue of direct and indirect measurement of genetic variants also has important implications for statistical power. In many current GWAS and GEI studies, the true susceptibility loci involved in the disease etiology is not known (or unavailable) and the linkage disequilibrium (LD) between marker alleles and the actual disease loci is used to study associations between gene variants and the phenotype under study [539,540]. Since this is an indirect approach, the effect estimate will be underestimated if the LD between the two variants is incomplete ($r^2 < 1$) [541]. Previous studies have demonstrated that the sample size requirements of GEI studies can be strongly influenced by the marker allele frequency, disease allele frequency, the LD between these loci, as well as the main genetic and environmental effect, the prevalence/impact of the environmental exposure and the magnitude of the interaction [540,541]. A summary of these methodological issues is summarized in Table 18.

Future directions for gene-environment interactions and obesity

Given that specific environments can greatly impact the magnitude of genetic predisposition to obesity, the systematic study of gene-environment interactions constitutes an important field of investigation in order to inform public health strategies to prevent and manage obesity and other complex diseases. Gene-environment interaction studies in the context of various forms of obesity (monogenic, polygenic) and in diverse experimental designs (observational, interventional) [50] may lead to a better understanding of the protective or detrimental environmental exposures that modify the impact of

certain genetic variants. Existing interactions need to be studied in additional obesity-prone (e.g. response to smoking cessation, response to insulin therapy in diabetic subjects) or obesity-protective (e.g. lifestyle intervention, response to the anti-obesity drug orlistat administration or to bariatric surgery) conditions. Gene-environment interaction studies are complementary to observational epidemiology, interventional study or clinical trials, and will certainly help to elaborate efficient strategies to reverse the obesity epidemic.

Currently, GWAS for obesity-related traits have focused on the marginal gene effect ignoring gene-environment interaction entirely [542]. Gene-environment interactions are nevertheless frequent in obesity, and statistical models that do not properly account for gene-environment interactions may attenuate the marginal effect size and reduce the power to detect true associations [51,543]. Applying a joint test for a main genotype effect and gene-environment interaction may increase the power to identify an individual SNP associated with a disease outcome [544-546]. As many completed GWAS for obesity have been conducted on samples with large amounts of existing environmental data, performing gene-environment-wide interaction studies (GEWIS) in these existing datasets is a cost-effective strategy to find additional obesity-associated gene variants that interact with specific environments but have been missed by conventional GWAS [547]. Since large sample sizes and meta-analytical approaches are required to reliably detect SNPs with subtle gene-environment interaction patterns [548], GEWIS for obesity have been initiated in the context of large international obesity consortiums like GIANT [549]. Although these methods show promise, recent simulations indicate that this technique is more appropriate for analyzing interactions between genetic risk scores rather than

individual SNPs, due to the reduced power when analyzing the small effect sizes of individual SNPs [550]. As a potential solution, Marigorta and Gibson suggest selecting participants who are at a high-risk for obesity based on environmental exposure [550]. This strategy has the potential to identify environmental exposures that can modulate the impact of specific variants associated with obesity [550].

While hypothesis-free GEWIS have potential to identify gene variants that are less amenable with GWAS, there are limitations to this technique. These include the challenges of identifying adequately sized cohorts with appropriate genetic and phenotypic data, as well as issues with statistical power. As a novel alternative to these techniques, variance prioritization was developed as a method to model genetic associations with genetic variance, without requiring knowledge of the interacting variables [551]. The main effects of gene variants involved in interactions are typically associated with a large degree of variance [551]. This strategy exploits this pattern to rank and prioritize variance estimates to identify gene variants whose variance per genotype significantly varies [551].

Bayesian methods have also been developed to integrate variations in multiple SNPs within a given gene/region, and examine how an environmental exposure moderates the risk of these genetic profiles [552]. This method was applied to the Environment and Genetics of Lung Cancer Etiology (EAGLE) study and detected a smoking x genetic profile interaction that was not detected by conventional interaction tests [552]. Artificial neural networks have been applied to interaction analyses and simulations suggest that this technique may be particularly valuable for detecting non-linear penetrance and interaction effects [553]. Other analytical approaches have been

developed to test interactions while addressing the common concern of statistical power. These techniques, termed "cocktail methods," involve a three-step approach to testing genome-wide gene-environment interactions while preserving the type 1 error rate and increasing power by 30-40% under certain circumstances [554]. These three steps include screening, multiple testing and GxE testing, and current simulations of this technique have been applied to binary environmental variables, although this approach is applicable to continuous environmental data [554]. While early analysis of these novel techniques has been positive, further real data application of these methods will reveal the generalizability of these approaches.

Recent GWAS for obesity have collected phenotypic information in individuals living in a broad range of environments. While successful, this approach may fail to identify potential gene variants associated with obesity-related traits in a context dependent manner. Gene identification efforts may therefore be targeted in populations that display homogeneous environment and lifestyle factors across time and across the community, as observed in the Plain people [555]. Performing genetic association studies for adiposity change in response to a standard major environment modification (antipsychotic drug use, smoking cessation, intensive caloric restriction, anti-obesity drug therapy, obesity surgery) is another valuable way to control the environmental exposure, lower sources of heterogeneity and provide a more comprehensive molecular basis for genetic predisposition to obesity.

In order to refine the search for interaction variants, statistical GEI tests could be combined with methylation quantitative trait loci (meQTL), expression quantitative trait loci (eQTL), and protein quantitative trait loci (pQTL) to focus on SNPs with a plausible

biological mechanism for interaction [286]. Specifically, a joint test could be applied to identify genetic variants that statistically interact with a given environmental exposure (e.g. physical activity level) to modulate an outcome (e.g. BMI), with the same genetic variants being also eQTL, meQTL and/or pQTL for a given locus. Ideally, the methylation expression and protein level of the same locus would be modulated by the same environmental factor [273]. A similar test could be applied to analyse the interaction between an individual SNP and multiple environmental factors. Since methylation is influenced by several environmental exposures (physical activity [273], diet [278], sleep [556]) identifying SNPs that redundantly interact with multiple exposures may be a method to exploit this pattern. The 'Identifying REdundant Gene-environment InteractionS' (REGIS) method may increase the probability of detecting 'true' and replicable gene-environment interactions. Another avenue for future research is to study gene-environment interactions jointly in mouse and human studies [557]. The development of the clustered regularly interspaced short palindromic repeat (CRISPR) system for gene targeting and editing creates a new opportunity to study 'humanized' genetically modified mice carrying human mutations [558,559]. Combining this biological data from animal studies with statistical evidence of interaction form human epidemiological studies is also likely to improve the validity of gene-environment interaction studies [51].

Conclusion

A prolific period of discovery is foreseen in this fast-moving field, especially with the many methodological innovations that attempt to address the 'missing heritability' of obesity. To effectively tackle this knowledge gap, prospective studies need to incorporate

current evidence related to statistical modelling, confounding, biological assumptions, temporal aspects and variable measurement to optimize the validity of emerging evidence. Given the potential of GEI research to identify high-risk subgroups of the population that experience greater benefit from specific environmental exposures, the application of this knowledge may improve targeting of public health policies and further develop the field of personalized medicine to combat the obesity epidemic.

Table 18. Summary of methodological issues and solutions for gene-environment interaction studies in obesity.

Methodological Issue	Suggested Solution	Reference (Lead author
Modelling the G x E cross-product terms	Include an additional coefficient to model non-linear genetic effects ($\beta_4 G^2$), and a second to account for non-linear interaction effects ($\beta_5 G^2$ x E)	Aliev, *Bavav Genet,* 2014
Comparing biological frameworks (e.g. diathesis-stress model *vs.* differential susceptibility framework)	Adjust the parameters in the regression equation to compare alternate theoretical frameworks	Belsky, *Psychol Bull*, 2009 Widamen, *Psychol Methods*, 2012
Selection of interaction scale (e.g. additive *vs.* multiplicative)	Consider the application of the interaction test *a priori*. Additive scales have been recommended for identifying heterogeneous effects across subgroups in public health settings, while multiplicative scales are suggested for studying disease etiology	Ottman, *Prev Med*, 1996
Confounding of the G x E interaction term	Include all covariate x gene and covariate x environment interaction terms	Keller, *Biol Psychiatry*, 2014
Shared heritability between the outcome and covariates	Avoid the inclusion of heritable covariates that are associated with the gene variant being tested	Aschard, *Am J Hum Genet*, 2015
Correlation between the gene variant under study and the interacting environmental factor	Directly analyze the relationship between the interacting gene variant and environmental exposure to ensure that they are not correlated	VanderWeele, *Am J Epidemiol*, 2013
Variations in gene expression/silencing, and changing the heritability of BMI throughout development	Use a repeated measures design or include a G x E x Time term if the sample size is sufficient	Liu, *Environ Health*, 2012
Changing heritability of BMI throughout development	Use existing gene x age interactions to identify variants with differential effects across the lifespan	Elks, *Front Endocrinol*, 2012 Winkler, *PLoS Genet*, 2015
Measurement error associated with the environmental exposure and outcome	Consider more accurate measurement tools or repeated measures in favour of large sample sizes with less accurate measures	Wong, *Int J Epidemiol*, 2003

REFERENCES

1 Finucane, M. M. *et al.* National, regional, and global trends in body-mass index since 1980: systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9.1 million participants. *Lancet* **377**, 557-567, doi:10.1016/S0140-6736(10)62037-5 (2011).

2 Dixon, J. B. The effect of obesity on health outcomes. *Mol Cell Endocrinol* **316**, 104-108, doi:10.1016/j.mce.2009.07.008 (2010).

3 Fontaine, K. R., Redden, D. T., Wang, C., Westfall, A. O. & Allison, D. B. Years of life lost due to obesity. *Jama* **289**, 187-193 (2003).

4 Muller-Riemenschneider F, R. T., Berghofer A, Willich S. Health-economic burden of obesity in Europe. *Eur J Epidemiol* **23**, 499–509 (2008).

5 McAllister, E. J. *et al.* Ten putative contributors to the obesity epidemic. *Crit Rev Food Sci Nutr* **49**, 868-913, doi:10.1080/10408390903372599 (2009).

6 Traversy, G. & Chaput, J. P. Alcohol Consumption and Obesity: An Update. *Curr Obes Rep* **4**, 122-130, doi:10.1007/s13679-014-0129-4 (2015).

7 Choquet, H. & Meyre, D. Molecular basis of obesity: current status and future prospects. *Curr Genomics* **12**, 154-168, doi:10.2174/138920211795677921 (2011).

8 Pigeyre, M., Yazdi, F. T., Kaur, Y. & Meyre, D. Recent progress in genetics, epigenetics and metagenomics unveils the pathophysiology of human obesity. *Clinical science* **130**, 943-986, doi:10.1042/CS20160136 (2016).

9 Blakemore, A. I. & Froguel, P. Investigation of Mendelian forms of obesity holds out the prospect of personalized medicine. *Ann N Y Acad Sci* **1214**, 180-189, doi:10.1111/j.1749-6632.2010.05880.x (2010).

10 Hinney, A. & Hebebrand, J. Polygenic obesity in humans. *Obes Facts* **1**, 35-42, doi:10.1159/000113935 (2008).

11 Hinney, A., Vogel, C. I. & Hebebrand, J. From monogenic to polygenic obesity: recent advances. *Eur Child Adolesc Psychiatry* **19**, 297-310, doi:10.1007/s00787-010-0096-6 (2010).

12 Locke, A. E. *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature* **518**, 197-206, doi:10.1038/nature14177 (2015).

13 Speliotes, E. K. *et al.* Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet* **42**, 937-948, doi:10.1038/ng.686 (2010).

14 Day, F. R. & Loos, R. J. Developments in obesity genetics in the era of genome-wide association studies. *J Nutrigenet Nutrigenomics* **4**, 222-238, doi:10.1159/000332158 (2011).

15 Franks, P. W. & Poveda, A. Gene-lifestyle and gene-pharmacotherapy interactions in obesity and its cardiovascular consequences. *Curr Vasc Pharmacol* **9**, 401-456 (2011).

16 Andreasen, C. H. *et al.* Low physical activity accentuates the effect of the FTO rs9939609 polymorphism on body fat accumulation. *Diabetes* **57**, 95-101, doi:10.2337/db07-0910 (2008).

17 Sonestedt, E. *et al.* Fat and carbohydrate intake modify the association between genetic variation in the FTO genotype and obesity. *Am J Clin Nutr* **90**, 1418-1425, doi:10.3945/ajcn.2009.27958 (2009).
18 Kilpelainen, T. O. *et al.* Physical activity attenuates the influence of FTO variants on obesity risk: a meta-analysis of 218,166 adults and 19,268 children. *PLoS Med* **8**, e1001116, doi:10.1371/journal.pmed.1001116 (2011).
19 Corella, D. *et al.* A high intake of saturated fatty acids strengthens the association between the fat mass and obesity-associated gene and BMI. *J Nutr* **141**, 2219-2225, doi:10.3945/jn.111.143826 (2011).
20 Fontaine, K. R., Redden, D. T., Wang, C., Westfall, A. O. & Allison, D. B. Years of life lost due to obesity. *Jama* **289**, 187-193 (2003).
21 World Health Organization (WHO). Obesity: preventing and managing the global epidemic: Report of a WHO consultation. Geneva, Switzerland: World Health Organization, 2000. (Technical report series no. 894).
22 World Health Organization. Obesity and overweight: Fact sheet. In: WHO Media centre WHO. 2013. http://www.who.int/mediacentre/ factsheets/fs311/en/. Accessed June 04 2014.
23 Wang, Y. Epidemiology of childhood obesity--methodological aspects and guidelines: what is new? *International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity* **28 Suppl 3**, S21-28, doi:10.1038/sj.ijo.0802801 (2004).
24 Song, Y. M., Sung, J., Davey Smith, G. & Ebrahim, S. Body mass index and ischemic and hemorrhagic stroke: a prospective study in Korean men. *Stroke; a journal of cerebral circulation* **35**, 831-836, doi:10.1161/01.STR.0000119386.22691.1C (2004).
25 Manson, J. E. *et al.* A prospective study of obesity and risk of coronary heart disease in women. *N Engl J Med* **322**, 882-889, doi:10.1056/NEJM199003293221303 (1990).
26 Calle, E. E., Rodriguez, C., Walker-Thurmond, K. & Thun, M. J. Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. *N Engl J Med* **348**, 1625-1638, doi:10.1056/NEJMoa021423 (2003).
27 Reeves, G. K. *et al.* Cancer incidence and mortality in relation to body mass index in the Million Women Study: cohort study. *Bmj* **335**, 1134, doi:10.1136/bmj.39367.495995.AE (2007).
28 Whitlock, G. *et al.* Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies. *Lancet* **373**, 1083-1096, doi:10.1016/S0140-6736(09)60318-4 (2009).
29 WHO Global InfoBase team. The SuRF Report 2. Surveillance of chronic disease risk factors: Country-level data and comparable estimates. Geneva: World Health Organization, 2005.
30 Finkelstein, E. A. *et al.* Obesity and severe obesity forecasts through 2030. *Am J Prev Med* **42**, 563-570, doi:10.1016/j.amepre.2011.10.026 (2012).
31 Finkelstein, E. A. *et al.* The lifetime medical cost burden of overweight and obesity: implications for obesity prevention. *Obesity (Silver Spring)* **16**, 1843-1848, doi:10.1038/oby.2008.290 (2008).
32 Haslam, D. W. & James, W. P. Obesity. *Lancet* **366**, 1197-1209 (2005).

