Genetic studies of salivary cortisol profiles and their influence on chronic disease risk factors

by

Erin K. Payne

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Epidemiological Science)
in the University of Michigan
2013

Doctoral Committee:

Professor Sharon L.R. Kardia, Co-Chair
Professor Ana V. Diez-Roux, Co-Chair
Associate Professor Bhramar Mukherjee
Professor Patricia A. Peyser
Assistant Research Scientist Jennifer A. Smith

Acknowledgements

This dissertation would not have been possible without with generous support, guidance, and counsel from my advisor, Sharon Kardia, and the other members of my dissertation committee, Ana Diez-Roux, Bhramar Mukherjee, Pat Peyser, and Jennifer Smith. I would like to thank the past and current members of the Kardia Research Lab. Tracy for her assistance when I was an analysis novice, Koji for teaching me to navigate the servers, Wei and Josh for their programming prowess, and Linda for keeping us organized and her reminders to keep moving forward.

A few special recognitions are needed. Thank you to Jen for extending her time, knowledge, mentorship, support, and a shoulder to lean on. To Erin, Alicia, and Stephanie, thank you for being the team that saw me through the highs and the lows. I will be forever grateful. To Caitlin, Colleen, and Casey, for answering the endless calls on how to finish the dissertation and start saving the world. To Steve, for his unfaltering love, enduring patience, and unending faith. To my brothers, for their constant inspiration. Lastly, to my parents, for their infinite wisdom, love, support, encouragement, and limitless belief in me. I love you all.

Table of Contents

Acknowledgements ii
List of Tables vi
List of Figures ix
Abstract xiii
CHAPTER 1: INTRODUCTION 1
Introduction 1
Psychosocial stress 3
Cortisol 6
Epidemiologic studies of cortisol 7
Stress response genes 8
Biological mechanisms linking stressors and chronic disease risk factors 11
CHAPTER 2: STUDY POPULATION AND VARIABLE MEASUREMENTS 13
Study Population 13
The MESA Study 13
The MESA Stress Study 15
Cortisol Samples and Cortisol Features 17
Daily Salivary Cortisol Samples 17
Cortisol Features 19
Genetic Data 35
Chronic Disease Risk Factors 40
Body Mass Index 40
Fasting Glucose 43
Inflammatory Measures 47
CHAPTER 3: VARIATION IN STRESS RESPONSE GENES IS RELATED TO FEATURES OF DIURNAL CORTISOL CURVES IN THE MULTI-ETHNIC STUDY OF ATHEROSCLEROSIS 55
Introduction 55
Methods 57
Study Population 57
Cortisol Sample Collection 58
Cortisol Features 58
Genetic Data 60
Stress Response Genes 61

Statistical Analysis 61
Gene-level analyses 61
Meta-analysis 62
Power Calculations 63
Results 65
Discussion 73
CHAPTER 4: INTERACTIONS BETWEEN CORTISOL LEVELS AND STRESS RESPONSE GENES IN PREDICTING CHRONIC DISEASE RISK FACTORS IN THE MULTI-ETHNIC STUDY OF ATHEROSCLEROSIS 77
Introduction 77
Methods 79
Study Population 79
Cortisol Sample Collection 79
Cortisol Features 80
Genetic Data 82
Stress Response Genes 82
Outcome variables 83
Statistical Analysis 84
Power Calculations 85
Results 87
Discussion 103
CHAPTER 5: A GENOME-WIDE ASSOCIATION STUDY (GWAS) OF SALIVARY CORTISOL CONCENTRATIONS: THE MULTI-ETHNIC STUDY OF ATHEROSCLEROSIS 107
Introduction 107
Methods 110
Study Population 110
Cortisol Sample Collection 111
Cortisol Features 111
Genetic Data 113
Statistical Strategy 113
Results 115
GWAS Results 116
Meta-analysis Results 119
Discussion 121
CHAPTER 6: CONCLUSION 127
Gene-based Association 127
Gene-by-Environment Interaction Studies 130
Genome-wide Association Studies of Complex Traits 134
Limitations 138

Concluding Remarks 139
APPENDIX 140
REFERENCES 200

List of Tables

Table 1: Chromosomal locations of the stress response genes. 9
Table 2: Selected MESA Exam components 14
Table 3: Characteristics of MESA Stress Study participants. 16
Table 4: Features of the diurnal cortisol curve. 21
Table 5: Number (percentage) of missing cortisol feature observations. 26
Table 6: Distributions of cortisol features by ethnic group. 26
Table 7: Pearson's product moment correlation coefficients between cortisol summary features. 31
Table 8: Univariate associations with cortisol features. 33
Table 9: Associations of race with cortisol features. 33
Table 10: Racial/ethnic group stratified multivariable associations of age, sex, and education with cortisol features. 34
Table 11: Multivariable associations of the top 10 principal components in each ethnic group on Wakeup. 36
Table 12: Multivariable associations of the top 10 principal components in each ethnic group on Bedtime. 37
Table 13: Multivariable associations of the top 10 principal components in each ethnic group on CAR. 37
Table 14: Multivariable associations of the top 10 principal components in each ethnic group on AUC. 38
Table 15: Multivariable associations of the top 10 principal components in each ethnic group on EDSlope. 38
Table 16: Multivariable associations of the top 10 principal components in each ethnic group on LDSlope. 39
Table 17: Multivariable associations of the top 10 principal components in each ethnic group on ODSlope. 39
Table 18: Ethnic group specific distributions of chronic disease risk factors. 42
Table 19: The univariate associations of covariates and cortisol features on average BMI, ln(Glucose), ln(IL6 + 1), and ln(TNF-α +1). 53
Table 20: Features of the diurnal cortisol curve 59
Table 21: Gene-level power calculations. 64
Table 22: Pearson's product moment correlation coefficients between cortisol summary features, all ethnic groups combined. 64
Table 23: Characteristics of MESA Stress Study participants. 65
Table 24: Distributions of cortisol summary features 66
Table 25: Gene-level main effect and meta-analysis results for *ADRA2A*. 67
Table 26: Gene-level main effect and meta-analysis results for *ADRB2*. 68
Table 27: Gene-level main effect and meta-analysis results for *SLC6A4*. 69

Table 28: Comparison of *SLC6A4* index SNPs across ethnic groups. 73
Table 29: Features of the diurnal cortisol curve 81
Table 30: Gene-level power calculations. 86
Table 31: Pearson's product moment correlation coefficients between cortisol summary features, all ethnic groups combined. 86
Table 32: Characteristics of MESA Stress Study participants. 87
Table 33: Distributions of cortisol summary features 88
Table 34: Distribution of outcome variables. 88
Table 35: Ethnic group specific distributions of chronic disease risk factors. 89
Table 36: Multivariable associations on chronic disease risk factors for AUC. 89
Table 37: Multivariable associations on chronic disease risk factors for Bedtime 89
Table 38: Multivariable associations on chronic disease risk factors for CAR. 90
Table 39: Multivariable associations on chronic disease risk factors for EDSlope. 90
Table 40: Multivariable associations on chronic disease risk factors for LDSlope. 90
Table 41: Multivariable associations on chronic disease risk factors for ODSlope. 91
Table 42: Multivariable associations on chronic disease risk factors for Wakeup. 91
Table 43: Gene-by-cortisol interactions for BMI. 93
Table 44: Gene-by-cortisol interactions for ln(Glucose). 94
Table 45: Gene-by-cortisol interactions for ln(IL-6 + 1). 95
Table 46: Gene-by-cortisol interactions for ln(TNF-a + 1). 96
Table 47: Individual SNP-by-ODSlope interactions in predicting ln(Glucose). 97
Table 48: Two degree of freedom test of individual SNP-by-ODSlope interactions in predicting ln(Glucose) 101
Table 49: Features of the diurnal cortisol curve 112
Table 50: Characteristics of MESA Stress Study participants. 115
Table 51: Distributions of cortisol summary features 116
Table 52: Most strongly associated SNPs for each cortisol summary feature 118
Table 53: METAL meta-analysis results. 120
Table 54: METAL assessment of heterogeneity. 121
Table A1: Chromosomal location of each stress response gene 140
Table A2: Start and end positions of genes of interest, ±5kb, and over size of gene regions. 140
Table A3: Number of SNPs with a minor allele frequency (MAF) of 5% or greater in each gene region for each ethnic group 140
Table A4: Gene-level main effect and meta-analysis results for *NR3C1* 141
Table A5: Gene-level main effect and meta-analysis results for *NR3C2* 142
Table A6: Gene-level main effect and meta-analysis results for *TH* 143
Table A7: Individual SNP-by-Bedtimee interactions in *ADRB2* in predicting ln(IL-6 + 1), with an interaction p-value < 0.05, in order of significance by ethnic group 144
Table A8: Individual SNP-by-Bedtime interactions in *NR3C2* in predicting ln(IL-6 + 1), with an interaction p-value < 0.05, in order of significance by ethnic group 145
Table A9: Individual SNP-by-EDSlope interactions in *NR3C2* in predicting ln(IL-6 + 1), with an interaction p-value < 0.05, in order of significance by ethnic group 147

Table A10: Individual SNP-by-Bedtime interactions in *NR3C1* in predicting ln(TNF-α + 1), with an interaction p-value < 0.05, in order of significance by ethnic group..... 149
Table A11: Individual SNP-by-Wakeup interactions in *NR3C2* in predicting ln(TNF-α + 1), with an interaction p-value < 0.05, in order of significance by ethnic group..... 151
Table A12: Two degree of freedom test of individual SNP-by-Bedtime interactions in predicting ln(IL-6 + 1) 153
Table A13: Two degree of freedom test of individual SNP-by-Bedtime interactions in predicting ln(IL-6 + 1) 155
Table A14: Two degree of freedom test of individual SNP-by-EDSlope interactions in predicting ln(IL-6 + 1) 158
Table A15: Two degree of freedom test of individual SNP-by-Bedtime interactions in predicting ln(TNF-α + 1) 161
Table A16: Two degree of freedom test of individual SNP-by-Bedtime interactions in predicting ln(TNF-α + 1) 165

List of Figures

Figure 1: The multi-step process of the HPA axis showing the production and feedback of glucocorticoids. 5
Figure 2: Distribution of cortisol collection times, presented in hours since wakeup. 18
Figure 3: Median log-transformed salivary cortisol concentration across three days (n=16,342 cortisol samples), presented in hours since wakeup. 20
Figure 4: Representation of the diurnal cortisol curve describing the cortisol features of interest. 20
Figure 5: Distributions of log-transformed cortisol features among African Americans. 27
Figure 6: Distributions of log-transformed cortisol features among European Americans. 28
Figure 7: Distributions of log-transformed cortisol features among Hispanic Americans. 29
Figure 8: Distribution of averaged BMI in kg/m^2 41
Figure 9: Distribution of averaged BMI in kg/m^2, by racial/ethnic group. 41
Figure 10: Distribution of averaged fasting glucose concentrations in mg/dl. 43
Figure 11: Distribution of natural log-transformed average fasting glucose concentrations. 44
Figure 12: Distribution of ln(Glucose) concentrations, by racial/ethnic group. 45
Figure 13: Scatter plot of natural log-transformed average fasting glucose concentrations versus age, by racial/ethnic group. 46
Figure 14: Distribution of IL-6 concentrations in pg/ml. 48
Figure 15: Distribution of ln(IL-6 + 1) in pg/ml. 49
Figure 16: Distribution of ln(IL-6 + 1) concentrations, by racial/ethnic group. 49
Figure 17: Distribution of TNF-α in pg/ml. 50
Figure 18: Distribution of ln(TNF-α + 1)in pg/ml. 51
Figure 19: Distribution of ln(TNF-α + 1) concentrations, by racial/ethnic group. 51
Figure 20: Representation of the diurnal cortisol curve describing the cortisol features of interest 59
Figure 21: LocusZoom plot of the correlation between loci of the *SLC6A4* gene region among European Americans in predicting EDSlope. 70
Figure 22: LocusZoom plot of the correlation between loci of the *SLC6A4* gene region among Hispanic Americans in predicting EDSlope. 71
Figure 23: LocusZoom plot of the correlation between loci of the *SLC6A4* gene region among African Americans in predicting EDSlope. 71
Figure 24: Representation of the diurnal cortisol curve describing our summary features of interest 81
Figure 25: LocusZoom plot of the interaction between loci of the *SLC6A4* gene region and EDSlope among African Americans in predicting ln(Glucose). 99