33 McAllister, E. J. *et al.* Ten putative contributors to the obesity epidemic. *Critical reviews in food science and nutrition* **49**, 868-913 (2009).
34 Levin, B. E. Factors promoting and ameliorating the development of obesity. *Physiology & behavior* **86**, 633-639 (2005).
35 Choquet, H. & Meyre, D. Molecular Basis of Obesity: Current Status and Future Prospects. *Current genomics* **vol.12 (3)**, 154-168 (2011).
36 Montague, C. T. *et al.* Congenital leptin deficiency is associated with severe early-onset obesity in humans. *Nature* **387**, 903-908. (1997).
37 Katsanis, N. *et al.* Mutations in MKKS cause obesity, retinal dystrophy and renal malformations associated with Bardet-Biedl syndrome. *Nat Genet* **26**, 67-70 (2000).
38 Frayling, T. M. *et al.* A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* **316**, 889-894 (2007).
39 Dina, C. *et al.* Variation in FTO contributes to childhood obesity and severe adult obesity. *Nat Genet* **39**, 724-726 (2007).
40 Khoury, M. J. & Wacholder, S. Invited commentary: from genome-wide association studies to gene-environment-wide interaction studies--challenges and opportunities. *American journal of epidemiology* **169**, 227-230; discussion 234-225 (2009).
41 Khoury, M. J., Beaty, T. H. & Cohen, B. H. Scope and strategies of genetic epidemiology: analysis of articles published in Genetic Epidemiology, 1984-1991. *Genet Epidemiol* **10**, 321-329, doi:10.1002/gepi.1370100505 (1993).
42 Narod, S. A. Modifiers of risk of hereditary breast and ovarian cancer. *Nat Rev Cancer* **2**, 113-123, doi:10.1038/nrc726 (2002).
43 Ichihara, S. & Yamada, Y. Genetic factors for human obesity. *Cell Mol Life Sci* **65**, 1086-1098, doi:10.1007/s00018-007-7453-8 (2008).
44 Chung, W. K. & Leibel, R. L. Molecular physiology of syndromic obesities in humans. *Trends Endocrinol Metab* **16**, 267-272, doi:10.1016/j.tem.2005.06.009 (2005).
45 Choquet, H. & Meyre, D. Molecular basis of obesity: current status and future prospects. *Current genomics* **12**, 154-168, doi:10.2174/138920211795677921 (2011).
46 Stutzmann, F. *et al.* Prevalence of melanocortin-4 receptor deficiency in Europeans and their age-dependent penetrance in multigenerational pedigrees. *Diabetes* **57**, 2511-2518, doi:10.2337/db08-0153 (2008).
47 Choquet, H. & Meyre, D. Genetics of Obesity: What have we Learned? *Current genomics* **12**, 169-179, doi:10.2174/138920211795677895 (2011).
48 Stahl, E. A. *et al.* Bayesian inference analyses of the polygenic architecture of rheumatoid arthritis. *Nat Genet* **44**, 483-489, doi:10.1038/ng.2232 (2012).
49 Haldane, J. B. The interaction of nature and nurture. *Ann Eugen* **13**, 197-205 (1946).
50 Hunter, D. J. Gene-environment interactions in human diseases. *Nat Rev Genet* **6**, 287-298 (2005).

51 Franks, P. W., Pearson, E. & Florez, J. C. Gene-environment and gene-treatment interactions in type 2 diabetes: progress, pitfalls, and prospects. *Diabetes care* **36**, 1413-1421, doi:10.2337/dc12-2211 (2013).
52 Rothman, K. J., Greenland, S. & Walker, A. M. Concepts of interaction. *American journal of epidemiology* **112**, 467-470 (1980).
53 Bayoumi, R. A. *et al.* Heritability of determinants of the metabolic syndrome among healthy Arabs of the Oman family study. *Obesity (Silver Spring)* **15**, 551-556 (2007).
54 Stunkard, A. J., Foch, T. T. & Hrubec, Z. A twin study of human obesity. *Jama* **256**, 51-54 (1986).
55 Stunkard, A. J. *et al.* An adoption study of human obesity. *N Engl J Med* **314**, 193-198 (1986).
56 Bouchard, C. *et al.* The response to long-term overfeeding in identical twins. *N Engl J Med* **322**, 1477-1482 (1990).
57 Whitaker, R. C. Predicting preschooler obesity at birth: the role of maternal obesity in early pregnancy. *Pediatrics* **114**, e29-36 (2004).
58 Murrin, C. M., Kelly, G. E., Tremblay, R. E. & Kelleher, C. C. Body mass index and height over three generations: evidence from the Lifeways cross-generational cohort study. *BMC public health* **12**, 81 (2012).
59 Lawlor, D. A., Lichtenstein, P., Fraser, A. & Langstrom, N. Does maternal weight gain in pregnancy have long-term effects on offspring adiposity? A sibling study in a prospective cohort of 146,894 men from 136,050 families. *The American journal of clinical nutrition* **94**, 142-148 (2011).
60 Pietilainen, K. H. *et al.* Distribution and heritability of BMI in Finnish adolescents aged 16y and 17y: a study of 4884 twins and 2509 singletons. *International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity* **23**, 107-115. (1999).
61 Schousboe, K. *et al.* Sex differences in heritability of BMI: a comparative study of results from twin studies in eight countries. *Twin Res* **6**, 409-421 (2003).
62 Shungin, D., Wnkler, T., Croteau-Chonka, D., Ferreira, T., Locke, A., Magi, R., Strawbridge, R., Pers, T., Fischer, K. New genetic loci link adipose and insulin biology to body fat distribution. *Nature* **518**, 187-193 (2015).
63 Dubois, L. *et al.* Genetic and environmental contributions to weight, height, and BMI from birth to 19 years of age: an international study of over 12,000 twin pairs. *PLoS One* **7**, e30153, doi:10.1371/journal.pone.0030153 (2012).
64 Demerath, E. W. *et al.* Genetic and environmental influences on infant weight and weight change: the Fels Longitudinal Study. *Am J Hum Biol* **19**, 692-702 (2007).
65 Haworth, C. M. *et al.* Increasing heritability of BMI and stronger associations with the FTO gene over childhood. *Obesity (Silver Spring)* **16**, 2663-2668 (2008).
66 Lajunen, H. R. *et al.* Genetic and environmental effects on body mass index during adolescence: a prospective study among Finnish twins. *Int J Obes (Lond)* **33**, 559-567 (2009).
67 Nan, C. *et al.* Heritability of body mass index in pre-adolescence, young adulthood and late adulthood. *European journal of epidemiology* (2012).
68 North, K. E. *et al.* Genetic Epidemiology of BMI and Body Mass Change From Adolescence to Young Adulthood. *Obesity (Silver Spring)* (2009).

69 Hjelmborg, J. B. *et al.* Genetic influences on growth traits of BMI: a longitudinal study of adult twins. *Obesity (Silver Spring)* **16**, 847-852 (2008).
70 Allison, D. B., Faith, M. S. & Nathan, J. S. Risch's lambda values for human obesity. *International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity* **20**, 990-999. (1996).
71 Rokholm, B. *et al.* Increased genetic variance of BMI with a higher prevalence of obesity. *PLoS One* **6**, e20816, doi:10.1371/journal.pone.0020816 (2011).
72 Wardle, J., Carnell, S., Haworth, C. M. & Plomin, R. Evidence for a strong genetic influence on childhood adiposity despite the force of the obesogenic environment. *The American journal of clinical nutrition* **87**, 398-404 (2008).
73 Rokholm, B. *et al.* Increasing genetic variance of body mass index during the Swedish obesity epidemic. *PLoS One* **6**, e27135, doi:10.1371/journal.pone.0027135 (2011).
74 Mustelin, L., Silventoinen, K., Pietilainen, K., Rissanen, A. & Kaprio, J. Physical activity reduces the influence of genetic effects on BMI and waist circumference: a study in young adult twins. *Int J Obes (Lond)* **33**, 29-36 (2009).
75 McCaffery, J. M., Papandonatos, G. D., Bond, D. S., Lyons, M. J. & Wing, R. R. Gene X environment interaction of vigorous exercise and body mass index among male Vietnam-era twins. *The American journal of clinical nutrition* **89**, 1011-1018 (2009).
76 Williams, P. T. Attenuating effect of vigorous physical activity on the risk for inherited obesity: a study of 47,691 runners. *PLoS One* **7**, e31436, doi:10.1371/journal.pone.0031436 (2012).
77 Roskam, A. J. *et al.* Comparative appraisal of educational inequalities in overweight and obesity among adults in 19 European countries. *Int J Epidemiol* **39**, 392-404 (2010).
78 Lajunen, H. R., Kaprio, J., Rose, R. J., Pulkkinen, L. & Silventoinen, K. Genetic and environmental influences on BMI from late childhood to adolescence are modified by parental education. *Obesity (Silver Spring)* **20**, 583-589 (2012).
79 Cappuccio, F. P. *et al.* Meta-analysis of short sleep duration and obesity in children and adults. *Sleep* **31**, 619-626 (2008).
80 Watson, N. F. *et al.* Sleep duration and body mass index in twins: a gene-environment interaction. *Sleep* **35**, 1-7 (2012).
81 Maher, A. R. *et al.* Efficacy and comparative effectiveness of atypical antipsychotic medications for off-label uses in adults: a systematic review and meta-analysis. *Jama* **306**, 1359-1369 (2011).
82 Correll, C. U. & Malhotra, A. K. Pharmacogenetics of antipsychotic-induced weight gain. *Psychopharmacology* **174**, 477-489 (2004).
83 Theisen, F. M. *et al.* Clozapine-induced weight gain: a study in monozygotic twins and same-sex sib pairs. *Psychiatric genetics* **15**, 285-289 (2005).
84 Gebhardt, S. *et al.* Body weight gain induced by atypical antipsychotics: an extension of the monozygotic twin and sib pair study. *J Clin Pharm Ther* **35**, 207-211 (2010).
85 Hainer, V. *et al.* Intrapair resemblance in very low calorie diet-induced weight loss in female obese identical twins. *International journal of obesity and related*

metabolic disorders : journal of the International Association for the Study of Obesity **24**, 1051-1057 (2000).
86 Bouchard, C. *et al.* The response to exercise with constant energy intake in identical twins. *Obes Res* **2**, 400-410 (1994).
87 Hatoum, I. J. *et al.* Heritability of the weight loss response to gastric bypass surgery. *J Clin Endocrinol Metab* **96**, E1630-1633 (2011).
88 Guo, G., Liu, H., Wang, L., Shen, H. & Hu, W. The Genome-Wide Influence on Human BMI Depends on Physical Activity, Life Course, and Historical Period. *Demography* **52**, 1651-1670, doi:10.1007/s13524-015-0421-2 (2015).
89 Vaisse, C. *et al.* Melanocortin-4 receptor mutations are a frequent and heterogeneous cause of morbid obesity. *J Clin Invest* **106**, 253-262. (2000).
90 Moore, S. J. *et al.* Clinical and genetic epidemiology of Bardet-Biedl syndrome in Newfoundland: a 22-year prospective, population-based, cohort study. *American journal of medical genetics. Part A* **132**, 352-360 (2005).
91 Dudley, O., McManus, B., Vogels, A., Whittington, J. & Muscatelli, F. Cross-cultural comparisons of obesity and growth in Prader-Willi syndrome. *Journal of intellectual disability research : JIDR* **52**, 426-436 (2008).
92 Bachmann-Gagescu, R. *et al.* Recurrent 200-kb deletions of 16p11.2 that include the SH2B1 gene are associated with developmental delay and obesity. *Genetics in medicine : official journal of the American College of Medical Genetics* **12**, 641-647 (2010).
93 Farooqi, I. S. *et al.* Clinical spectrum of obesity and mutations in the melanocortin 4 receptor gene. *N Engl J Med* **348**, 1085-1095. (2003).
94 Carmi, R., Elbedour, K., Stone, E. M. & Sheffield, V. C. Phenotypic differences among patients with Bardet-Biedl syndrome linked to three different chromosome loci. *American journal of medical genetics* **59**, 199-203 (1995).
95 Sebire, N. J. *et al.* Maternal obesity and pregnancy outcome: a study of 287,213 pregnancies in London. *International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity* **25**, 1175-1182 (2001).
96 Crozier, S. R. *et al.* Weight gain in pregnancy and childhood body composition: findings from the Southampton Women's Survey. *The American journal of clinical nutrition* **91**, 1745-1751 (2010).
97 Oken, E., Rifas-Shiman, S. L., Field, A. E., Frazier, A. L. & Gillman, M. W. Maternal gestational weight gain and offspring weight in adolescence. *Obstet Gynecol* **112**, 999-1006 (2008).
98 Reynolds, R. M., Osmond, C., Phillips, D. I. & Godfrey, K. M. Maternal BMI, Parity, and Pregnancy Weight Gain: Influences on Offspring Adiposity in Young Adulthood. *J Clin Endocrinol Metab* (2010).
99 Nizard, J., Dommergue, M. & Clement, K. Pregnancy in a woman with a leptin-receptor mutation. *N Engl J Med* **366**, 1064-1065 (2012).
100 Lawlor, D. A. *et al.* Maternal and offspring adiposity-related genetic variants and gestational weight gain. *The American journal of clinical nutrition* **94**, 149-155 (2011).
101 Lee, K. W. *et al.* Prenatal exposure to cigarette smoke interacts with to modulate dietary preference for fat. *J Psychiatry Neurosci* **39**, 130263 (2014).

102 Ogden, C. L. *et al.* Prevalence of overweight and obesity in the United States, 1999-2004. *Jama* **295**, 1549-1555 (2006).
103 Dempfle, A. *et al.* Large quantitative effect of melanocortin-4 receptor gene mutations on body mass index. *J Med Genet* **41**, 795-800 (2004).
104 Lindgren, C. M. *et al.* Genome-wide association scan meta-analysis identifies three Loci influencing adiposity and fat distribution. *PLoS Genet* **5**, e1000508 (2009).
105 Heid, I. M. *et al.* Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nat Genet* **42**, 949-960 (2010).
106 Winkler, T. W. *et al.* The Influence of Age and Sex on Genetic Associations with Adult Body Size and Shape: A Large-Scale Genome-Wide Interaction Study. *PLoS Genet* **11**, e1005378, doi:10.1371/journal.pgen.1005378 (2015).
107 Butler, J. V., Whittington, J. E., Holland, A. J., McAllister, C. J. & Goldstone, A. P. The transition between the phenotypes of Prader-Willi syndrome during infancy and early childhood. *Dev Med Child Neurol* **52**, e88-93 (2010).
108 Walters, R. G. *et al.* A new highly penetrant form of obesity due to deletions on chromosome 16p11.2. *Nature* **463**, 671-675, doi:10.1038/nature08727 (2010).
109 Yu, Y. *et al.* Age- and gender-dependent obesity in individuals with 16p11.2 deletion. *Journal of genetics and genomics = Yi chuan xue bao* **38**, 403-409 (2011).
110 Kaakinen, M. *et al.* Life-Course Analysis of a Fat Mass and Obesity-Associated (FTO) Gene Variant and Body Mass Index in the Northern Finland Birth Cohort 1966 Using Structural Equation Modeling. *American journal of epidemiology* (2010).
111 Sovio, U. *et al.* Association between Common Variation at the FTO Locus and Changes in Body Mass Index from Infancy to Late Childhood: The Complex Nature of Genetic Association through Growth and Development. *PLoS Genet* **7**, e1001307 (2011).
112 Hertel, J. K. *et al.* FTO, Type 2 Diabetes, and Weight Gain Throughout Adult Life: A Meta-Analysis of 41,504 Subjects From the Scandinavian HUNT, MDC, and MPP Studies. *Diabetes* (2011).
113 Rzehak, P. *et al.* Associations between BMI and the FTO gene are age dependent: results from the GINI and LISA birth cohort studies up to age 6 years. *Obes Facts* **3**, 173-180 (2010).
114 Qi, L. *et al.* Fat mass-and obesity-associated (FTO) gene variant is associated with obesity: longitudinal analyses in two cohort studies and functional test. *Diabetes* **57**, 3145-3151 (2008).
115 Wangensteen, T. *et al.* FTO genotype and weight gain in obese and normal weight adults from a Norwegian population based cohort (the HUNT study). *Experimental and clinical endocrinology & diabetes : official journal, German Society of Endocrinology [and] German Diabetes Association* **118**, 649-652 (2010).
116 Vimaleswaran, K. S. *et al.* Association Between FTO Variant and Change in Body Weight and Its Interaction With Dietary Factors: The DiOGenes Study. *Obesity (Silver Spring)* (2012).

117 Jacobsson, J. A. *et al.* Detailed analysis of variants in FTO in association with body composition in a cohort of 70-year-olds suggests a weakened effect among elderly. *PLoS One* **6**, e20158, doi:10.1371/journal.pone.0020158 (2011).
118 Elks, C. E. *et al.* Genetic markers of adult obesity risk are associated with greater early infancy weight gain and growth. *PLoS Med* **7**, e1000284 (2010).
119 Kilpelainen, T. O. *et al.* Obesity-susceptibility loci have a limited influence on birth weight: a meta-analysis of up to 28,219 individuals. *The American journal of clinical nutrition* **93**, 851-860 (2011).
120 Elks, C. E. *et al.* Adult obesity susceptibility variants are associated with greater childhood weight gain and a faster tempo of growth: the 1946 British Birth Cohort Study. *The American journal of clinical nutrition* (2012).
121 Kilpelainen, T. O., Bingham, S. A., Khaw, K. T., Wareham, N. J. & Loos, R. J. Association of variants in the PCSK1 gene with obesity in the EPIC-Norfolk study. *Hum Mol Genet* **18**, 3496-3501, doi:10.1093/hmg/ddp280 (2009).
122 Choquet, H., Kasberger, J., Hamidovic, A. & Jorgenson, E. Contribution of Common PCSK1 Genetic Variants to Obesity in 8,359 Subjects from Multi-Ethnic American Population. *PLoS One* **8**, e57857 (2013).
123 Nead, K. T. *et al.* Contribution of common non-synonymous variants in PCSK1 to body mass index variation and risk of obesity: a systematic review and meta-analysis with evidence from up to 331 175 individuals. *Hum Mol Genet* **24**, 3582-3594, doi:10.1093/hmg/ddv097 (2015).
124 Wangensteen, T. *et al.* Mutations in the melanocortin 4 receptor (MC4R) gene in obese patients in Norway. *Experimental and clinical endocrinology & diabetes : official journal, German Society of Endocrinology [and] German Diabetes Association* **117**, 266-273 (2009).
125 Rosenquist, J. N. *et al.* Cohort of birth modifies the association between FTO genotype and BMI. *Proc Natl Acad Sci U S A* **112**, 354-359, doi:10.1073/pnas.1411893111 (2015).
126 Taylor, A. E. *et al.* Associations of FTO and MC4R Variants with Obesity Traits in Indians and the Role of Rural/Urban Environment as a Possible Effect Modifier. *J Obes* **2011**, 307542 (2011).
127 Vasan, S. K. *et al.* Associations of Variants in FTO and Near MC4R With Obesity Traits in South Asian Indians. *Obesity (Silver Spring)* (2012).
128 Hennig, B. J. *et al.* FTO gene variation and measures of body mass in an African population. *BMC Med Genet* **10**, 21, doi:10.1186/1471-2350-10-21 (2009).
129 Demerath, E. W. *et al.* The positive association of obesity variants with adulthood adiposity strengthens over an 80-year period: a gene-by-birth year interaction. *Human Heredity* **75**, 175-185, doi:10.1159/000351742 (2013).
130 Richardson, A. S. *et al.* Moderate to vigorous physical activity interactions with genetic variants and body mass index in a large US ethnically diverse cohort. *Pediatr Obes* **9**, 35-46, doi:10.1111/j.2047-6310.2013.00152.x (2013).
131 Vimaleswaran, K. S. *et al.* Physical activity attenuates the body mass index-increasing influence of genetic variation in the FTO gene. *Am J Clin Nutr* **90**, 425-428, doi:10.3945/ajcn.2009.27652 (2009).

132 Demerath, E. W. *et al.* Interaction of FTO and physical activity level on adiposity in African-American and European-American adults: the ARIC study. *Obesity (Silver Spring)* **19**, 1866-1872, doi:10.1038/oby.2011.131 (2011).
133 Sonestedt, E. *et al.* Association between fat intake, physical activity and mortality depending on genetic variation in FTO. *Int J Obes (Lond)* **35**, 1041-1049, doi:10.1038/ijo.2010.263 (2011).
134 Ruiz, J. R. *et al.* Attenuation of the effect of the FTO rs9939609 polymorphism on total and central body fat by physical activity in adolescents: the HELENA study. *Arch Pediatr Adolesc Med* **164**, 328-333, doi:10.1001/archpediatrics.2010.29 (2010).
135 Moore, S. C. *et al.* Common genetic variants and central adiposity among Asian-Indians. *Obesity (Silver Spring)* **20**, 1902-1908, doi:10.1038/oby.2011.238 (2012).
136 Lee, H. J. *et al.* Effects of common FTO gene variants associated with BMI on dietary intake and physical activity in Koreans. *Clin Chim Acta* **411**, 1716-1722, doi:10.1016/j.cca.2010.07.010 (2010).
137 Li, S. *et al.* Physical activity attenuates the genetic predisposition to obesity in 20,000 men and women from EPIC-Norfolk prospective population study. *PLoS Med* **7**, e1000332, doi:10.1371/journal.pmed.1000332 (2010).
138 Qi, Q. *et al.* Television watching, leisure time physical activity, and the genetic predisposition in relation to body mass index in women and men. *Circulation* **126**, 1821-1827, doi:10.1161/CIRCULATIONAHA.112.098061 (2012).
139 Rampersaud, E. *et al.* Physical activity and the association of common FTO gene variants with body mass index and obesity. *Arch Intern Med* **168**, 1791-1797, doi:10.1001/archinte.168.16.1791 (2008).
140 Rankinen, T., Rice, T., Teran-Garcia, M., Rao, D. C. & Bouchard, C. FTO genotype is associated with exercise training-induced changes in body composition. *Obesity (Silver Spring)* **18**, 322-326, doi:10.1038/oby.2009.205 (2010).
141 Ahmad, T. *et al.* Lifestyle interaction with fat mass and obesity-associated (FTO) genotype and risk of obesity in apparently healthy U.S. women. *Diabetes care* **34**, 675-680, doi:10.2337/dc10-0948 (2011).
142 Xi, B. *et al.* The common rs9939609 variant of the fat mass and obesity-associated gene is associated with obesity risk in children and adolescents of Beijing, China. *BMC Med Genet* **11**, 107, doi:10.1186/1471-2350-11-107 (2010).
143 Mitchell, J. A. *et al.* FTO genotype and the weight loss benefits of moderate intensity exercise. *Obesity (Silver Spring)* **18**, 641-643, doi:10.1038/oby.2009.311 (2010).
144 Corella, D. *et al.* Statistical and biological gene-lifestyle interactions of MC4R and FTO with diet and physical activity on obesity: new effects on alcohol consumption. *PLoS One* **7**, e52344, doi:10.1371/journal.pone.0052344 (2012).
145 Scott, R. A. *et al.* FTO genotype and adiposity in children: physical activity levels influence the effect of the risk genotype in adolescent males. *European journal of human genetics : EJHG* **18**, 1339-1343, doi:10.1038/ejhg.2010.131 (2010).
146 Cauchi, S. *et al.* Combined effects of MC4R and FTO common genetic variants on obesity in European general populations. *J Mol Med (Berl)* **87**, 537-546, doi:10.1007/s00109-009-0451-6 (2009).

147 Ahmad, S. *et al.* Physical activity, smoking, and genetic predisposition to obesity in people from Pakistan: the PROMIS study. *BMC Med Genet* **16**, 114, doi:10.1186/s12881-015-0259-x (2015).
148 Reddon, H. *et al.* Physical activity and genetic predisposition to obesity in a multiethnic longitudinal study. *Sci Rep* **6**, 18672, doi:10.1038/srep18672 (2016).
149 Nakamura, S. *et al.* Gene-environment interactions in obesity: implication for future applications in preventive medicine. *J Hum Genet*, doi:10.1038/jhg.2015.148 (2015).
150 Kilpelainen, T. O. *et al.* Physical activity attenuates the influence of FTO variants on obesity risk: a meta-analysis of 218,166 adults and 19,268 children. *PLoS medicine* **8**, e1001116, doi:10.1371/journal.pmed.1001116 (2011).
151 Reddon, H. *et al.* Physical activity and genetic predisposition to obesity in a multiethnic longitudinal study. *Scientific reports* **6**, 18672, doi:10.1038/srep18672 (2016).
152 Zhu, J. *et al.* Associations of genetic risk score with obesity and related traits and the modifying effect of physical activity in a Chinese Han population. *PLoS One* **9**, e91442, doi:10.1371/journal.pone.0091442 (2014).
153 Ahmad, S. *et al.* Gene x physical activity interactions in obesity: combined analysis of 111,421 individuals of European ancestry. *PLoS Genet* **9**, e1003607, doi:10.1371/journal.pgen.1003607 (2013).
154 Graff, M. *et al.* Screen time behaviours may interact with obesity genes, independent of physical activity, to influence adolescent BMI in an ethnically diverse cohort. *Pediatr Obes* **8**, e74-79, doi:http://dx.doi.org/10.1111/j.2047-6310.2013.00195.x (2013).
155 Klimentidis, Y. C., Arora, A., Chougule, A., Zhou, J. & Raichlen, D. A. FTO association and interaction with time spent sitting. *Int J Obes (Lond)* **40**, 411-416, doi:10.1038/ijo.2015.190 (2015).
156 Rouskas, K. *et al.* Loss-of-Function Mutations in MC4R Are Very Rare in the Greek Severely Obese Adult Population. *Obesity (Silver Spring)* (2012).
157 Sonestedt, E. *et al.* Association between fat intake, physical activity and mortality depending on genetic variation in FTO. *Int J Obes (Lond)* (2011).
158 Corella, D. *et al.* A high intake of saturated fatty acids strengthens the association between the fat mass and obesity-associated gene and BMI. *The Journal of nutrition* **141**, 2219-2225, doi:10.3945/jn.111.143826 (2011).
159 Moleres, A. *et al.* Dietary fatty acid distribution modifies obesity risk linked to the rs9939609 polymorphism of the fat mass and obesity-associated gene in a Spanish case-control study of children. *The British journal of nutrition* **107**, 533-538, doi:10.1017/S0007114511003424 (2012).
160 Phillips, C. M. *et al.* High dietary saturated fat intake accentuates obesity risk associated with the fat mass and obesity-associated gene in adults. *The Journal of nutrition* **142**, 824-831, doi:10.3945/jn.111.153460 (2012).
161 Phillips, C. M. *et al.* High dietary saturated fat intake accentuates obesity risk associated with the fat mass and obesity-associated gene in adults. *J Nutr* **142**, 824-831, doi:10.3945/jn.111.153460 (2012).
162 Moleres, A. *et al.* Dietary fatty acid distribution modifies obesity risk linked to the rs9939609 polymorphism of the fat mass and obesity-associated gene in a Spanish

case-control study of children. *The British journal of nutrition* **107**, 533-538, doi:10.1017/S0007114511003424 (2012).
163 Lappalainen, T. *et al.* Association of the fat mass and obesity-associated (FTO) gene variant (rs9939609) with dietary intake in the Finnish Diabetes Prevention Study. *The British journal of nutrition* **108**, 1859-1865, doi:10.1017/S0007114511007410 (2012).
164 Qi, Q. *et al.* Fried food consumption, genetic risk, and body mass index: gene-diet interaction analysis in three US cohort studies. *Bmj* **348**, g1610, doi:10.1136/bmj.g1610 (2014).
165 Pennell, C. E. *et al.* Gene-environment interactions between the fat mass and obesity-associated (FTO) gene and nutrition modify the association between FTO and obesity in childhood and adolescence. *J Paediatr Child Health* **50**, 56 (2014).
166 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* **24**, 4728-4738, doi:10.1093/hmg/ddv186 (2015).
167 Qi, Q. *et al.* FTO genetic variants, dietary intake and body mass index: insights from 177,330 individuals. *Hum Mol Genet* **23**, 6961-6972, doi:10.1093/hmg/ddu411 (2014).
168 Corella, D. *et al.* APOA2, dietary fat, and body mass index: replication of a gene-diet interaction in 3 independent populations. *Archives of internal medicine* **169**, 1897-1906, doi:10.1001/archinternmed.2009.343 (2009).
169 Corella, D. *et al.* Association between the APOA2 promoter polymorphism and body weight in Mediterranean and Asian populations: replication of a gene-saturated fat interaction. *Int J Obes (Lond)* **35**, 666-675, doi:10.1038/ijo.2010.187 (2011).
170 Corella, D. *et al.* APOA2, dietary fat, and body mass index: replication of a gene-diet interaction in 3 independent populations. *Arch Intern Med* **169**, 1897-1906, doi:10.1001/archinternmed.2009.343 (2009).
171 Smith, C. E., Ordovas, J. M., Sanchez-Moreno, C., Lee, Y. C. & Garaulet, M. Apolipoprotein A-II polymorphism: relationships to behavioural and hormonal mediators of obesity. *Int J Obes (Lond)* **36**, 130-136, doi:10.1038/ijo.2011.24 (2012).
172 Prieur, X., Coste, H. & Rodriguez, J. C. The human apolipoprotein AV gene is regulated by peroxisome proliferator-activated receptor-alpha and contains a novel farnesoid X-activated receptor response element. *J Biol Chem* **278**, 25468-25480, doi:10.1074/jbc.M301302200 (2003).
173 Prieur, X. *et al.* Thyroid hormone regulates the hypotriglyceridemic gene APOA5. *J Biol Chem* **280**, 27533-27543, doi:10.1074/jbc.M503139200 (2005).
174 Aberle, J., Evans, D., Beil, F. U. & Seedorf, U. A polymorphism in the apolipoprotein A5 gene is associated with weight loss after short-term diet. *Clin Genet* **68**, 152-154, doi:10.1111/j.1399-0004.2005.00463.x (2005).
175 Corella, D. *et al.* APOA5 gene variation modulates the effects of dietary fat intake on body mass index and obesity risk in the Framingham Heart Study. *J Mol Med (Berl)* **85**, 119-128, doi:10.1007/s00109-006-0147-0 (2007).

176 Sanchez-Moreno, C. *et al.* APOA5 gene variation interacts with dietary fat intake to modulate obesity and circulating triglycerides in a Mediterranean population. *J Nutr* **141**, 380-385, doi:10.3945/jn.110.130344 (2011).
177 Marti, A., Corbalan, M. S., Martinez-Gonzalez, M. A., Forga, L. & Martinez, J. A. CHO intake alters obesity risk associated with Pro12Ala polymorphism of PPARgamma gene. *J Physiol Biochem* **58**, 219-220 (2002).
178 Robitaille, J., Despres, J. P., Perusse, L. & Vohl, M. C. The PPAR-gamma P12A polymorphism modulates the relationship between dietary fat intake and components of the metabolic syndrome: results from the Quebec Family Study. *Clin Genet* **63**, 109-116 (2003).
179 Papoutsakis, C. Gene-Diet Interactions and Obesity Indices. *Current Nutrition Reports* **1**, 142-152, doi:http://dx.doi.org/10.1007/s13668-012-0019-x (2012).
180 Memisoglu, A. *et al.* Interaction between a peroxisome proliferator-activated receptor gamma gene polymorphism and dietary fat intake in relation to body mass. *Hum Mol Genet* **12**, 2923-2929, doi:10.1093/hmg/ddg318 (2003).
181 Garaulet, M., Smith, C. E., Hernandez-Gonzalez, T., Lee, Y. C. & Ordovas, J. M. PPARgamma Pro12Ala interacts with fat intake for obesity and weight loss in a behavioural treatment based on the Mediterranean diet. *Mol Nutr Food Res* **55**, 1771-1779, doi:10.1002/mnfr.201100437 (2011).
182 Anderson, A. L. *et al.* Relationships of dietary patterns with body composition in older adults differ by gender and PPAR-gamma Pro12Ala genotype. *Eur J Nutr* **49**, 385-394, doi:10.1007/s00394-010-0096-9 (2010).
183 Dedoussis, G. V. *et al.* An age-dependent diet-modified effect of the PPARgamma Pro12Ala polymorphism in children. *Metabolism* **60**, 467-473, doi:10.1016/j.metabol.2010.04.007 (2011).
184 Lindi, V. I. *et al.* Association of the Pro12Ala polymorphism in the PPAR-gamma2 gene with 3-year incidence of type 2 diabetes and body weight change in the Finnish Diabetes Prevention Study. *Diabetes* **51**, 2581-2586 (2002).
185 Franks, P. W. *et al.* The Pro12Ala variant at the peroxisome proliferator-activated receptor gamma gene and change in obesity-related traits in the Diabetes Prevention Program. *Diabetologia* **50**, 2451-2460, doi:10.1007/s00125-007-0826-6 (2007).
186 Adamo, K. B. *et al.* Peroxisome proliferator-activated receptor gamma 2 and acyl-CoA synthetase 5 polymorphisms influence diet response. *Obesity (Silver Spring)* **15**, 1068-1075, doi:10.1038/oby.2007.630 (2007).
187 Nicklas, B. J. *et al.* Genetic variation in the peroxisome proliferator-activated receptor-gamma2 gene (Pro12Ala) affects metabolic responses to weight loss and subsequent weight regain. *Diabetes* **50**, 2172-2176 (2001).
188 Riedel, C. *et al.* Interactions of genetic and environmental risk factors with respect to body fat mass in children: results from the ALSPAC study. *Obesity (Silver Spring)* **21**, 1238-1242, doi:10.1002/oby.20196 (2013).
189 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* **23**, 854-862, doi:10.1038/ejhg.2014.189 (2014).