Figure 26: LocusZoom plot of the interaction between loci of the *SLC6A4* gene region and EDSlope among European Americans in predicting ln(Glucose). 99
Figure 27: Representation of the diurnal cortisol curve describing our summary features of interest 112
Figure 28: GWAS results for cortisol awakening response (CAR) in African Americans. 117
Figure A1: LocusZoom plot of the association between loci of the ADRA2A gene region among European Americans in predicting AUC 168
Figure A2: LocusZoom plot of the association between loci of the ADRA2A gene region among Hispanic Americans in predicting AUC 168
Figure A3: LocusZoom plot of the association between loci of the ADRA2A gene region among Hispanic Americans in predicting Bedtime. 169
Figure A4: LocusZoom plot of the association between loci of the ADRA2A gene region among African Americans in predicting CAR 169
Figure A5: LocusZoom plot of the association between loci of the ADRA2A gene region among European Americans in predicting EDSlope 170
Figure A6: LocusZoom plot of the association between loci of the ADRB2 gene region among European Americans in predicting Bedtime 170
Figure A7: LocusZoom plot of the association between loci of the *ADRB2* gene region among European Americans in predicting CAR 171
Figure A8: LocusZoom plot of the association between loci of the *ADRB2* gene region among European Americans in predicting ODSlope 171
Figure A9: LocusZoom plot of the association between loci of the *SLC6A4* gene region among European Americans in predicting CAR 172
Figure A10: LocusZoom plot of the association between loci of the *TH* gene region among African Americans in predicting EDSlope 172
Figure A11: LocusZoom plot of the association between loci of the *TH* gene region among African Americans in predicting LDSlope 173
Figure A12: Linkage disequilibrium plot for *SLC6A4* in European Americans. 174
Figure A13: Linkage disequilibrium plot for *SLC6A4* in Hispanic Americans 175
Figure A14: Linkage disequilibrium plot for *SLC6A4* in African Americans 176
Figure A15: LocusZoom plot of the interaction between loci of the *ADRB2* gene region and Bedtime among African Americans in predicting ln(IL-6 + 1) 177
Figure A16: LocusZoom plot of the interaction between loci of the *ADRB2* gene region and Bedtime among Hispanic Americans in predicting ln(IL-6 + 1) 177
Figure A17: LocusZoom plot of the interaction between loci of the *NR3C2* gene region and Bedtime among African Americans in predicting ln(IL-6 + 1) 178
Figure A18: LocusZoom plot of the interaction between loci of the *NR3C2* gene region and Bedtime among European Americans in predicting ln(IL-6 + 1) 178
Figure A19: LocusZoom plot of the interaction between loci of the *NR3C2* gene region and EDSlope among African Americans in predicting ln(IL-6 + 1) 179
Figure A20: LocusZoom plot of the interaction between loci of the *NR3C2* gene region and EDSlope among European Americans in predicting ln(IL-6 + 1) 179

Figure A21: LocusZoom plot of the interaction between loci of the *NR3C1* gene region and Bedtime among European Americans in predicting ln(TNF-α + 1) 180
Figure A22: LocusZoom plot of the interaction between loci of the *NR3C1* gene region and Bedtime among Hispanic Americans in predicting ln(TNF-α + 1) 180
Figure A23: LocusZoom plot of the interaction between loci of the *NR3C2* gene region and Wakeup among European Americans in predicting ln(TNF-α + 1) 181
Figure A24: Quantile-Quantile plot of the observed and expected p-values for associations with AUC among European Americans 182
Figure A25: Quantile-Quantile plot of the observed and expected p-values for associations with Bedtime among European Americans 182
Figure A26: Quantile-Quantile plot of the observed and expected p-values for associations with CAR among European Americans 183
Figure A27: Quantile-Quantile plot of the observed and expected p-values for associations with EDSlope among European Americans 183
Figure A28: Quantile-Quantile plot of the observed and expected p-values for associations with LDSlope among European Americans 184
Figure A29: Quantile-Quantile plot of the observed and expected p-values for associations with ODSlope among European Americans 184
Figure A30: Quantile-Quantile plot of the observed and expected p-values for associations with Wakeup among European Americans 185
Figure A31: Quantile-Quantile plot of the observed and expected p-values for associations with AUC among African Americans 185
Figure A32: Quantile-Quantile plot of the observed and expected p-values for associations with Bedtime among African Americans 186
Figure A33: Quantile-Quantile plot of the observed and expected p-values for associations with CAR among African Americans 186
Figure A34: Quantile-Quantile plot of the observed and expected p-values for associations with EDSlope among African Americans 187
Figure A35: Quantile-Quantile plot of the observed and expected p-values for associations with LDSlope among African Americans 187
Figure A36: Quantile-Quantile plot of the observed and expected p-values for associations with ODSlope among African Americans 188
Figure A37: Quantile-Quantile plot of the observed and expected p-values for associations with Wakeup among African Americans 188
Figure A38: Quantile-Quantile plot of the observed and expected p-values for associations with AUC among Hispanic Americans 189
Figure A39: Quantile-Quantile plot of the observed and expected p-values for associations with Bedtime among Hispanic Americans 189
Figure A40: Quantile-Quantile plot of the observed and expected p-values for associations with CAR among Hispanic Americans 190
Figure A41: Quantile-Quantile plot of the observed and expected p-values for associations with EDSlope among Hispanic Americans 190
Figure A42: Quantile-Quantile plot of the observed and expected p-values for associations with LDSlope among Hispanic Americans 191

Figure A43: Quantile-Quantile plot of the observed and expected p-values for associations with ODSlope among Hispanic Americans 191
Figure A44: Quantile-Quantile plot of the observed and expected p-values for associations with Wakeup among Hispanic Americans 192
Figure A45: GWAS results for Wakeup. Analyses were run separately for African Americans (AFA), European Americans (EA), and Hispanic Americans (HIS) 193
Figure A46: GWAS results for Bedtime. Analyses were run separately for African Americans (AFA), European Americans (EA), and Hispanic Americans (HIS) 194
Figure A47: GWAS results for cortisol awakening response (CAR). Analyses were run separately for African Americans (AFA), European Americans (EA), and Hispanic Americans (HIS) 195
Figure A48: GWAS results for Area under the Curve (AUC). Analyses were run separately for African Americans (AFA), European Americans (EA), and Hispanic Americans (HIS) 196
Figure A49: GWAS results for Early Decline Slope (EDSlope). Analyses were run separately for African Americans (AFA), European Americans (EA), and Hispanic Americans (HIS) 197
Figure A50: GWAS results for Late Decline Slope (LDSlope). Analyses were run separately for African Americans (AFA), European Americans (EA), and Hispanic Americans (HIS) 198
Figure A51: GWAS results for Overall Decline Slope (ODSlope). Analyses were run separately for African Americans (AFA), European Americans (EA), and Hispanic Americans (HIS) 199