190 Singh, A. *et al.* Computing a Synthetic Chronic Psychosocial Stress Measurement in Multiple Datasets and its Application in the Replication of G x E Interactions of the EBF1 Gene. *Genet Epidemiol* **39**, 489-497, doi:10.1002/gepi.21910 (2015).
191 Corella, D. *et al.* Education modulates the association of the FTO rs9939609 polymorphism with body mass index and obesity risk in the Mediterranean population. *Nutrition, metabolism, and cardiovascular diseases : NMCD*.
192 Foraita, R. *et al.* Does the FTO gene interact with the socioeconomic status on the obesity development among young European children? Results from the IDEFICS study. *Int J Obes (Lond)* **39**, 1-6, doi:10.1038/ijo.2014.156 (2015).
193 Freathy, R. M. *et al.* Genetic variation at CHRNA5-CHRNA3-CHRNB4 interacts with smoking status to influence body mass index. *Int J Epidemiol* **40**, 1617-1628, doi:10.1093/ije/dyr077 (2011).
194 Ahmad, S. *et al.* A novel interaction between the FLJ33534 locus and smoking in obesity: a genome-wide study of 14 131 Pakistani adults. *Int J Obes (Lond)* **40**, 186-190, doi:10.1038/ijo.2015.152 (2016).
195 Edwards, T. L. *et al.* HTR1B, ADIPOR1, PPARGC1A, and CYP19A1 and obesity in a cohort of Caucasians and African Americans: an evaluation of gene-environment interactions and candidate genes. *Am J Epidemiol* **175**, 11-21, doi:10.1093/aje/kwr272 (2012).
196 Beyerlein, A., von Kries, R., Ness, A. R. & Ong, K. K. Genetic markers of obesity risk: stronger associations with body composition in overweight compared to normal-weight children. *PLoS One* **6**, e19057 (2011).
197 Mitchell, J. A., Hakonarson, H., Rebbeck, T. R. & Grant, S. F. Obesity-susceptibility loci and the tails of the pediatric BMI distribution. *Obesity (Silver Spring)* **21**, 1256-1260, doi:10.1002/oby.20319 (2013).
198 Williams, P. T. Quantile-specific penetrance of genes affecting lipoproteins, adiposity and height. *PLoS One* **7**, e28764, doi:10.1371/journal.pone.0028764 (2012).
199 Luppino, F. S. *et al.* Overweight, obesity, and depression: a systematic review and meta-analysis of longitudinal studies. *Archives of general psychiatry* **67**, 220-229 (2010).
200 Rivera, M. *et al.* Depressive disorder moderates the effect of the FTO gene on body mass index. *Mol Psychiatry* (2011).
201 Barber, T. M., McCarthy, M. I., Wass, J. A. & Franks, S. Obesity and polycystic ovary syndrome. *Clinical endocrinology* **65**, 137-145 (2006).
202 Kowalska, I. *et al.* The FTO gene modifies weight, fat mass and insulin sensitivity in women with polycystic ovary syndrome, where its role may be larger than in other phenotypes. *Diabetes & metabolism* **35**, 328-331 (2009).
203 Tan, S. *et al.* Large effects on body mass index and insulin resistance of fat mass and obesity associated gene (FTO) variants in patients with polycystic ovary syndrome (PCOS). *BMC Med Genet* **11**, 12 (2010).
204 Bonfig, W., Dokoupil, K. & Schmidt, H. A special, strict, fat-reduced, and carbohydrate-modified diet leads to marked weight reduction even in overweight adolescents with Prader-Willi syndrome (PWS). *TheScientificWorldJournal* **9**, 934-939 (2009).

205 Wigren, M. & Hansen, S. Prader-Willi syndrome: clinical picture, psychosocial support and current management. *Child: care, health and development* **29**, 449-456 (2003).
206 Vogels, A. & Fryns, J. P. Age at diagnosis, body mass index and physical morbidity in children and adults with the Prader-Willi syndrome. *Genet Couns* **15**, 397-404 (2004).
207 Silverthorn, K. H. & Hornak, J. E. Beneficial effects of exercise on aerobic capacity and body composition in adults with Prader-Willi syndrome. *Am J Ment Retard* **97**, 654-658 (1993).
208 Eiholzer, U. *et al.* Improving body composition and physical activity in Prader-Willi Syndrome. *J Pediatr* **142**, 73-78 (2003).
209 Butler, M. G., Theodoro, M. F., Bittel, D. C. & Donnelly, J. E. Energy expenditure and physical activity in Prader-Willi syndrome: comparison with obese subjects. *American journal of medical genetics. Part A* **143**, 449-459 (2007).
210 Santoro, N. *et al.* Weight loss in obese children carrying the proopiomelanocortin R236G variant. *J Endocrinol Invest* **29**, 226-230 (2006).
211 Reinehr, T. *et al.* Lifestyle intervention in obese children with variations in the melanocortin 4 receptor gene. *Obesity (Silver Spring)* **17**, 382-389 (2009).
212 Muller, T. D. *et al.* 'Fat mass and obesity associated' gene (FTO): no significant association of variant rs9939609 with weight loss in a lifestyle intervention and lipid metabolism markers in German obese children and adolescents. *BMC Med Genet* **9**, 85 (2008).
213 Franks, P. W. *et al.* Assessing gene-treatment interactions at the FTO and INSIG2 loci on obesity-related traits in the Diabetes Prevention Program. *Diabetologia* **51**, 2214-2223 (2008).
214 Lappalainen, T. J. *et al.* The common variant in the FTO gene did not modify the effect of lifestyle changes on body weight: the Finnish Diabetes Prevention Study. *Obesity (Silver Spring)* **17**, 832-836 (2009).
215 Matsuo, T., Nakata, Y., Murotake, Y., Hotta, K. & Tanaka, K. Effects of FTO Genotype on Weight Loss and Metabolic Risk Factors in Response to Calorie Restriction Among Japanese Women. *Obesity (Silver Spring)* (2011).
216 Woehning, A. *et al.* The A-allele of the common FTO gene variant rs9939609 complicates weight maintenance in severe obese patients. *Int J Obes (Lond)* (2012).
217 Grau, K. *et al.* Macronutrient-specific effect of FTO rs9939609 in response to a 10-week randomized hypo-energetic diet among obese Europeans. *Int J Obes (Lond)* **33**, 1227-1234 (2009).
218 Razquin, C. *et al.* A 3-year intervention with a Mediterranean diet modified the association between the rs9939609 gene variant in FTO and body weight changes. *Int J Obes (Lond)* **34**, 266-272 (2010).
219 Zhang, X. *et al.* FTO genotype and 2-year change in body composition and fat distribution in response to weight-loss diets: the POUNDS LOST Trial. *Diabetes* **61**, 3005-3011, doi:10.2337/db11-1799 (2012).

220 Huang, T. *et al.* FTO genotype, dietary protein, and change in appetite: the Preventing Overweight Using Novel Dietary Strategies trial. *Am J Clin Nutr* **99**, 1126-1130, doi:10.3945/ajcn.113.082164 (2014).
221 Mitchell, J. A. *et al.* FTO genotype and the weight loss benefits of moderate intensity exercise. *Obesity (Silver Spring)* **18**, 641-643.
222 Rankinen, T., Rice, T., Teran-Garcia, M., Rao, D. C. & Bouchard, C. FTO genotype is associated with exercise training-induced changes in body composition. *Obesity* **18**, 322-326, doi:10.1038/oby.2009.205 (2010).
223 Scherag, A. *et al.* Two new Loci for body-weight regulation identified in a joint analysis of genome-wide association studies for early-onset extreme obesity in French and german study groups. *PLoS Genet* **6**, e1000916 (2010).
224 Scherag, A. *et al.* SDCCAG8 obesity alleles and reduced weight loss after a lifestyle intervention in overweight children and adolescents. *Obesity (Silver Spring)* **20**, 466-470 (2012).
225 Delahanty, L. M. *et al.* Genetic predictors of weight loss and weight regain after intensive lifestyle modification, metformin treatment, or standard care in the Diabetes Prevention Program. *Diabetes Care* **35**, 363-366 (2012).
226 Tai, E. S. & Ordovas, J. M. The role of perilipin in human obesity and insulin resistance. *Curr Opin Lipidol* **18**, 152-156, doi:10.1097/MOL.0b013e328086aeab (2007).
227 Corella, D. *et al.* Obese subjects carrying the 11482G>A polymorphism at the perilipin locus are resistant to weight loss after dietary energy restriction. *The Journal of clinical endocrinology and metabolism* **90**, 5121-5126, doi:10.1210/jc.2005-0576 (2005).
228 Jang, Y. *et al.* Genetic variation at the perilipin locus is associated with changes in serum free fatty acids and abdominal fat following mild weight loss. *Int J Obes (Lond)* **30**, 1601-1608, doi:10.1038/sj.ijo.0803312 (2006).
229 Ruiz, J. R. *et al.* Preliminary findings on the role of PLIN1 polymorphisms on body composition and energy metabolism response to energy restriction in obese women. *The British journal of nutrition* **106**, 486-490, doi:10.1017/S0007114511000432 (2011).
230 Smith, C. E. *et al.* Perilipin polymorphism interacts with dietary carbohydrates to modulate anthropometric traits in hispanics of Caribbean origin. *J Nutr* **138**, 1852-1858 (2008).
231 Zelissen, P. M. *et al.* Effect of three treatment schedules of recombinant methionyl human leptin on body weight in obese adults: a randomized, placebo-controlled trial. *Diabetes, obesity & metabolism* **7**, 755-761 (2005).
232 Farooqi, I. S. *et al.* Beneficial effects of leptin on obesity, T cell hyporesponsiveness, and neuroendocrine/metabolic dysfunction of human congenital leptin deficiency. *J Clin Invest* **110**, 1093-1103 (2002).
233 Licinio, J. *et al.* Phenotypic effects of leptin replacement on morbid obesity, diabetes mellitus, hypogonadism, and behavior in leptin-deficient adults. *Proc Natl Acad Sci U S A* **101**, 4531-4536, doi:10.1073/pnas.0308767101 (2004).
234 Paz-Filho, G., Wong, M. L. & Licinio, J. Ten years of leptin replacement therapy. *Obesity reviews : an official journal of the International Association for the Study of Obesity*.

235 Heymsfield, S. B. *et al.* Recombinant leptin for weight loss in obese and lean adults: a randomized, controlled, dose-escalation trial. *Jama* **282**, 1568-1575 (1999).
236 Farooqi, I. S. *et al.* Partial leptin deficiency and human adiposity. *Nature* **414**, 34-35 (2001).
237 Baumgart, D. *et al.* G protein beta3 subunit 825T allele and enhanced coronary vasoconstriction on alpha(2)-adrenoceptor activation. *Circ Res* **85**, 965-969 (1999).
238 Hsiao, D. J., Wu, L. S., Huang, S. Y. & Lin, E. Weight loss and body fat reduction under sibutramine therapy in obesity with the C825T polymorphism in the GNB3 gene. *Pharmacogenetics and genomics* **19**, 730-733 (2009).
239 Grudell, A. B. *et al.* A controlled pharmacogenetic trial of sibutramine on weight loss and body composition in obese or overweight adults. *Gastroenterology* **135**, 1142-1154 (2008).
240 Tiwari, H. K. *et al.* Association of allelic variation in genes mediating aspects of energy homeostasis with weight gain during administration of antipsychotic drugs (CATIE Study). *Frontiers in genetics* **2** (2011).
241 Chowdhury, N. I. *et al.* Genetic association study between antipsychotic-induced weight gain and the melanocortin-4 receptor gene. *The pharmacogenomics journal* (2012).
242 Roffeei, S. N. *et al.* Association of FTO, LEPR and MTHFR gene polymorphisms with metabolic syndrome in schizophrenia patients receiving antipsychotics. *Pharmacogenomics* **15**, 477-485, doi:10.2217/pgs.13.220 (2014).
243 Czerwensky, F., Leucht, S. & Steimer, W. Association of the common MC4R rs17782313 polymorphism with antipsychotic-related weight gain. *J Clin Psychopharmacol* **33**, 74-79, doi:10.1097/JCP.0b013e31827772db (2013).
244 Czerwensky, F., Leucht, S. & Steimer, W. MC4R rs489693: a clinical risk factor for second generation antipsychotic-related weight gain? *Int J Neuropsychopharmacol* **16**, 2103-2109, doi:10.1017/S1461145713000849 (2013).
245 Sjostrom, L. *et al.* Lifestyle, diabetes, and cardiovascular risk factors 10 years after bariatric surgery. *N Engl J Med* **351**, 2683-2693 (2004).
246 Scheimann, A. O., Butler, M. G., Gourash, L., Cuffari, C. & Klish, W. Critical analysis of bariatric procedures in Prader-Willi syndrome. *Journal of pediatric gastroenterology and nutrition* **46**, 80-83 (2008).
247 Aslan, I. R. *et al.* Bariatric surgery in a patient with complete MC4R deficiency. *Int J Obes (Lond)* **35**, 457-461 (2011).
248 Stutzmann, F. *et al.* Variability of the effect of bariatric surgery according to the genotype of the MC4R. *Diabetes & metabolism* **36**, A23-A24 (2010).
249 Aslan, I. R. *et al.* Weight Loss after Roux-en-Y Gastric Bypass in Obese Patients Heterozygous for MC4R Mutations. *Obes Surg* (2011).
250 Hatoum, I. J. *et al.* Melanocortin-4 Receptor Signaling Is Required for Weight Loss after Gastric Bypass Surgery. *J Clin Endocrinol Metab* (2012).
251 Censani, M. *et al.* Weight loss after bariatric surgery in morbidly obese adolescents with MC4R mutations. *Obesity (Silver Spring)* **22**, 225-231, doi:10.1002/oby.20511 (2014).

252 Sarzynski, M. A. *et al.* Associations of markers in 11 obesity candidate genes with maximal weight loss and weight regain in the SOS bariatric surgery cases. *Int J Obes (Lond)* **35**, 676-683 (2011).
253 Bandstein, M. *et al.* The Role of FTO and Vitamin D for the Weight Loss Effect of Roux-en-Y Gastric Bypass Surgery in Obese Patients. *Obes Surg* **25**, 2071-2077, doi:10.1007/s11695-015-1644-4 (2015).
254 Liou, T. H. *et al.* ESR1, FTO, and UCP2 genes interact with bariatric surgery affecting weight loss and glycemic control in severely obese patients. *Obes Surg* **21**, 1758-1765 (2011).
255 Franks, P. W. & Ling, C. Epigenetics and obesity: the devil is in the details. *BMC Med* **8**, 88, doi:10.1186/1741-7015-8-88 (2010).
256 van Dijk, S. J., Molloy, P. L., Varinli, H., Morrison, J. L. & Muhlhausler, B. S. Epigenetics and human obesity. *Int J Obes (Lond)* **39**, 85-97, doi:10.1038/ijo.2014.34 (2015).
257 Feil R, F. M. Epigenetics and the environment: emerging patterns and implications. *Nature Reviews. Genetics.* **13**, 97-109 (2011).
258 Youngson, N. A. & Morris, M. J. What obesity research tells us about epigenetic mechanisms. *Philos Trans R Soc Lond B Biol Sci* **368**, 20110337, doi:10.1098/rstb.2011.0337 (2013).
259 Law, J. A. & Jacobsen, S. E. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet* **11**, 204-220, doi:10.1038/nrg2719 (2010).
260 Rakyan, V. K., Down, T. A., Balding, D. J. & Beck, S. Epigenome-wide association studies for common human diseases. *Nat Rev Genet* **12**, 529-541, doi:10.1038/nrg3000 (2011).
261 Dick, K. J. *et al.* DNA methylation and body-mass index: a genome-wide analysis. *Lancet* **383**, 1990-1998, doi:10.1016/S0140-6736(13)62674-4 (2014).
262 Demerath, E. W. *et al.* Epigenome-wide association study (EWAS) of BMI, BMI change and waist circumference in African American adults identifies multiple replicated loci. *Hum Mol Genet* **24**, 4464-4479, doi:10.1093/hmg/ddv161 (2015).
263 Pan, H. *et al.* HIF3A association with adiposity: the story begins before birth. *Epigenomics* **7**, 937-950, doi:10.2217/epi.15.45 (2015).
264 Ronn, T. *et al.* Impact of age, BMI and HbA1c levels on the genome-wide DNA methylation and mRNA expression patterns in human adipose tissue and identification of epigenetic biomarkers in blood. *Hum Mol Genet* **24**, 3792-3813, doi:10.1093/hmg/ddv124 (2015).
265 Fraga, M. F. *et al.* Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci U S A* **102**, 10604-10609, doi:10.1073/pnas.0500398102 (2005).
266 Ronn, T. *et al.* Age influences DNA methylation and gene expression of COX7A1 in human skeletal muscle. *Diabetologia* **51**, 1159-1168, doi:10.1007/s00125-008-1018-8 (2008).
267 Tobi, E. W. *et al.* DNA methylation signatures link prenatal famine exposure to growth and metabolism. *Nat Commun* **5**, 5592, doi:10.1038/ncomms6592 (2014).

268 Yajnik, C. S. & Deshmukh, U. S. Fetal programming: maternal nutrition and role of one-carbon metabolism. *Rev Endocr Metab Disord* **13**, 121-127, doi:10.1007/s11154-012-9214-8 (2012).
269 Yajnik, C. S. *et al.* Vitamin B12 and folate concentrations during pregnancy and insulin resistance in the offspring: the Pune Maternal Nutrition Study. *Diabetologia* **51**, 29-38, doi:10.1007/s00125-007-0793-y (2008).
270 Gueant, J. L., Namour, F., Gueant-Rodriguez, R. M. & Daval, J. L. Folate and fetal programming: a play in epigenomics? *Trends Endocrinol Metab* **24**, 279-289, doi:10.1016/j.tem.2013.01.010 (2013).
271 Ling, C. & Ronn, T. Epigenetic adaptation to regular exercise in humans. *Drug Discov Today* **19**, 1015-1018, doi:10.1016/j.drudis.2014.03.006 (2014).
272 Nitert, M. D. *et al.* Impact of an exercise intervention on DNA methylation in skeletal muscle from first-degree relatives of patients with type 2 diabetes. *Diabetes* **61**, 3322-3332, doi:10.2337/db11-1653 (2012).
273 Ronn, T. *et al.* A six months exercise intervention influences the genome-wide DNA methylation pattern in human adipose tissue. *PLoS Genet* **9**, e1003572, doi:10.1371/journal.pgen.1003572 (2013).
274 Voisin, S., Eynon, N., Yan, X. & Bishop, D. J. Exercise training and DNA methylation in humans. *Acta Physiol (Oxf)* **213**, 39-59, doi:10.1111/apha.12414 (2015).
275 Gillberg, L., Jacobsen, S. C., Ronn, T., Brons, C. & Vaag, A. PPARGC1A DNA methylation in subcutaneous adipose tissue in low birth weight subjects--impact of 5 days of high-fat overfeeding. *Metabolism* **63**, 263-271, doi:10.1016/j.metabol.2013.10.003 (2014).
276 Jacobsen, S. C. *et al.* Effects of short-term high-fat overfeeding on genome-wide DNA methylation in the skeletal muscle of healthy young men. *Diabetologia* **55**, 3341-3349, doi:10.1007/s00125-012-2717-8 (2012).
277 Jacobsen, S. C. *et al.* Young men with low birthweight exhibit decreased plasticity of genome-wide muscle DNA methylation by high-fat overfeeding. *Diabetologia* **57**, 1154-1158, doi:10.1007/s00125-014-3198-8 (2014).
278 Voisin, S. *et al.* Dietary fat quality impacts genome-wide DNA methylation patterns in a cross-sectional study of Greek preadolescents. *European journal of human genetics : EJHG* **23**, 654-662, doi:10.1038/ejhg.2014.139 (2015).
279 Barres, R. *et al.* Weight loss after gastric bypass surgery in human obesity remodels promoter methylation. *Cell Rep* **3**, 1020-1027, doi:10.1016/j.celrep.2013.03.018 (2013).
280 Benton, M. C. *et al.* An analysis of DNA methylation in human adipose tissue reveals differential modification of obesity genes before and after gastric bypass and weight loss. *Genome Biol* **16**, 8, doi:10.1186/s13059-014-0569-x (2015).
281 Huang, Y. T. *et al.* Epigenetic patterns in successful weight loss maintainers: a pilot study. *Int J Obes (Lond)* **39**, 865-868, doi:10.1038/ijo.2014.213 (2015).
282 Multhaup, M. L. *et al.* Mouse-human experimental epigenetic analysis unmasks dietary targets and genetic liability for diabetic phenotypes. *Cell Metab* **21**, 138-149, doi:10.1016/j.cmet.2014.12.014 (2015).

283 Teh, A. L. *et al.* The effect of genotype and in utero environment on interindividual variation in neonate DNA methylomes. *Genome Res* **24**, 1064-1074, doi:10.1101/gr.171439.113 (2014).
284 Grundberg, E. *et al.* Global analysis of DNA methylation variation in adipose tissue from twins reveals links to disease-associated variants in distal regulatory elements. *Am J Hum Genet* **93**, 876-890, doi:10.1016/j.ajhg.2013.10.004 (2013).
285 Bell, C. G. *et al.* Integrated genetic and epigenetic analysis identifies haplotype-specific methylation in the FTO type 2 diabetes and obesity susceptibility locus. *PLoS One* **5**, e14040, doi:10.1371/journal.pone.0014040 (2010).
286 Voisin, S. *et al.* Many obesity-associated SNPs strongly associate with DNA methylation changes at proximal promoters and enhancers. *Genome Med* **7**, 103, doi:10.1186/s13073-015-0225-4 (2015).
287 Dalgaard, K. *et al.* Trim28 Haploinsufficiency Triggers Bi-stable Epigenetic Obesity. *Cell* **164**, 353-364, doi:10.1016/j.cell.2015.12.025 (2016).
288 Blakemore, A. I. & Froguel, P. Investigation of Mendelian forms of obesity holds out the prospect of personalized medicine. *Annals of the New York Academy of Sciences* **1214**, 180-189 (2010).
289 Doyle, Y. G., Furey, A. & Flowers, J. Sick individuals and sick populations: 20 years later. *Journal of epidemiology and community health* **60**, 396-398 (2006).
290 Bloss, C. S., Jeste, D. V. & Schork, N. J. Genomics for disease treatment and prevention. *Psychiatr Clin North Am* **34**, 147-166 (2011).
291 Choquet, H. & Meyre, D. Genetics of Obesity: What have we Learned? *Curr Genomics* **12**, 169-179, doi:10.2174/138920211795677895 (2011).
292 Robinson, M. R. *et al.* Genotype-covariate interaction effects and the heritability of adult body mass index. *Nature genetics*, doi:10.1038/ng.3912 (2017).
293 Ladabaum, U., Mannalithara, A., Myer, P. A. & Singh, G. Obesity, Abdominal Obesity, Physical Activity, and Caloric Intake in U.S. Adults: 1988-2010. *Am J Med* **127**, 717-727, doi:10.1016/j.amjmed.2014.02.026 (2014).
294 Speakman, J. R. & O'Rahilly, S. Fat: an evolving issue. *Dis Model Mech* **5**, 569-573, doi:10.1242/dmm.010553 (2012).
295 Pate, R. R. *et al.* Physical activity and public health. A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. *Jama* **273**, 402-407 (1995).
296 Paffenbarger, R. S., Jr., Hyde, R. T., Wing, A. L. & Hsieh, C. C. Physical activity, all-cause mortality, and longevity of college alumni. *N Engl J Med* **314**, 605-613, doi:10.1056/NEJM198603063141003 (1986).
297 Powell, K. E., Thompson, P. D., Caspersen, C. J. & Kendrick, J. S. Physical activity and the incidence of coronary heart disease. *Annu Rev Public Health* **8**, 253-287, doi:10.1146/annurev.pu.08.050187.001345 (1987).
298 *The Practical Guide Identification, Evaluation, and Treatment of Overweight and Obesity in Adults.* Washington, DC: National Heart, Lung and Blood Institutes, National Institutes of Health; 2000.
299 *AcceleratingProgressinObesityPrevention:SolvingtheWeightofthe Nation.* Washington, DC: Institute of Medicine; 2012.

300 Owen, N., Healy, G. N., Matthews, C. E. & Dunstan, D. W. Too much sitting: the population health science of sedentary behavior. *Exerc Sport Sci Rev* **38**, 105-113, doi:10.1097/JES.0b013e3181e373a2 (2010).
301 Tremblay, M. Sedentary Behaviour Research N. *Letter to the editor: standardized use of the terms "sedentary" and "sedentary behaviours."* 37:540–2. *Appl Physiol Nutr Metab* **37**, 540-542 (2012).
302 World-Health-Organization. Global recommendations on physical activity for health. Switzerland. 2010. Available from: http://whqlibdoc.who.int/publications/2010/9789241599979_eng.pdf.
303 Physical Activity Guidelines for Americans. 2008. Available from: http://www.health.gov/paguidelines.
304 Chau, J. Y. *et al.* Daily sitting time and all-cause mortality: a meta-analysis. *PloS one* **8**, e80000, doi:10.1371/journal.pone.0080000 (2013).
305 Healy, G. N. *et al.* Television time and continuous metabolic risk in physically active adults. *Med Sci Sports Exerc* **40**, 639-645, doi:10.1249/MSS.0b013e3181607421 (2008).
306 Katzmarzyk, P. T., Church, T. S., Craig, C. L. & Bouchard, C. Sitting time and mortality from all causes, cardiovascular disease, and cancer. *Med Sci Sports Exerc* **41**, 998-1005, doi:10.1249/MSS.0b013e3181930355 (2009).
307 Tremblay, M. S., Colley, R. C., Saunders, T. J., Healy, G. N. & Owen, N. Physiological and health implications of a sedentary lifestyle. *Appl Physiol Nutr Metab* **35**, 725-740, doi:10.1139/H10-079 (2010).
308 Biddle, S. J., Gorely, T., Marshall, S. J., Murdey, I. & Cameron, N. Physical activity and sedentary behaviours in youth: issues and controversies. *J R Soc Promot Health* **124**, 29-33 (2004).
309 Ekelund, U. *et al.* TV viewing and physical activity are independently associated with metabolic risk in children: the European Youth Heart Study. *PLoS Med* **3**, e488, doi:10.1371/journal.pmed.0030488 (2006).
310 Dunstan, D. W. *et al.* Television viewing time and mortality: the Australian Diabetes, Obesity and Lifestyle Study (AusDiab). *Circulation* **121**, 384-391, doi:10.1161/CIRCULATIONAHA.109.894824 (2010).
311 Grontved, A. & Hu, F. B. Television viewing and risk of type 2 diabetes, cardiovascular disease, and all-cause mortality: a meta-analysis. *Jama* **305**, 2448-2455, doi:10.1001/jama.2011.812 (2011).
312 Thorp, A. A., Owen, N., Neuhaus, M. & Dunstan, D. W. Sedentary behaviors and subsequent health outcomes in adults a systematic review of longitudinal studies, 1996-2011. *Am J Prev Med* **41**, 207-215, doi:10.1016/j.amepre.2011.05.004 (2011).
313 Wijndaele, K. *et al.* Television viewing time independently predicts all-cause and cardiovascular mortality: the EPIC Norfolk study. *Int J Epidemiol* **40**, 150-159, doi:10.1093/ije/dyq105 (2011).
314 Brownson, R. C., Boehmer, T. K. & Luke, D. A. Declining rates of physical activity in the United States: what are the contributors? *Annu Rev Public Health* **26**, 421-443, doi:10.1146/annurev.publhealth.26.021304.144437 (2005).

315 Church, T. S. *et al.* Trends over 5 decades in U.S. occupation-related physical activity and their associations with obesity. *PLoS One* **6**, e19657, doi:10.1371/journal.pone.0019657 (2011).
316 Stamatakis, E., Ekelund, U. & Wareham, N. J. Temporal trends in physical activity in England: the Health Survey for England 1991 to 2004. *Prev Med* **45**, 416-423, doi:10.1016/j.ypmed.2006.12.014 (2007).
317 Knuth, A. G. & Hallal, P. C. Temporal trends in physical activity: a systematic review. *J Phys Act Health* **6**, 548-559 (2009).
318 Hallal, P. C. *et al.* Global physical activity levels: surveillance progress, pitfalls, and prospects. *Lancet* **380**, 247-257, doi:10.1016/S0140-6736(12)60646-1 (2012).
319 Nielsen Co. TV usage trends: Q3 and Q4 2010. http://www.nielsen.com/content/dam/corporate/us /en/reports-downloads/2011-Reports/State%20of %20the%20Media%20TV%20Q3%20Q4%202010 .pdf. Accessed March 11, 2015.
320 Lee, I. M. *et al.* Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet* **380**, 219-229, doi:10.1016/S0140-6736(12)61031-9 (2012).
321 Petersen, C. B. *et al.* Total sitting time and risk of myocardial infarction, coronary heart disease and all-cause mortality in a prospective cohort of Danish adults. *Int J Behav Nutr Phy* **11**, doi:Artn 13
Doi 10.1186/1479-5868-11-13 (2014).
322 Chau, J. Y. *et al.* Cross-sectional associations of total sitting and leisure screen time with cardiometabolic risk in adults. Results from the HUNT Study, Norway. *Journal of science and medicine in sport / Sports Medicine Australia* **17**, 78-84, doi:10.1016/j.jsams.2013.03.004 (2014).
323 Young A, W. F., Donnelly P. . Multiple novel gene-by-environment interactions modify the effect of FTO variants on body mass index. *Nat Commun* **7**, 12724 (2016).
324 Ahmad, S. *et al.* Gene x physical activity interactions in obesity: combined analysis of 111,421 individuals of European ancestry. *PLoS Genet* **9**, e1003607, doi:10.1371/journal.pgen.1003607 (2013).
325 Keating, B. J. *et al.* Concept, design and implementation of a cardiovascular gene-centric 50 k SNP array for large-scale genomic association studies. *PLoS One* **3**, e3583 (2008).
326 Gerstein, H. C., Yusuf, S., Holman, R., Bosch, J. & Pogue, J. Rationale, design and recruitment characteristics of a large, simple international trial of diabetes prevention: the DREAM trial. *Diabetologia* **47**, 1519-1527, doi:10.1007/s00125-004-1485-5 (2004).
327 Genuth, S. *et al.* Follow-up report on the diagnosis of diabetes mellitus. *Diabetes Care* **26**, 3160-3167 (2003).
328 Janssens, A. C. *et al.* The impact of genotype frequencies on the clinical validity of genomic profiling for predicting common chronic diseases. *Genet Med* **9**, 528-535, doi:10.1097GIM.0b013e31812eece0 (2007).
329 Robiou-du-Pont, S. *et al.* Contribution of 24 obesity-associated genetic variants to insulin resistance, pancreatic beta-cell function and type 2 diabetes risk in the