Abstract

A growing body of work has examined the contribution of stress to various health outcomes. The hormone cortisol is likely to be a key mediator of the stress response that influences multiple physiologic systems that are involved in common chronic disease (including the cardiovascular system, immune system, and metabolism). An individual's daily cortisol response (e.g. waking, peak, end of day) has been shown to be patterned by race/ethnicity as well as socioeconomic factors. Despite evidence of associations of various risk factors with cortisol levels, considerable intra- and inter-individual variability in cortisol remains unexplained. Alone or through interaction with environmental features, genetic factors could contribute to unexplained variability in cortisol concentrations or cortisol responsivity. Furthermore, genetic factors may influence how cortisol affects a wide range of anthropometric, metabolic, and inflammatory processes underlying chronic disease risk. In this dissertation, both a candidate gene approach and a genome-wide association study (GWAS) were used to investigate genetic contributions to cortisol variability and its physiological effects in a sample of European Americans, African Americans, and Hispanic Americans from the Multi-Ethnic Study of Atherosclerosis (MESA). The sequence kernel association test (SKAT) was used for gene-based analysis. In the gene-based analysis of six stress response genes, we found that three genes had significant influence on cortisol features in

at least one ethnic group. Only one gene, *SLC6A4*, had a significant effect across ethnic groups in meta-analysis ($p < 0.05$). Extending this work to an analysis of gene-by-cortisol interaction effects on BMI, glucose, and inflammatory factors, we used SKAT and its interaction based extension to identify four genes (*ADRB2*, *NR3C1*, *NR3C2*, *SLC6A4*) that have significant evidence of interaction with cortisol features to influence anthropometric, metabolic, and inflammatory markers ($p < 0.05$). In GWAS, we found 18 regions with $p < 1x10^{-6}$ across the seven cortisol features evaluated in the three ethnic groups. Meta-analyses across ethnic groups identified only five genomic regions with $p < 5x10^{-6}$; none of the GWAS results replicated in meta-analysis. Overall, this dissertation illustrates that genetic analyses across ethnic groups can provide new insights into the role of genes in cortisol features and their relationship with chronic disease risk factors.

CHAPTER 1

INTRODUCTION

Introduction

Cortisol concentrations follow a strong daily pattern, where they are high upon awakening, reach a maximum concentration approximately half an hour later, and slowly decrease throughout the rest of the day [1-3]. This natural diurnal cycle is affected by lifestyle choices and daily stressors, and it impacts many physiological systems that underlie the increased risk of chronic diseases. Numerous studies have explored the environmental, psychological, social, and lifestyle factors that influence cortisol levels [4-7]. Only recently have we begun to investigate the potential genetic causes and modifications of cortisol and its influence on the pathophysiology of chronic diseases.

Several population-based studies have linked daily cortisol patterns to health outcomes, including elevated blood pressure, abdominal obesity, and coronary calcification [8-10]. For example, in a sample of 718 middle-aged black and white adults from the Coronary Artery Risk Development in Young Adults (CARDIA) study, individuals in the quartile with the flattest cortisol declines had higher odds of coronary artery calcification compared to individuals in other quartiles after adjustment for demographic, behavioral, and chronic disease risk factors (OR=2.58; 95% CI=1.26-5.30)

[10]. Cortisol concentrations and features of the cortisol curve have also been associated with diabetes [11] and inflammation [12]. Previous work in the Multi-ethnic Study of Atherosclerosis (MESA) has shown that persons with higher levels of interleukin-6 have a flatter cortisol awakening response and flatter decline [12].

Despite evidence of associations of several demographic factors with cortisol, including age, sex, and race/ethnicity[5, 6, 13, 14], considerable inter-individual variability in cortisol remains unexplained. This unexplained variation has led to increased interest in identifying genetic predictors of cortisol phenotypes [15]. Twin studies have shown a range of heritability estimates for cortisol concentrations, ranging from 0% to 84% [16-19]. However, in a combined analysis of cortisol heritability in multiple twin studies, basal cortisol concentrations had an estimated heritability of 62% [20]. These findings imply that cortisol concentrations have a genetic component.

Most of the genetic investigations to date have focused on candidate gene associations of cortisol, notably the glucocorticoid receptor gene (*NR3C1*) and the mineralocorticoid receptor gene (*NR3C2*), which code for receptors involved in the action of cortisol on the brain [21-23]. Polymorphisms in these receptor genes have been inconsistently associated with a range of factors, including body composition and insulin response [22, 24-26]. While a growing body of work has investigated European American and African American populations, limited information is available regarding other racial/ethnic groups [5, 6].

The MESA Study is a longitudinal cohort study focused on investigating the early stages of atherosclerosis. This multi-site study began in 2000, and aimed to identify risk

factors of subclinical cardiovascular diseases. The study includes more than 6,000 men and women recruited from six communities across the United States. A wide range of demographic, anthropometric, psychosocial, biochemical, physiological, genetic, and clinical data has been collected on MESA participants. In addition to the overarching MESA study, the MESA Stress Study collected multiple salivary cortisol samples across multiple days in a subsample of 1,000 MESA participants. The goal of the MESA Stress Study was to examine biological stress markers, such as salivary cortisol and salivary amylase, in relationship to the other psychosocial and chronic disease risk factors collected by MESA.

Psychosocial stress

The field of public health has been concerned with the impacts of stress for more than 50 years. Over the course of the last half-century the conceptualization of stress has grown. In the 1950's there was recognition that stress played a role in how the relationships between mental and physical health states could lead to clinically apparent disease [27]. A few years later, there was a shift in the understanding of stress from that of a negative force being exerted on the body to a potentially beneficial and necessary force. In a 1958 address, Dr. Howard Rusk stated that "[s]tress must be used as a therapeutic friend. It's a tool in our hands which I think we will find as valuable in the field of aging as antibiotics and some of the great new therapeutic treasures of the last decade seem to us now in the management of infectious diseases" [28].

In the last few decades the understanding of stress has expanded. It is now accepted that stress can manifest in multiple forms (acute, chronic) and can be the result of many different causes (physiological, psychological, and psychosocial). An individual's response to a stressor can also take multiple forms. The first is considered the fight-or-flight response, due to the activation of the sympathetic nervous system, while the second is an emotional, depressive response as a result of hypothalamic-pituitary-adrenal (HPA) axis activation [29]. The brain is the central hub of this multifaceted system. It is responsible for evaluating whether a given stressor is an acute or chronic signal, whether it is positive (health-promoting) or negative (health-damaging), and then finally for determining the appropriate response to that stressor. These responses could involve a wide range of behavioral or physiological reactions and recruit participation of the cardiovascular system, immune system, or metabolic system through the release of neurotransmitters (e.g. catecholamines, adrenaline, glucocorticoids) [30].

Much of the cascade of neurotransmitters in response to a stressor occurs through a multi-step pathway, the hypothalamic-pituitary-adrenal (HPA) axis. Figure 1 briefly summarizes this pathway where the hypothalamus sends corticotrophin releasing factor (CRF) to the pituitary gland, which releases adrenocorticotrophic hormone (ACTH) that is in turn picked up by the adrenal glands, which then release glucocorticoids. In addition to these factors having other peripheral effects, the glucocorticoids are involved in a negative feedback loop that turns off the HPA axis once a sufficient response to a stressor has occurred [31]. In the hypothalamus, the paraventricular nucleus releases CRF, which is

transported to the anterior pituitary, where it causes the release of ACTH into the blood stream. ACTH stimulates the adrenal cortex to synthesize and release the glucocorticoids cortisol (humans) or corticosterone (rodents). Glucocorticoids have a feedback mechanism at the level of the hippocampus, hypothalamus and pituitary to dampen excess activation of the HPA axis [31].

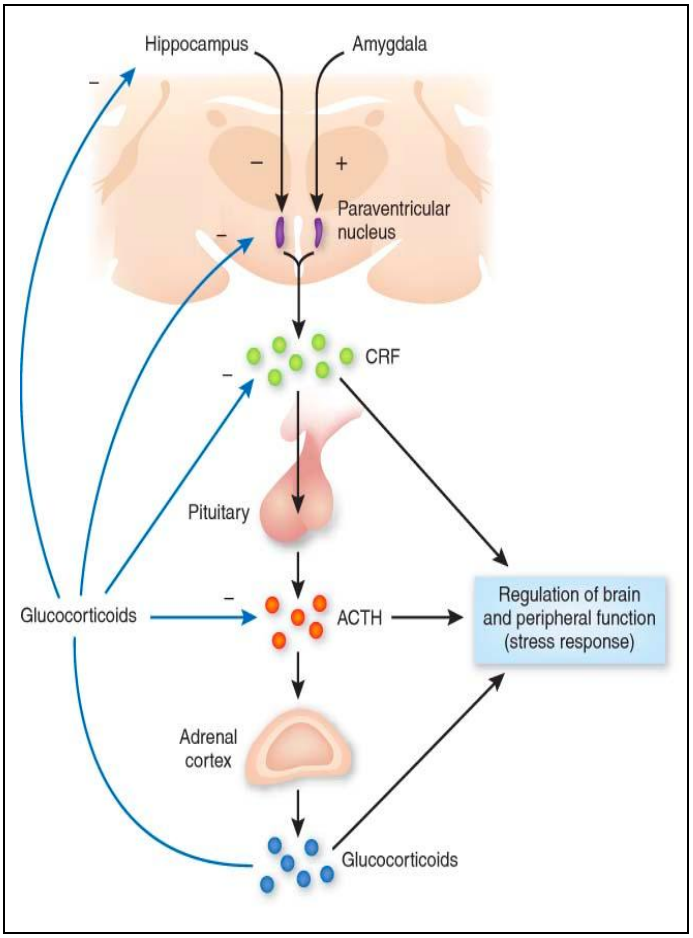


Figure 1: The multi-step process of the HPA axis showing the production and feedback of glucocorticoids [31].

Cortisol

Cortisol concentrations follow a strong daily pattern, where they are high upon awakening, reach a maximum concentration approximately half an hour later, and slowly decrease throughout the rest of the day [1-3]. Additionally, cortisol concentrations increase in response to stressors [32]. When exposed to a stressor the body activates the sympathetic nervous system and the HPA axis. Once activated, the HPA axis increases production of stress hormones, including cortisol, which are released from the adrenal glands according to an ultradian rhythm of hourly pulses [33]. After release from the adrenal glands, cortisol as a ligand is taken up by receptors in the brain (negative feedback loop).

In brain tissue, cortisol binds to two types of receptors, the mineralocorticoid receptors and the glucocorticoid receptors. The mineralocorticoid receptors are occupied under basal conditions of the ultradian rhythm [34, 35]. The glucocorticoid receptors, conversely, are only occupied when cortisol concentrations are high (e.g. in response to a stressor), due to the lower affinity of these receptors [36]. Polymorphisms in both the mineralocorticoid and glucocorticoid receptor genes have been associated with stress responsivity [37-39]. A more detailed discussion of the variation in these and other genes follows in a later section (Stress response genes).

The nature of within-person variability of cortisol concentrations over the day, as a result of these basal and stress related releases, requires the use of multiple cortisol measurements to completely assess an individual's pattern. Cortisol concentrations can be measured from multiple biological samples: urine, blood serum, and saliva. Urine and blood collection methods are difficult in population studies because of the need for

repeated collection, making the ability to measure cortisol concentrations in saliva samples an ideal alternative [40]. Salivary cortisol concentrations have been shown to be highly correlated with blood serum cortisol concentrations, with correlations ranging from 0.71-0.96 [2, 40-43].

Epidemiologic studies of cortisol

In addition to within-person variability of cortisol concentrations throughout the day, there is considerable variability in cortisol concentrations across individuals and between groups. In the control groups of a recent meta-analysis of studies comparing individuals with depression to non-depressed controls, average morning salivary cortisol concentrations ranged from a minimum of 7.8 nmol/L (SD=1.87) to 33.5 nmol/L (SD=9.5) while average evening salivary cortisol concentrations ranged from 2 nmol/L (SD=1.7) to 9.8 nmol/L (SD=5.9) [4]. Age has been shown to be a significant predictor of cortisol concentrations, where concentrations increase with age [13]. Additionally, there are gender differences with respect to salivary cortisol concentrations, with men having significantly higher mean levels than women [14]. Of particular note among women, salivary cortisol concentrations varied based on menstrual cycle phase and oral contraceptive use [14].

Moving beyond age and gender, there is evidence that an individual's daily cortisol profile is associated with race/ethnicity as well as socioeconomic factors. Flatter declines later in the day (less steep slopes) have been observed in African Americans and Hispanic Americans compared to European Americans. This pattern of flatter afternoon

decline has also been shown in lower socioeconomic status groups relative to higher socioeconomic status groups [5-7]. It has been suggested that chronic stress may explain the flatter declines in these individuals [44, 45]. Consistent with this hypothesis, cortisol patterns are also affected by psychosocial factors such as clinical hostility [46]and neighborhood sources of stress such as neighborhood violence [47]. One of the goals of this dissertation is to extend the body of knowledge beyond demographic and psychosocial factors to investigate the potential influence of genetic factors on cortisol patterns (Aim 3).

Stress response genes

The cortisol metabolic pathway suggests several key genes whose variation could affect cortisol levels or responsivity. In this dissertation we will examine six selected stress response genes: a glucocorticoid receptor gene (*NR3C1*), a mineralocorticoid receptor gene (*NR3C2*), the tyrosine hydroxylase gene (*TH*), the alpha-2A-adrenergic receptor gene (*ADRA2A*), the beta-2-adrenergic receptor gene (*ADRB2*), and the serotonin transporter gene (*SLC6A4*). Details on the chromosomal locations for each of these stress response genes can be found in Table 1.

Table1: Chromosomal locations of the stress response genes.