French population. *Int J Obes (Lond)* **37**, 980-985, doi:10.1038/ijo.2012.175 (2013).
330 Wilmot, E. G. *et al.* Sedentary time in adults and the association with diabetes, cardiovascular disease and death: systematic review and meta-analysis. *Diabetologia* **55**, 2895-2905, doi:10.1007/s00125-012-2677-z (2012).
331 Haskell, W. L. *et al.* Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. *Circulation* **116**, 1081-1093, doi:10.1161/CIRCULATIONAHA.107.185649 (2007).
332 Larsen, R. N. *et al.* Breaking up prolonged sitting reduces resting blood pressure in overweight/obese adults. *Nutr Metab Cardiovasc Dis* **24**, 976-982, doi:10.1016/j.numecd.2014.04.011 (2014).
333 Dunstan, D. W. *et al.* Breaking up prolonged sitting reduces postprandial glucose and insulin responses. *Diabetes Care* **35**, 976-983, doi:10.2337/dc11-1931 (2012).
334 Thorp, A. A., Kingwell, B. A., Owen, N. & Dunstan, D. W. Breaking up workplace sitting time with intermittent standing bouts improves fatigue and musculoskeletal discomfort in overweight/obese office workers. *Occup Environ Med* **71**, 765-771, doi:10.1136/oemed-2014-102348 (2014).
335 Young, A. I., Wauthier, F. & Donnelly, P. Multiple novel gene-by-environment interactions modify the effect of FTO variants on body mass index. *Nature communications* **7**, 12724, doi:10.1038/ncomms12724 (2016).
336 Speakman, J. R. *et al.* Set points, settling points and some alternative models: theoretical options to understand how genes and environments combine to regulate body adiposity. *Disease models & mechanisms* **4**, 733-745, doi:10.1242/dmm.008698 (2011).
337 Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & R-Development-Core-Team. nlme: Linear and Nonlinear Mixed Effects Models. *R package version 3.*, 1-105. (2012).
338 Dee, A. *et al.* The direct and indirect costs of both overweight and obesity: a systematic review. *BMC research notes* **7**, 242, doi:10.1186/1756-0500-7-242 (2014).
339 Blomain, E. S., Dirhan, D. A., Valentino, M. A., Kim, G. W. & Waldman, S. A. Mechanisms of Weight Regain following Weight Loss. *ISRN obesity* **2013**, 210524, doi:10.1155/2013/210524 (2013).
340 Glazer, G. Long-term pharmacotherapy of obesity 2000: a review of efficacy and safety. *Arch Intern Med* **161**, 1814-1824 (2001).
341 Kral, J. G., Kava, R. A., Catalano, P. M. & Moore, B. J. Severe Obesity: The Neglected Epidemic. *Obesity facts* **5**, 254-269, doi:10.1159/000338566 (2012).
342 Gortmaker, S. L. *et al.* Changing the future of obesity: science, policy, and action. *Lancet* **378**, 838-847, doi:10.1016/S0140-6736(11)60815-5 (2011).
343 Chaput, J. P., Bouchard, C. & Tremblay, A. Change in sleep duration and visceral fat accumulation over 6 years in adults. *Obesity* **22**, E9-12, doi:10.1002/oby.20701 (2014).
344 Taheri, S., Lin, L., Austin, D., Young, T. & Mignot, E. Short sleep duration is associated with reduced leptin, elevated ghrelin, and increased body mass index. *PLoS medicine* **1**, e62, doi:10.1371/journal.pmed.0010062 (2004).

345 Jean-Louis, G. *et al.* Associations between inadequate sleep and obesity in the US adult population: analysis of the national health interview survey (1977-2009). *BMC public health* **14**, 290, doi:10.1186/1471-2458-14-290 (2014).
346 Yiengprugsawan, V., Banwell, C., Seubsman, S. A. & Sleigh, A. C. Short sleep and obesity in a large national cohort of Thai adults. *BMJ Open* **2**, e000561, doi:10.1136/bmjopen-2011-000561 (2012).
347 Theorell-Haglow, J., Berglund, L., Berne, C. & Lindberg, E. Both habitual short sleepers and long sleepers are at greater risk of obesity: a population-based 10-year follow-up in women. *Sleep medicine* **15**, 1204-1211, doi:10.1016/j.sleep.2014.02.014 (2014).
348 Westerlund, A. *et al.* Habitual sleep patterns and the distribution of body mass index: cross-sectional findings among Swedish men and women. *Sleep medicine* **15**, 1196-1203, doi:10.1016/j.sleep.2014.06.012 (2014).
349 Singh, M., Drake, C. L., Roehrs, T., Hudgel, D. W. & Roth, T. The association between obesity and short sleep duration: a population-based study. *J Clin Sleep Med* **1**, 357-363 (2005).
350 Wong, R. J., Chou, C. & Ahmed, A. Long Term Trends and Racial/Ethnic Disparities in the Prevalence of Obesity. *Journal of community health*, doi:10.1007/s10900-014-9870-6 (2014).
351 Norden-Krichmar, T. M. *et al.* Correlation analysis of genetic admixture and social identification with body mass index in a Native American Community. *American journal of human biology : the official journal of the Human Biology Council* **26**, 347-360, doi:10.1002/ajhb.22521 (2014).
352 Elks, C. E. *et al.* Variability in the heritability of body mass index: a systematic review and meta-regression. *Frontiers in endocrinology* **3**, 29, doi:10.3389/fendo.2012.00029 (2012).
353 Yazdi, F. T., Clee, S. M. & Meyre, D. Obesity genetics in mouse and human: back and forth, and back again. *PeerJ* **3**, e856, doi:10.7717/peerj.856 (2015).
354 Ahmad, S., Varga, T. V. & Franks, P. W. Gene x environment interactions in obesity: the state of the evidence. *Human heredity* **75**, 106-115, doi:10.1159/000351070 (2013).
355 Watson, N. F. *et al.* Sleep duration and body mass index in twins: a gene-environment interaction. *Sleep* **35**, 597-603, doi:10.5665/sleep.1810 (2012).
356 Anand, S. S. *et al.* Glucose levels are associated with cardiovascular disease and death in an international cohort of normal glycaemic and dysglycaemic men and women: the EpiDREAM cohort study. *Eur J Prev Cardiol* **19**, 755-764, doi:10.1177/1741826711409327 (2012).
357 Salzano, F. M. & Sans, M. Interethnic admixture and the evolution of Latin American populations. *Genet Mol Biol* **37**, 151-170 (2014).
358 Wu, Y., Zhai, L. & Zhang, D. Sleep duration and obesity among adults: a meta-analysis of prospective studies. *Sleep medicine* **15**, 1456-1462, doi:10.1016/j.sleep.2014.07.018 (2014).
359 Dudbridge, F. Power and predictive accuracy of polygenic risk scores. *PLoS Genet* **9**, e1003348, doi:10.1371/journal.pgen.1003348 (2013).

360 Jean-Louis, G. *et al.* Associations between inadequate sleep and obesity in the US adult population: analysis of the national health interview survey (1977-2009). *BMC public health* **14**, 290, doi:10.1186/1471-2458-14-290 (2014).
361 Martinez, S. M. *et al.* Is it time for bed? Short sleep duration increases risk of obesity in Mexican American children. *Sleep medicine* **15**, 1484-1489, doi:10.1016/j.sleep.2014.09.009 (2014).
362 Xiao, Q., Arem, H., Moore, S. C., Hollenbeck, A. R. & Matthews, C. E. A large prospective investigation of sleep duration, weight change, and obesity in the NIH-AARP Diet and Health Study cohort. *American journal of epidemiology* **178**, 1600-1610, doi:10.1093/aje/kwt180 (2013).
363 Ford, E. S. *et al.* Sleep duration and body mass index and waist circumference among U.S. adults. *Obesity* **22**, 598-607, doi:10.1002/oby.20558 (2014).
364 Singh, M., Drake, C. L., Roehrs, T., Hudgel, D. W. & Roth, T. The association between obesity and short sleep duration: a population-based study. *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine* **1**, 357-363 (2005).
365 Cappuccio, F. P., D'Elia, L., Strazzullo, P. & Miller, M. A. Sleep duration and all-cause mortality: a systematic review and meta-analysis of prospective studies. *Sleep* **33**, 585-592 (2010).
366 Gallicchio, L. & Kalesan, B. Sleep duration and mortality: a systematic review and meta-analysis. *Journal of sleep research* **18**, 148-158, doi:10.1111/j.1365-2869.2008.00732.x (2009).
367 Spiegel, K., Leproult, R. & Van Cauter, E. Impact of sleep debt on metabolic and endocrine function. *Lancet* **354**, 1435-1439, doi:10.1016/S0140-6736(99)01376-8 (1999).
368 Reaven, G. M., Lithell, H. & Landsberg, L. Hypertension and associated metabolic abnormalities--the role of insulin resistance and the sympathoadrenal system. *The New England journal of medicine* **334**, 374-381, doi:10.1056/NEJM199602083340607 (1996).
369 St-Onge, M. P. *et al.* Short sleep duration increases energy intakes but does not change energy expenditure in normal-weight individuals. *The American journal of clinical nutrition* **94**, 410-416, doi:10.3945/ajcn.111.013904 (2011).
370 Youngstedt, S. D. & Kripke, D. F. Long sleep and mortality: rationale for sleep restriction. *Sleep medicine reviews* **8**, 159-174, doi:10.1016/j.smrv.2003.10.002 (2004).
371 Logue, E. E., Scott, E. D., Palmieri, P. A. & Dudley, P. Sleep duration, quality, or stability and obesity in an urban family medicine center. *Journal of clinical sleep medicine : JCSM* **10**, 177-182 (2014).
372 Massart, R. The genome-wide landscape of DNA methylation and hydroxymethylation in response to sleep deprivation impacts on synaptic plasticity genes. *Translational Psychiatry* **4** (2014).
373 Celis-Morales, C. *et al.* Sleep characteristics modify the association of genetic predisposition with obesity and anthropometric measurements in 119,679 UK Biobank participants. *The American journal of clinical nutrition* **105**, 980-990, doi:10.3945/ajcn.116.147231 (2017).

374 Pollack, K. M. & Cheskin, L. J. Obesity and workplace traumatic injury: does the science support the link? *Inj Prev* **13**, 297-302, doi:10.1136/ip.2006.014787 (2007).
375 Hammond, R. A. & Levine, R. The economic impact of obesity in the United States. *Diabetes, metabolic syndrome and obesity : targets and therapy* **3**, 285-295, doi:10.2147/DMSOTT.S7384 (2010).
376 Finkelstein, E. A., DiBonaventura, M., Burgess, S. M. & Hale, B. C. The costs of obesity in the workplace. *J Occup Environ Med* **52**, 971-976, doi:10.1097/JOM.0b013e3181f274d2 (2010).
377 Fock, K. M. & Khoo, J. Diet and exercise in management of obesity and overweight. *J Gastroenterol Hepatol* **28 Suppl 4**, 59-63, doi:10.1111/jgh.12407 (2013).
378 Yanovski, S. Z. & Yanovski, J. A. Long-term drug treatment for obesity: a systematic and clinical review. *JAMA* **311**, 74-86, doi:10.1001/jama.2013.281361 (2014).
379 Birch, L. L. & Ventura, A. K. Preventing childhood obesity: what works? *Int J Obes (Lond)* **33 Suppl 1**, S74-81 (2009).
380 Keith, S. W. *et al.* Putative contributors to the secular increase in obesity: exploring the roads less traveled. *Int J Obes (Lond)* **30**, 1585-1594 (2006).
381 Baum, C. L. The effects of cigarette costs on BMI and obesity. *Health Econ* **18**, 3-19, doi:10.1002/hec.1340 (2009).
382 Yang, T. C., Matthews, S. A. & Chen, V. Y. Stochastic variability in stress, sleep duration, and sleep quality across the distribution of body mass index: insights from quantile regression. *Int J Behav Med* **21**, 282-291, doi:10.1007/s12529-013-9293-2 (2014).
383 Lahti-Koski, M., Pietinen, P., Heliovaara, M. & Vartiainen, E. Associations of body mass index and obesity with physical activity, food choices, alcohol intake, and smoking in the 1982-1997 FINRISK Studies. *Am J Clin Nutr* **75**, 809-817 (2002).
384 Suter, P. M. The paradox of the alcohol-paradox - another step towards the resolution of the 'alcohol energy wastage' controversy. *Eur J Clin Invest* **30**, 749-750 (2000).
385 Yeomans, M. R. Alcohol, appetite and energy balance: is alcohol intake a risk factor for obesity? *Physiol Behav* **100**, 82-89, doi:10.1016/j.physbeh.2010.01.012 (2010).
386 Sobczyk-Kopciol, A. *et al.* Inverse association of the obesity predisposing FTO rs9939609 genotype with alcohol consumption and risk for alcohol dependence. *Addiction* **106**, 739-748, doi:10.1111/j.1360-0443.2010.03248.x (2011).
387 French, M. T., Norton, E. C., Fang, H. & Maclean, J. C. Alcohol consumption and body weight. *Health Econ* **19**, 814-832, doi:10.1002/hec.1521 (2010).
388 Yeomans, M. R. Alcohol, appetite and energy balance: is alcohol intake a risk factor for obesity? *Physiol Behav* **100**, 82-89, doi:10.1016/j.physbeh.2010.01.012 (2010).
389 Cheng, C. Y. *et al.* Admixture mapping of 15,280 African Americans identifies obesity susceptibility loci on chromosomes 5 and X. *PLoS Genet* **5**, e1000490 (2009).

390 Klimentidis, Y. C., Miller, G. F. & Shriver, M. D. The relationship between European genetic admixture and body composition among Hispanics and Native Americans. *Am J Hum Biol* **21**, 377-382 (2009).
391 Pigeyre, M., Yazdi, F. T., Kaur, Y. & Meyre, D. Recent progress in genetics, epigenetics and metagenomics unveils the pathophysiology of human obesity. *Clin Sci (Lond)* **130**, 943-986, doi:10.1042/CS20160136 (2016).
392 Reddon, H., Gueant, J. L. & Meyre, D. The importance of gene-environment interactions in human obesity. *Clin Sci (Lond)* **130**, 1571-1597, doi:10.1042/CS20160221 (2016).
393 Kalsi, G., Prescott, C. A., Kendler, K. S. & Riley, B. P. Unraveling the molecular mechanisms of alcohol dependence. *Trends Genet* **25**, 49-55, doi:10.1016/j.tig.2008.10.005 (2009).
394 Barry, D., Clarke, M. & Petry, N. M. Obesity and its relationship to addictions: is overeating a form of addictive behavior? *Am J Addict* **18**, 439-451, doi:10.3109/10550490903205579 (2009).
395 Trinko, R., Sears, R. M., Guarnieri, D. J. & DiLeone, R. J. Neural mechanisms underlying obesity and drug addiction. *Physiol Behav* **91**, 499-505, doi:10.1016/j.physbeh.2007.01.001 (2007).
396 Wang, L. *et al.* Genetic variants in the fat mass- and obesity-associated (FTO) gene are associated with alcohol dependence. *Journal of molecular neuroscience : MN* **51**, 416-424, doi:10.1007/s12031-013-0044-2 (2013).
397 Hubacek, J. A. *et al.* Fat mass and obesity-associated (fto) gene and alcohol intake. *Addiction* **107**, 1185-1186, doi:10.1111/j.1360-0443.2012.03825.x (2012).
398 Takeuchi, F. *et al.* Confirmation of ALDH2 as a Major locus of drinking behavior and of its variants regulating multiple metabolic phenotypes in a Japanese population. *Circulation journal : official journal of the Japanese Circulation Society* **75**, 911-918 (2011).
399 Wen, W. *et al.* Meta-analysis of genome-wide association studies in East Asian-ancestry populations identifies four new loci for body mass index. *Hum Mol Genet* **23**, 5492-5504, doi:10.1093/hmg/ddu248 (2014).
400 Anand, S. S. *et al.* Glucose levels are associated with cardiovascular disease and death in an international cohort of normal glycaemic and dysglycaemic men and women: the EpiDREAM cohort study. *Eur J Prev Cardiol* **19**, 755-764, doi:10.1177/1741826711409327 (2012).
401 Gerstein, H. C. *et al.* Rationale, design and recruitment characteristics of a large, simple international trial of diabetes prevention: the DREAM trial. *Diabetologia* **47**, 1519-1527, doi:10.1007/s00125-004-1485-5 (2004).
402 Robiou-du-Pont, S., Li, A., Christie, S., Sohani, Z. N. & Meyre, D. Should we have blind faith in bioinformatics software? Illustrations from the SNAP web-based tool. *PLoS One* **10**, e0118925, doi:10.1371/journal.pone.0118925 (2015).
403 Genuth, S. *et al.* Follow-up report on the diagnosis of diabetes mellitus. *Diabetes Care* **26**, 3160-3167 (2003).
404 Dudbridge, F. Power and predictive accuracy of polygenic risk scores. *PLoS Genet* **9**, e1003348, doi:10.1371/journal.pgen.1003348 (2013).
405 Robiou-du-Pont, S. *et al.* Contribution of 24 obesity-associated genetic variants to insulin resistance, pancreatic beta-cell function and type 2 diabetes risk in the