Stress Response Gene	**Location**
Alpha-2A-adrenergic receptor gene (*ADRA2A*)	10q24-q26
Beta-2-adrenergic receptor gene (*ADRB2*)	5q31-q32
Glucocorticoid receptor gene (*NR3C1*)	5q31.3
Mineralocorticoid receptor gene (*NR3C2*)	4q31.1
Serotonin transporter gene (*SLC6A4*)	17q11.1-q12
Tyrosine hydroxylase gene (*TH*)	11p5.5

Mineralocorticoid and glucocorticoid receptors are located in brain tissue, in the hippocampus and dentate gyrus, where they influence stress reactivity, both through the downstream effects of cortisol and through down regulation of the HPA axis after responding to a stressor [22,23]. These two receptors operate together in responding to cortisol levels [22]. More specifically, *NR3C2* (mineralocorticoid receptor gene) is involved under basal cortisol conditions and the early stages of response to a stressor, while *NR3C1* is involved when cortisol concentrations are higher, later in the stress response after the mineralocorticoid receptors are filled [21]. Polymorphisms in both receptors have been associated with the stress response [37-39], which supports the interplay just described. As a specific example, in the glucocorticoid receptor gene a well-studied SNP, N363S (an amino acid substitution from asparagine to serine; minor allele frequency 3%-7% [48]), has been found to increase cortisol responses to a psychosocial stressor [22].

The tyrosine hydroxylase (*TH*) gene, the alpha-2A-adrenergic receptor gene (*ADRA2A*) and the beta-2-adrenergic receptor gene (*ADRB2*) were selected as stress response candidate genes given prior work demonstrating associations between

polymorphisms in these genes and stress responsivity. Tyrosine hydroxylase is involved in catecholamine biosynthesis [49]. In response to a stressor, catecholamine synthesis is up-regulated as a result of tyrosine hydroxylase (*TH*) transcription [50]. Genetic polymorphisms of *TH* have been previously associated with catecholamine excretion [51] as well as hemodynamic responses to stress [51, 52], another demonstration of how stress has downstream implications on the cardiovascular system.

The two adrenergic receptor genes also have hemodynamic implications. Located in the brain stem, alpha-2 adrenergic receptors impact blood vessels through the release of noradrenaline and adrenaline [53]. Polymorphisms of this gene (a restriction fragment length polymorphism, resulting in a 6.3- or 6.7-kb allele, with the shorter allele being the minor allele) have been associated with the response to environmental stressors [54]. Vasodilation occurs when adrenaline fills beta-2 adrenergic receptors, which can offset the hypertensive effects induced by the sympathetic nervous system in response to stressors [55]. Polymorphisms of *ADRB2*, the most functionally relevant being Arg16/Gly, have been previously shown to be related to blood pressure under conditions of physical and mental stress [56].

The serotonin transporter gene (*SLC6A4* or *5-HTT*) gained widespread recognition after demonstration of gene-by- psychological stressor interaction. However, difficulties with replication of promoter region (designated *5-HTTLPR*) polymorphisms demonstrating interaction effects with social and psychological stressors [57-61] has led to the hypothesis *5-HTTLPR* may more generally have implications for stress responsivity through serotoninergic actions [62], rather than having specific implications for psychiatric

outcomes, which was the focus of the original research. This hypothesis of the promoter region polymorphism conferring increased stress responsivity or susceptibility has been supported by brain imaging studies [63]. It has been suggested that serotoninergic activity may be related to stress responsivity through activation of the sympathetic nervous system [64]. Additionally, the *5-HTTLPR* polymorphisms have also been shown to be associated with cortisol features, specifically the cortisol awakening response [62]

Biological mechanisms linking stressors and chronic disease risk factors

The pathway through which stress has downstream impacts on chronic disease risk factors involves the activation of the HPA axis. Cushing's syndrome, which is the results of chronic hypercortisolism, has a variety of anthropometric and metabolic characteristics, which include altered fat distribution as well as glucose intolerances and diabetes [65]. Even under less extreme circumstances, cortisol concentrations influence these anthropometric and metabolic systems. Under conditions of chronic stress, which would result in prolonged exposure to increased circulating stress hormones and even HPA axis feedback dysregulation, the implications for glucose metabolism and anthropometric consequences may be more pronounced [22, 29].

Beyond Cushing's syndrome, cortisol levels have been previously associated with obesity [29, 66, 67] as well as diabetes-related outcomes [8, 68-70]; however, findings have not always been consistent [71, 72]. There are a number of possible reasons for the inconsistent findings. First, studies are generally limited in sample size. Second, studies are often limited by the available cortisol measures, perhaps with samples collected only for one

day or for only one point in time (e.g. wakeup concentrations). The lack of large, population-based samples limit the generalizability of findings and the restricted availability of cortisol measures throughout the day limit the interpretation of associations with anthropometric and metabolic factors.

In addition to anthropometric and metabolic risk factors, inflammatory factors are also of interest. There has been recent evidence that stress, particularly chronic stress, has implications for inflammatory systems, resulting in increased inflammatory marker concentrations [73-75]. Two recent studies have shown that chronic stress results in up-regulation of inflammatory responses, even when cortisol concentrations are not heightened [76, 77]. The work on anthropometric, metabolic, and inflammatory factors combined suggests that cortisol concentrations do not have to be extreme to have downstream implications on chronic disease risk factors.

In this dissertation features of daily cortisol profiles as well as a wealth of epidemiologic data from a sample of European American, African American, and Hispanic American participants in the MESA Stress Study were used to examine gene-level associations between the known stress response genes and cortisol features (Chapter 3), assess the influence of cortisol and known stress response genes on anthropometric, metabolic, and inflammatory factors (Chapter 4) and identify novel genome-wide loci associated with cortisol features (Chapter 5).

CHAPTER 2

STUDY POPULATION AND VARIABLE MEASUREMENTS

Study Population

The MESA Study

The Multi-Ethnic study of Atherosclerosis (MESA) is a longitudinal cohort study focused on investigating the early stages of atherosclerosis. This multisite study began participant recruitment in 2000 from six communities across the United States: Baltimore, MD; Chicago, IL; Forsyth, NC; Los Angeles, CA; New York, NY; and St. Paul, MN. In order to be eligible to participate in the MESA Study, prospective participants had to be two criteria: be 45-84 years of age and free from history of cardiovascular disease. At baseline, there were a total of 6,814 eligible men and women enrolled. The MESA Study was designed to be a multi-ethnic study; as such an equal number of men and women from at least two ethnic groups, with a target of approximately 1,100 participants, were recruited from each of the six communities. Multiple ethnic groups were recruited at each site in order to minimize confounding of ethnicity by site. Due to the variation in source population size and ethnic composition, a probability-based sample of participants was selected to achieve the desired age, gender, and ethnic group samples needed to reach an adequate number of new cardiovascular disease (CVD) events and to establish associations of risk factors with CVD events. A

variety of population-based recruitment approaches developed by each field center were used to create this sample (e.g. lists of area residents, random digit dialing, etc.).