French population. *Int J Obes (Lond)* **37**, 980-985, doi:10.1038/ijo.2012.175 (2013).
406 Vanderweele, T. J., Ko, Y.-A. & Mukherjee, B. Environmental confounding in gene-environment interaction studies. *Am J Epidemiol* **178**, 144-152 (2013).
407 Liangpunsakul, S. Relationship between alcohol intake and dietary pattern: findings from NHANES III. *World Journal of Gastroenterology* **16**, 4055-4060 (2010).
408 Wang, L., Lee, I. M., Manson, J. E., Buring, J. E. & Sesso, H. D. Alcohol consumption, weight gain, and risk of becoming overweight in middle-aged and older women. *Arch Intern Med* **170**, 453-461, doi:10.1001/archinternmed.2009.527 (2010).
409 Tolstrup, J. S. *et al.* Alcohol drinking frequency in relation to subsequent changes in waist circumference. *Am J Clin Nutr* **87**, 957-963 (2008).
410 Williamson, D. F. *et al.* Alcohol and body weight in United States adults. *Am J Public Health* **77**, 1324-1330 (1987).
411 Liu, S., Serdula, M. K., Williamson, D. F., Mokdad, A. H. & Byers, T. A prospective study of alcohol intake and change in body weight among US adults. *Am J Epidemiol* **140**, 912-920 (1994).
412 Thomson, C. A. *et al.* Alcohol consumption and body weight change in postmenopausal women: results from the Women's Health Initiative. *Int J Obes (Lond)* **36**, 1158-1164, doi:10.1038/ijo.2012.84 (2012).
413 Rohrer JE, R. B., Denison A, et al. . . 2005; 6:17. Frequency of alcohol use and obesity in community medicine patients. *BMC Fam Pract* **6** (2005).
414 Suter, P. M. Is alcohol consumption a risk factor for weight gain and obesity? *Crit Rev Clin Lab Sci* **42**, 197-227, doi:10.1080/10408360590913542 (2005).
415 Ruf, T., Nagel, G., Altenburg, H. P., Miller, A. B. & Thorand, B. Food and nutrient intake, anthropometric measurements and smoking according to alcohol consumption in the EPIC Heidelberg study. *Ann Nutr Metab* **49**, 16-25, doi:10.1159/000084173 (2005).
416 Colditz, G. A. *et al.* Alcohol intake in relation to diet and obesity in women and men. *Am J Clin Nutr* **54**, 49-55 (1991).
417 Thomson, M. *et al.* Alcohol consumption and nutrient intake in middle-aged Scottish men. *Am J Clin Nutr* **47**, 139-145 (1988).
418 Flechtner-Mors, M., Biesalski, H. K., Jenkinson, C. P., Adler, G. & Ditschuneit, H. H. Effects of moderate consumption of white wine on weight loss in overweight and obese subjects. *International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity* **28**, 1420-1426, doi:10.1038/sj.ijo.0802786 (2004).
419 Lieber, C. S. Perspectives: do alcohol calories count? *Am J Clin Nutr* **54**, 976-982 (1991).
420 Buemann, B. & Astrup, A. How does the body deal with energy from alcohol? *Nutrition* **17**, 638-641 (2001).
421 Suter, P. M., Schutz, Y. & Jequier, E. The effect of ethanol on fat storage in healthy subjects. *N Engl J Med* **326**, 983-987, doi:10.1056/NEJM199204093261503 (1992).

422 Klesges, R. C., Mealer, C. Z. & Klesges, L. M. Effects of alcohol intake on resting energy expenditure in young women social drinkers. *Am J Clin Nutr* **59**, 805-809 (1994).
423 Wang, L. *et al.* Genetic variants in the fat mass- and obesity-associated (FTO) gene are associated with alcohol dependence. *J Mol Neurosci* **51**, 416-424, doi:10.1007/s12031-013-0044-2 (2013).
424 Corella, D. *et al.* Education modulates the association of the FTO rs9939609 polymorphism with body mass index and obesity risk in the Mediterranean population. *Nutr Metab Cardiovasc Dis* **22**, 651-658, doi:10.1016/j.numecd.2010.10.006 (2012).
425 Yang, J. *et al.* FTO genotype is associated with phenotypic variability of body mass index. *Nature* **490**, 267-272, doi:10.1038/nature11401 (2012).
426 Fowler, A. K. *et al.* Alcohol-induced one-carbon metabolism impairment promotes dysfunction of DNA base excision repair in adult brain. *J Biol Chem* **287**, 43533-43542, doi:10.1074/jbc.M112.401497 (2012).
427 Giovannucci, E. Alcohol, one-carbon metabolism, and colorectal cancer: recent insights from molecular studies. *J Nutr* **134**, 2475S-2481S (2004).
428 Zakhari, S. Alcohol metabolism and epigenetics changes. *Alcohol Res* **35**, 6-16 (2013).
429 Mason, J. B. Biomarkers of nutrient exposure and status in one-carbon (methyl) metabolism. *J Nutr* **133 Suppl 3**, 941S-947S (2003).
430 Kruman, II & Fowler, A. K. Impaired one carbon metabolism and DNA methylation in alcohol toxicity. *J Neurochem* **129**, 770-780, doi:10.1111/jnc.12677 (2014).
431 Philibert, R. A., Plume, J. M., Gibbons, F. X., Brody, G. H. & Beach, S. R. The impact of recent alcohol use on genome wide DNA methylation signatures. *Frontiers in Genetics* **3**, 54, doi:10.3389/fgene.2012.00054 (2012).
432 Dogan, M. V., Lei, M. K., Beach, S. R., Brody, G. H. & Philibert, R. A. Alcohol and tobacco consumption alter hypothalamic pituitary adrenal axis DNA methylation. *Psychoneuroendocrinology* **66**, 176-184, doi:10.1016/j.psyneuen.2016.01.018 (2016).
433 World Health Organization: Obesity and overweight. http://www.who.int/mediacentre/factsheets/fs311/en/. (2015).
434 Fierabracci, P., Tamberi, A. & Santini, F. in *Minimally invasive bariatric and metabolic surgery: principles and technical aspects* (eds Marcello Lucchese & Nicola Scopinaro) (Springer, 2015).
435 Wadden, T. A., Butryn, M. L. & Wilson, C. Lifestyle modification for the management of obesity. *Gastroenterology* **132**, 2226-2238, doi:10.1053/j.gastro.2007.03.051 (2007).
436 Padwal, R. S. & Majumdar, S. R. Drug treatments for obesity: orlistat, sibutramine, and rimonabant. *Lancet* **369**, 71-77, doi:10.1016/S0140-6736(07)60033-6 (2007).
437 Bjoœrntorp, P. (John Wiley & Sons, 2001).
438 Obesity: preventing and managing the global epidemic. (2000).

439 Lipek, T., Igel, U., Gausche, R., Kiess, W. & Grande, G. Obesogenic environments: environmental approaches to obesity prevention. *J Pediatr Endocrinol Metab* **28**, 485-495, doi:10.1515/jpem-2015-0127 (2015).
440 Gangwisch, J. E., Malaspina, D., Boden-Albala, B. & Heymsfield, S. B. Inadequate sleep as a risk factor for obesity: analyses of the NHANES I. *Sleep* **28**, 1289-1296 (2005).
441 Dallman, M. F. Stress-induced obesity and the emotional nervous system. *Trends Endocrinol Metab* **21**, 159-165, doi:10.1016/j.tem.2009.10.004 (2010).
442 Fernández, J. R., Pearson, K. E., Kell, K. P. & Bohan Brown, M. M. Genetic admixture and obesity: recent perspectives and future applications. *Hum Hered* **75**, 98-105, doi:10.1159/000353180 (2013).
443 Li, A. & Meyre, D. Jumping on the Train of Personalized Medicine: A Primer for Non-Geneticist Clinicians: Part 2. Fundamental Concepts in Genetic Epidemiology. *Curr Psychiatry Rev* **10**, 101-117, doi:10.2174/1573400510666140319235334 (2014).
444 Choquet, H. & Meyre, D. Genomic insights into early-onset obesity. *Genome Med* **2**, 36, doi:10.1186/gm157 (2010).
445 Kaur, Y., de Souza, R. J., Gibson, W. T. & Meyre, D. A systematic review of genetic syndromes with obesity. *Obesity reviews : an official journal of the International Association for the Study of Obesity* **18**, 603-634, doi:10.1111/obr.12531 (2017).
446 Locke, A. E. *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature* **518**, 197-206, doi:10.1038/nature14177 (2015).
447 O'Rahilly, S. & Farooqi, I. S. Human obesity: a heritable neurobehavioral disorder that is highly sensitive to environmental conditions. *Diabetes* **57**, 2905-2910, doi:10.2337/db08-0210 (2008).
448 Hinney, A., Vogel, C. I. & Hebebrand, J. From monogenic to polygenic obesity: recent advances. *Eur Child Adolesc Psychiatry* **19**, 297-310, doi:10.1007/s00787-010-0096-6 (2010).
449 Mustelin, L., Silventoinen, K., Pietiläinen, K., Rissanen, A. & Kaprio, J. Physical activity reduces the influence of genetic effects on BMI and waist circumference: a study in young adult twins. *Int J Obes (Lond)* **33**, 29-36, doi:10.1038/ijo.2008.258 (2009).
450 Reddon, H., Gueant, J. L. & Meyre, D. The importance of gene-environment interactions in human obesity. *Clinical science* **130**, 1571-1597, doi:10.1042/CS20160221 (2016).
451 Okorodudu, D. O. *et al.* Diagnostic performance of body mass index to identify obesity as defined by body adiposity: a systematic review and meta-analysis. *Int J Obes (Lond)* **34**, 791-799, doi:10.1038/ijo.2010.5 (2010).
452 Shephard, R. J. Limits to the measurement of habitual physical activity by questionnaires. *Br J Sports Med* **37**, 197-206; discussion 206 (2003).
453 Qi, Q. *et al.* Sugar-sweetened beverages and genetic risk of obesity. *The New England journal of medicine* **367**, 1387-1396, doi:10.1056/NEJMoa1203039 (2012).
454 O'brien, R. A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality and Quantity* **41**, 673-690 (2007).

455 Yang, J. *et al.* FTO genotype is associated with phenotypic variability of body mass index. *Nature* **490**, 267-272, doi:10.1038/nature11401 (2012).
456 Bell, C. G. *et al.* Integrated genetic and epigenetic analysis identifies haplotype-specific methylation in the FTO type 2 diabetes and obesity susceptibility locus. *PloS one* **5**, e14040, doi:10.1371/journal.pone.0014040 (2010).
457 Ronn, T. *et al.* A six months exercise intervention influences the genome-wide DNA methylation pattern in human adipose tissue. *PLoS genetics* **9**, e1003572, doi:10.1371/journal.pgen.1003572 (2013).
458 Sonestedt, E. *et al.* Fat and carbohydrate intake modify the association between genetic variation in the FTO genotype and obesity. *The American journal of clinical nutrition* **90**, 1418-1425, doi:10.3945/ajcn.2009.27958 (2009).
459 Claussnitzer, M. *et al.* FTO Obesity Variant Circuitry and Adipocyte Browning in Humans. *The New England journal of medicine* **373**, 895-907, doi:10.1056/NEJMoa1502214 (2015).
460 Stanford, K. I., Middelbeek, R. J. & Goodyear, L. J. Exercise Effects on White Adipose Tissue: Beiging and Metabolic Adaptations. *Diabetes* **64**, 2361-2368, doi:10.2337/db15-0227 (2015).
461 Wu, L. *et al.* Influence of lifestyle on the FAIM2 promoter methylation between obese and lean children: a cohort study. *BMJ open* **5**, e007670, doi:10.1136/bmjopen-2015-007670 (2015).
462 Khera, A. V. *et al.* Genetic Risk, Adherence to a Healthy Lifestyle, and Coronary Disease. *The New England journal of medicine* **375**, 2349-2358, doi:10.1056/NEJMoa1605086 (2016).
463 White, P. D. Genes, the heart and destiny. *The New England journal of medicine* **256**, 965-969, doi:10.1056/NEJM195705232562101 (1957).
464 Nielsen, D. E. & El-Sohemy, A. Disclosure of genetic information and change in dietary intake: a randomized controlled trial. *PloS one* **9**, e112665, doi:10.1371/journal.pone.0112665 (2014).
465 Franks, P. W. Gene x environment interactions in type 2 diabetes. *Curr Diab Rep* **11**, 552-561, doi:10.1007/s11892-011-0224-9 (2011).
466 Cauchi, S. *et al.* Combined effects of MC4R and FTO common genetic variants on obesity in European general populations. *J Mol Med (Berl)* **87**, 537-546, doi:10.1007/s00109-009-0451-6 (2009).
467 Keller, M. C. Gene x environment interaction studies have not properly controlled for potential confounders: the problem and the (simple) solution. *Biol Psychiatry* **75**, 18-24 (2014).
468 Aliev, F., Latendresse, S. J., Bacanu, S.-A., Neale, M. C. & Dick, D. M. Testing for measured gene-environment interaction: problems with the use of cross-product terms and a regression model reparameterization solution. *Behav Genet* **44**, 165-181 (2014).
469 Aschard, H. A perspective on interaction effects in genetic association studies. *Genetic epidemiology* **40**, 678-688, doi:10.1002/gepi.21989 (2016).
470 Underwood, P. C. *et al.* Replication and meta-analysis of the gene-environment interaction between body mass index and the interleukin-6 promoter polymorphism with higher insulin resistance. *Metabolism: Clinical & Experimental* **61**, 667-671 (2012).

471 Munafo, M. R. & Flint, J. Replication and heterogeneity in gene x environment interaction studies. *Int J Neuropsychopharmacol* **12**, 727-729, doi:10.1017/S1461145709000479 (2009).
472 Duncan, L. E. & Keller, M. C. A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. *Am J Psychiatry* **168**, 1041-1049, doi:10.1176/appi.ajp.2011.11020191 (2011).
473 Franks, P. W. & Nettleton, J. A. Invited commentary: Gene X lifestyle interactions and complex disease traits--inferring cause and effect from observational data, sine qua non. *Am J Epidemiol* **172**, 992-997; discussion 998-999, doi:10.1093/aje/kwq280 (2010).
474 Fleming, J. L., Huang, T. H. & Toland, A. E. The role of parental and grandparental epigenetic alterations in familial cancer risk. *Cancer Res* **68**, 9116-9121, doi:10.1158/0008-5472.CAN-08-2184 (2008).
475 Liu, C. Y., Maity, A., Lin, X., Wright, R. O. & Christiani, D. C. Design and analysis issues in gene and environment studies. *Environ Health* **11**, 93, doi:10.1186/1476-069X-11-93 (2012).
476 Wong, M. Y., Day, N. E., Luan, J. A., Chan, K. P. & Wareham, N. J. The detection of gene-environment interaction for continuous traits: should we deal with measurement error by bigger studies or better measurement? *Int J Epidemiol* **32**, 51-57 (2003).
477 Hewitt, J. K. Editorial policy on candidate gene association and candidate gene-by-environment interaction studies of complex traits. *Behav Genet* **42**, 1-2, doi:10.1007/s10519-011-9504-z (2012).
478 Johnston C, L. B., Matthys W. Editorial policy for candidate gene studies. *J Abnorm Child Psychol* **41**, 511-514 (2013).
479 Pinelli, M., Scala, G., Amato, R., Cocozza, S. & Miele, G. Simulating gene-gene and gene-environment interactions in complex diseases: Gene-Environment iNteraction Simulator 2. *BMC Bioinformatics* **13**, 132 (2012).
480 Van Hulle, C. A., Lahey, B. B. & Rathouz, P. J. Operating characteristics of alternative statistical methods for detecting gene-by-measured environment interaction in the presence of gene-environment correlation in twin and sibling studies. *Behav Genet* **43**, 71-84 (2013).
481 Dick, D. M. *et al.* Candidate gene-environment interaction research: reflections and recommendations. *Perspect Psychol Sci* **10**, 37-59, doi:10.1177/1745691614556682 (2015).
482 Belsky, J., Pluess, M. & Widaman, K. F. Confirmatory and competitive evaluation of alternative gene-environment interaction hypotheses. *J Child Psychol Psychiatry* **54**, 1135-1143, doi:10.1111/jcpp.12075 (2013).
483 Zubin, J. & Spring, B. Vulnerability--a new view of schizophrenia. *J Abnorm Psychol* **86**, 103-126 (1977).
484 Burmeister, M., McInnis, M. G. & Zollner, S. Psychiatric genetics: progress amid controversy. *Nat Rev Genet* **9**, 527-540, doi:10.1038/nrg2381 (2008).
485 Belsky, J. & Pluess, M. Beyond diathesis stress: differential susceptibility to environmental influences. *Psychol Bull* **135**, 885-908, doi:10.1037/a0017376 (2009).

486 Belsky, J., Bakermans-Kranenburg, M.J., & van IJzendoorn, M.H. For better and for worse: Differential Susceptibility to environmental influences. *Current Directions in Psychological Science* **16**, 300-304 (2007).

487 Belsky, J. *et al.* Vulnerability genes or plasticity genes? *Mol Psychiatry* **14**, 746-754, doi:10.1038/mp.2009.44 (2009).

488 Ellis, B. J., Boyce, W. T., Belsky, J., Bakermans-Kranenburg, M. J. & van Ijzendoorn, M. H. Differential susceptibility to the environment: an evolutionary--neurodevelopmental theory. *Dev Psychopathol* **23**, 7-28, doi:10.1017/S0954579410000611 (2011).

489 Bakermans-Kranenburg, M. J. & van Ijzendoorn, M. H. Gene-environment interaction of the dopamine D4 receptor (DRD4) and observed maternal insensitivity predicting externalizing behavior in preschoolers. *Dev Psychobiol* **48**, 406-409, doi:10.1002/dev.20152 (2006).

490 Belsky, J. & Pluess, M. Beyond risk, resilience, and dysregulation: phenotypic plasticity and human development. *Dev Psychopathol* **25**, 1243-1261, doi:10.1017/S095457941300059X (2013).

491 Roisman, G. I. *et al.* Distinguishing differential susceptibility from diathesis-stress: recommendations for evaluating interaction effects. *Dev Psychopathol* **24**, 389-409, doi:10.1017/S0954579412000065 (2012).

492 Widaman, K. F. *et al.* Distinguishing ordinal and disordinal interactions. *Psychol Methods* **17**, 615-622, doi:10.1037/a0030003 (2012).

493 Bakermans-Kranenburg, M. J. & van Ijzendoorn, M. H. Research Review: genetic vulnerability or differential susceptibility in child development: the case of attachment. *J Child Psychol Psychiatry* **48**, 1160-1173, doi:10.1111/j.1469-7610.2007.01801.x (2007).

494 Brown, G. L., Mangelsdorf, S. C. & Neff, C. Father involvement, paternal sensitivity, and father-child attachment security in the first 3 years. *J Fam Psychol* **26**, 421-430, doi:10.1037/a0027836 (2012).

495 Blot, W. J. & Day, N. E. Synergism and interaction: are they equivalent? *Am J Epidemiol* **110**, 99-100 (1979).

496 Saracci, R. Interaction and synergism. *Am J Epidemiol* **112**, 465-466 (1980).

497 VanderWeele, T. J. & Robins, J. M. The identification of synergism in the sufficient-component-cause framework. *Epidemiology* **18**, 329-339, doi:10.1097/01.ede.0000260218.66432.88 (2007).

498 Vandenbroucke, J. P. New public health and old rhetoric. *Bmj* **308**, 994-995 (1994).

499 Ottman, R. Gene-environment interaction: definitions and study designs. *Prev Med* **25**, 764-770 (1996).

500 Koopman, J. S. Causal models and sources of interaction. *Am J Epidemiol* **106**, 439-444 (1977).

501 Kupper, L. L. & Hogan, M. D. Interaction in epidemiologic studies. *Am J Epidemiol* **108**, 447-453 (1978).

502 Walter, S. D. & Holford, T. R. Additive, multiplicative, and other models for disease risks. *Am J Epidemiol* **108**, 341-346 (1978).

503 Rothman, K. J., Greenland, S. & Walker, A. M. Concepts of interaction. *Am J Epidemiol* **112**, 467-470 (1980).

504 Siemiatycki, J. & Thomas, D. C. Biological models and statistical interactions: an example from multistage carcinogenesis. *Int J Epidemiol* **10**, 383-387 (1981).
505 Eaves, L. J. Genotype x Environment interaction in psychopathology: fact or artifact? *Twin Res Hum Genet* **9**, 1-8, doi:10.1375/183242706776403073 (2006).
506 Aschard, H., Vilhjalmsson, B. J., Joshi, A. D., Price, A. L. & Kraft, P. Adjusting for heritable covariates can bias effect estimates in genome-wide association studies. *Am J Hum Genet* **96**, 329-339, doi:10.1016/j.ajhg.2014.12.021 (2015).
507 Dudbridge, F. & Fletcher, O. Gene-environment dependence creates spurious gene-environment interaction. *Am J Hum Genet* **95**, 301-307, doi:10.1016/j.ajhg.2014.07.014 (2014).
508 Vanderweele, T. J., Ko, Y. A. & Mukherjee, B. Environmental confounding in gene-environment interaction studies. *Am J Epidemiol* **178**, 144-152, doi:10.1093/aje/kws439 (2013).
509 Amos, C. I. *et al.* Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat Genet* **40**, 616-622, doi:10.1038/ng.109 (2008).
510 Hung, R. J. *et al.* A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature* **452**, 633-637, doi:10.1038/nature06885 (2008).
511 Thorgeirsson, T. E. *et al.* A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature* **452**, 638-642, doi:10.1038/nature06846 (2008).
512 Nederhof, E. *et al.* Evidence for plasticity genotypes in a gene-gene-environment interaction: the TRAILS study. *Genes, Brain, & Behavior* **9**, 968-973 (2010).
513 Woodruff, T. J. Bridging epidemiology and model organisms to increase understanding of endocrine disrupting chemicals and human health effects. *J Steroid Biochem Mol Biol* **127**, 108-117, doi:10.1016/j.jsbmb.2010.11.007 (2011).
514 Buck Louis, G. M. *et al.* Designing prospective cohort studies for assessing reproductive and developmental toxicity during sensitive windows of human reproduction and development--the LIFE Study. *Paediatr Perinat Epidemiol* **25**, 413-424, doi:10.1111/j.1365-3016.2011.01205.x (2011).
515 Ege, M. J. & Strachan, D. P. Comparisons of power of statistical methods for gene-environment interaction analyses. *Eur J Epidemiol* **28**, 785-797 (2013).
516 Heo, M. & Leon, A. C. Sample sizes required to detect two-way and three-way interactions involving slope differences in mixed-effects linear models. *J Biopharm Stat* **20**, 787-802, doi:10.1080/10543401003618819 (2010).
517 Elks, C. E. *et al.* Variability in the heritability of body mass index: a systematic review and meta-regression. *Front Endocrinol (Lausanne)* **3**, 29, doi:10.3389/fendo.2012.00029 (2012).
518 Joseph, P. G., Pare, G. & Anand, S. S. Exploring gene-environment relationships in cardiovascular disease. *The Canadian journal of cardiology* **29**, 37-45, doi:10.1016/j.cjca.2012.10.009 (2013).
519 Luan, J. A., Wong, M. Y., Day, N. E. & Wareham, N. J. Sample size determination for studies of gene-environment interaction. *Int J Epidemiol* **30**, 1035-1040 (2001).

520 Armstrong BK, W. E., Saracci R. . *Principles of Exposure Measurement in Epidemiology.* . (Oxford University Press, 1994).
521 Carroll RJ, Crainiceanu C, Ruppert D, Stefanski LA: Measurement Error in Nonlinear Models: A Modern Perspective. Chapman and Hall; 2006.
522 Muller, M. J., Bosy-Westphal, A. & Krawczak, M. Genetic studies of common types of obesity: a critique of the current use of phenotypes. *Obes Rev* **11**, 612-618, doi:10.1111/j.1467-789X.2010.00734.x (2010).
523 Pastinen, T. *et al.* A system for specific, high-throughput genotyping by allele-specific primer extension on microarrays. *Genome Res* **10**, 1031-1042 (2000).
524 Prince, J. A. *et al.* Robust and accurate single nucleotide polymorphism genotyping by dynamic allele-specific hybridization (DASH): design criteria and assay validation. *Genome Res* **11**, 152-162 (2001).
525 Peters, T. *et al.* Validity of a short questionnaire to assess physical activity in 10 European countries. *Eur J Epidemiol* **27**, 15-25, doi:Doi 10.1007/S10654-011-9625-Y (2012).
526 Welk, G. J. *et al.* Validity of 24-h Physical Activity Recall: Physical Activity Measurement Survey. *Med Sci Sports Exerc* **46**, 2014-2024, doi:10.1249/MSS.0000000000000314 (2014).
527 Richardson, M. T., Ainsworth, B. E., Wu, H. C., Jacobs, D. R., Jr. & Leon, A. S. Ability of the Atherosclerosis Risk in Communities (ARIC)/Baecke Questionnaire to assess leisure-time physical activity. *Int J Epidemiol* **24**, 685-693 (1995).
528 Wareham, N. J. *et al.* Validity and repeatability of the EPIC-Norfolk Physical Activity Questionnaire. *Int J Epidemiol* **31**, 168-174 (2002).
529 Vinknes, K. J. *et al.* Evaluation of the body adiposity index in a Caucasian population: the Hordaland health study. *Am J Epidemiol* **177**, 586-592, doi:10.1093/aje/kws271 (2013).
530 Wang, K. *et al.* A genome-wide association study on obesity and obesity-related traits. *PLoS One* **6**, e18939 (2011).
531 Wen, W. *et al.* Meta-analysis identifies common variants associated with body mass index in east Asians. *Nat Genet* **44**, 307-311, doi:10.1038/ng.1087 (2012).
532 Franks, P. W., Mesa, J. L., Harding, A. H. & Wareham, N. J. Gene-lifestyle interaction on risk of type 2 diabetes. *Nutr Metab Cardiovasc Dis* **17**, 104-124, doi:10.1016/j.numecd.2006.04.001 (2007).
533 Livingstone, M. B. *et al.* Simultaneous measurement of free-living energy expenditure by the doubly labeled water method and heart-rate monitoring. *Am J Clin Nutr* **52**, 59-65 (1990).
534 Wareham, N. J., Wong, M. Y. & Day, N. E. Glucose intolerance and physical inactivity: the relative importance of low habitual energy expenditure and cardiorespiratory fitness. *Am J Epidemiol* **152**, 132-139 (2000).
535 Dhurandhar, N. V. *et al.* Energy balance measurement: when something is not better than nothing. *Int J Obes (Lond)* **39**, 1109-1113, doi:10.1038/ijo.2014.199 (2015).
536 Delude, C. M. Deep phenotyping: The details of disease. *Nature* **527**, S14-15, doi:10.1038/527S14a (2015).

537 Hodson, L., Skeaff, C. M. & Fielding, B. A. Fatty acid composition of adipose tissue and blood in humans and its use as a biomarker of dietary intake. *Prog Lipid Res* **47**, 348-380, doi:10.1016/j.plipres.2008.03.003 (2008).
538 Cameron, J. D. *et al.* Energy depletion by diet or aerobic exercise alone: impact of energy deficit modality on appetite parameters. *Am J Clin Nutr* **103**, 1008-1016, doi:10.3945/ajcn.115.115584 (2016).
539 Ottman, R. An epidemiologic approach to gene-environment interaction. *Genet Epidemiol* **7**, 177-185, doi:10.1002/gepi.1370070302 (1990).
540 Zondervan, K. T. & Cardon, L. R. The complex interplay among factors that influence allelic association. *Nat Rev Genet* **5**, 89-100, doi:10.1038/nrg1270 (2004).
541 Hein, R., Beckmann, L. & Chang-Claude, J. Sample size requirements for indirect association studies of gene-environment interactions (G x E). *Genet Epidemiol* **32**, 235-245, doi:10.1002/gepi.20298 (2008).
542 Williamson, E., Ponsonby, A. L., Carlin, J. & Dwyer, T. Effect of including environmental data in investigations of gene-disease associations in the presence of qualitative interactions. *Genet Epidemiol* **34**, 552-560 (2010).
543 Khoury, M. J., Adams, M. J., Jr. & Flanders, W. D. An epidemiologic approach to ecogenetics. *Am J Hum Genet* **42**, 89-95 (1988).
544 Kraft, P., Yen, Y. C., Stram, D. O., Morrison, J. & Gauderman, W. J. Exploiting gene-environment interaction to detect genetic associations. *Hum Hered* **63**, 111-119 (2007).
545 Dai, J. Y. *et al.* Simultaneously testing for marginal genetic association and gene-environment interaction. *Am J Epidemiol* **176**, 164-173, doi:10.1093/aje/kwr521 (2012).
546 Manning, A. K. *et al.* Meta-analysis of gene-environment interaction: joint estimation of SNP and SNP x environment regression coefficients. *Genet Epidemiol* **35**, 11-18, doi:10.1002/gepi.20546 (2011).
547 Murcray, C. E., Lewinger, J. P. & Gauderman, W. J. Gene-environment interaction in genome-wide association studies. *American journal of epidemiology* **169**, 219-226 (2009).
548 Aschard, H., Hancock, D. B., London, S. J. & Kraft, P. Genome-wide meta-analysis of joint tests for genetic and gene-environment interaction effects. *Hum Hered* **70**, 292-300 (2011).
549 Shungin, D. *et al.* New genetic loci link adipose and insulin biology to body fat distribution. *Nature* **518**, 187-196, doi:10.1038/nature14132 (2015).
550 Marigorta, U. M. & Gibson, G. A simulation study of gene-by-environment interactions in GWAS implies ample hidden effects. *Front Genet* **5**, 225, doi:10.3389/fgene.2014.00225 (2014).
551 Pare, G., Cook, N. R., Ridker, P. M. & Chasman, D. I. On the use of variance per genotype as a tool to identify quantitative trait interaction effects: a report from the Women's Genome Health Study. *PLoS Genet* **6**, e1000981, doi:10.1371/journal.pgen.1000981 (2010).
552 Yu, K. *et al.* A flexible Bayesian model for studying gene-environment interaction. *PLoS Genet* **8**, e1002482, doi:10.1371/journal.pgen.1002482 (2012).

553 Gunther, F., Pigeot, I. & Bammann, K. Artificial neural networks modeling gene-environment interaction. *BMC Genetics* **13**, 37, doi:10.1186/1471-2156-13-37 (2012).
554 Hsu, L. *et al.* Powerful cocktail methods for detecting genome-wide gene-environment interaction. *Genet Epidemiol* **36**, 183-194, doi:10.1002/gepi.21610 (2012).
555 Arcos-Burgos, M. & Muenke, M. Genetics of population isolates. *Clin Genet* **61**, 233-247 (2002).
556 Cedernaes, J. *et al.* Acute Sleep Loss Induces Tissue-Specific Epigenetic and Transcriptional Alterations to Circadian Clock Genes in Men. *The Journal of clinical endocrinology and metabolism* **100**, E1255-1261, doi:10.1210/JC.2015-2284 (2015).
557 Sonestedt, E. *et al.* Genetic variation in the glucose-dependent insulinotropic polypeptide receptor modifies the association between carbohydrate and fat intake and risk of type 2 diabetes in the Malmo Diet and Cancer cohort. *The Journal of clinical endocrinology and metabolism* **97**, E810-818, doi:10.1210/jc.2011-2444 (2012).
558 Sander, J. D. & Joung, J. K. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol* **32**, 347-355, doi:10.1038/nbt.2842 (2014).
559 Qin, W. *et al.* Generating Mouse Models Using CRISPR-Cas9-Mediated Genome Editing. *Current protocols in mouse biology* **6**, 39-66, doi:10.1002/9780470942390.mo150178 (2016).