Baseline examinations occurred from 2000 to 2002 and four follow-up examinations have occurred. Each examination consisted of questionnaire completion (in English, Spanish, or Chinese), a blood draw, anthropometric measurements, blood pressure measurement, and assessment of subclinical cardiovascular disease (Table 2). In an effort to minimize loss to follow-up, participants were contacted every nine to 12 months for information on cardiovascular endpoints, including diagnosis of new conditions, hospitalizations, treatments, interventions and behavioral changes in risk factors, as well as mortality [78].

Table 2: Selected MESA Exam components

Exam component	Exam 1	Exam 2	Exam 3	Exam 4	Exam 5
Year	2001*	2003	2004	2006	2011
Subclinical disease	X	X	X	X	X
Anthropometry	X	X	X	X	X
Blood Pressure	X	X	X	X	X
Phlebotomy	X	X	X	X	X
Questionnaires					
SES	X	X	X	X	X
Medical History	X	X	X	X	X
Medications	X	X	X	X	X
Diet	X				X
Physical Activity	X	X	X		X
Psychosocial	X	X	X	X	X

*When exam spans two years, midyear is shown.

The MESA Stress Study

In addition to the overarching MESA study an ancillary study, the MESA Stress Study, collected detailed stress hormone data, including multiple salivary cortisol samples over multiple days (see details below), on a subsample of 1002 MESA participants. MESA Stress Study participants were recruited from the New York and Los Angeles sites. The MESA Stress Study data collection occurred from 2004 to 2006, during the second and third follow-up examinations of the MESA cohort. Participants for the MESA Stress Study were African Americans, European Americans, and Hispanic Americans and were enrolled as they presented for follow-up, until approximately 500 participants were recruited at each of the two locations. Information on age, gender and race/ethnicity was obtained from all MESA participants at baseline. Race and ethnicity was characterized using participants' responses to questions modeled on the year 2000 census. Participants were classified as Hispanic, non-Hispanic white and non-Hispanic black. Basic demographic and health status information on the 1002 Stress Study participants at baseline is provided in Table 3.

Table 3: Characteristics of MESA Stress Study participants.

	Frequency (n=1002)
Site	
Columbia	52.2%
UCLA	47.8%
Age	
45-54	29.9%
55-64	27.7%
65-74	30.3%
75-84	12.1%
Race	
European American	18.6%
African American	28.6%
Hispanic American	52.8%
Gender	
Male	47.6%
Female	52.4%
Education Level	
Less than High School	27.0%
Completed High School	20.2%
Some College	29.7%
Bachelor's or higher	23.2%
Income	
< $20,000	29.3%
$20,000-34,999	27.5%
$35,000-$49,999	16.5%
$50,000 or higher	26.8%
Percent Current Smokers	11.3%
Percent Diabetic	13.5%
Body Mass Index (BMI) >=30	36.7%

Compared to other African American, European American, and Hispanic American participants at the New York and Los Angeles Field Centers, there was a smaller proportion of Stress Study participants in the oldest age category of 75 years or older (18.2% overall versus 12.1% in the Stress Study). There were also slightly more males in the Stress Study (47.6% versus 44.7%), and more Stress Study participants had

completed at least some college (29.7% versus 23.9%). Stress Study participants had a higher prevalence of obesity (BMI >=30) (36.7% versus 33.3%) and a lower prevalence of Diabetes (13.5% versus 17.6%).

Cortisol Samples and Cortisol Features

Daily Salivary Cortisol Samples

Each MESA Stress Study participant was given detailed instructions on the collection of daily salivary cortisol samples. Each participant was to collect six saliva samples per day over three consecutive weekdays, for a maximum of 18 samples per participant, using Salivette collection tubes. The samples were to be taken using the following schedule: the first sample taken upon waking and before getting out of bed (Sample 1); the second sample taken 30 minutes later (Sample 2); the third sample at 10:00am (Sample 3); the fourth sample at 12:00 noon or before lunch, whichever came first (Sample 4); the fifth sample at 6:00pm or before dinner, whichever came first (Sample 5); the sixth sample at bedtime (Sample 6). Each Salivette collection tube was equipped with a time tracking device on the cap (track-caps), which recorded the time when the swabs were removed for sample collection. Participants were aware of the time tracking device, and earlier work has shown that knowledge of the time tracking device improves sample collection compliance [79]. Saliva samples were stored at -20 degrees Celsius until being prepared for assay. Frozen samples were thawed and centrifuged at 3000 rmp for 3 minutes before cortisol levels were determined using a chemi-

luminescence assay with a high sensitivity of 0.16 ng/mL (IBL – Hamburg, Germany). Intra- and inter-assay coefficients of variation were <8%.

Figure 2 shows the distribution of the recorded time of each sample collected, relative to time since wakeup. The initial peak along the y-axis represents the number of samples taken at wakeup. The second peak should have been at 30 minutes after wakeup, however note that there is more of a cluster from 30 minutes to an hour and a half after wakeup. The first mode from approximately an hour and a half to 8 hours after wakeup, captures the 10:00am and lunchtime samples, while the second and third modes capture dinner and bedtime samples, respectively.

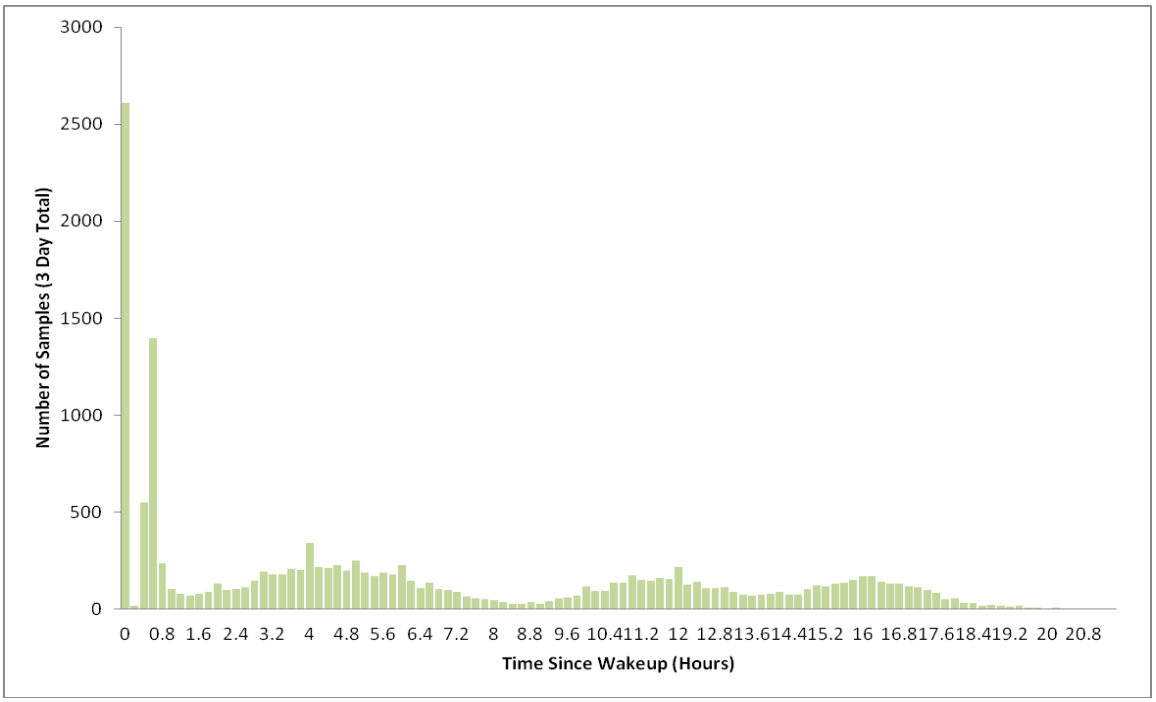


Figure 2: Distribution of cortisol collection times, presented in hours since wakeup.

The cortisol profile has a diurnal shape, rising and falling during the course of the day. In general, cortisol concentrations are high upon awakening (1st sample) and reach a peak approximately 30 minutes later (2nd sample). Concentrations then quickly decline between the 2nd and 3rd samples, after which the declining rate gradually flattens out through the 6th samples. Raw cortisol sample concentrations were measured in nmol per liter [80]. As a data preprocessing step, cortisol concentrations were log-transformed to more closely approximate a normal distribution [5, 12, 80]. Figure 3 shows the median log-transformed cortisol concentration, in time since wakeup. Median log-transformed cortisol concentrations among MESA Stress participants follow the expected pattern, fairly high at wakeup, peaking shortly thereafter, and then declining throughout the day. From hours 16 to 19 after wakeup concentrations plateau, then begin to climb again. When referencing cortisol concentrations from here forward, we mean the log-transformation concentrations, unless otherwise noted.

Cortisol Features

Specifically for these analyses, several cortisol features were computed: Wakeup, Bedtime, Cortisol awakening response (CAR), Area under the curve (AUC), Early Decline Slope, Late Decline Slope, and Overall Decline Slope. Figure 4 is graphical representation of these features. The Peak measurements were used for feature calculation, but were not used in analysis. Table 4 provides a description of the features used in this dissertation.

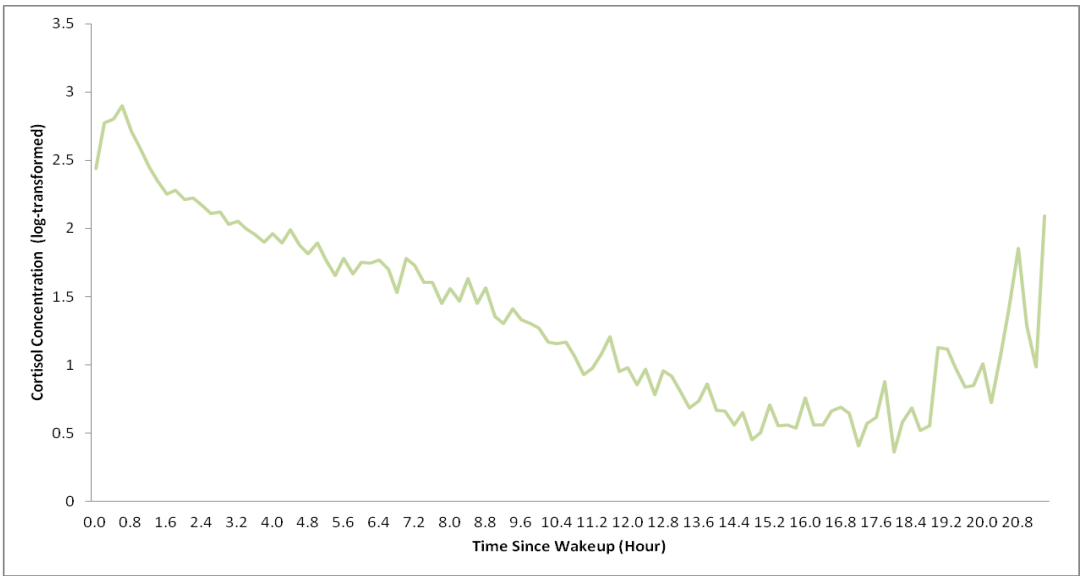
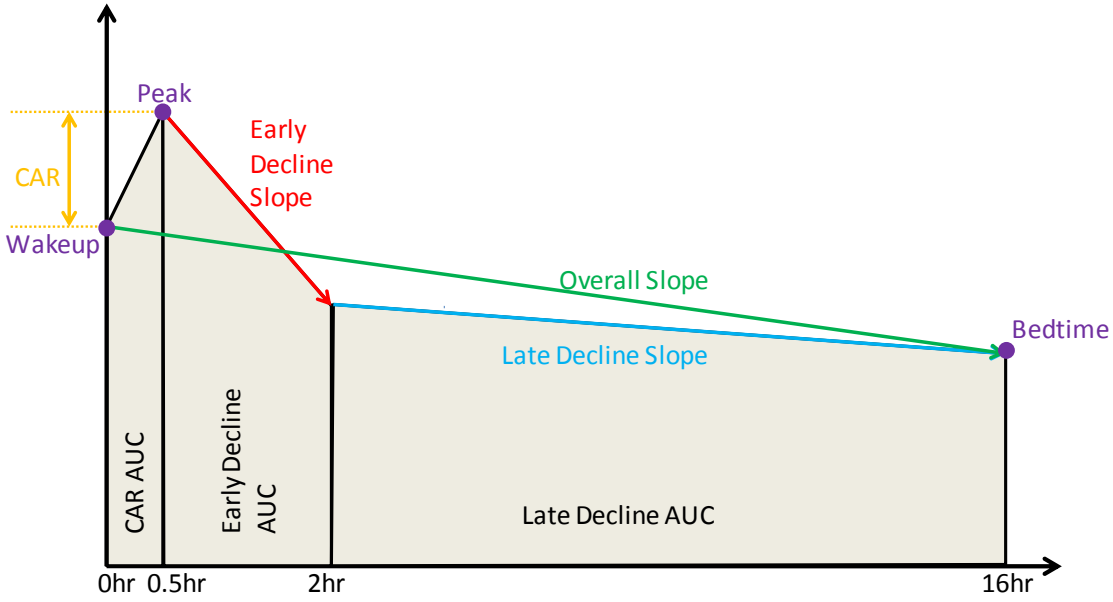


Figure 3: Median log-transformed salivary cortisol concentration across three days (n=16,342 cortisol samples), presented in hours since wakeup.



AUC 16hr = CAR AUC + Early Decline AUC + Late Decline AUC

Figure 4: Representation of the diurnal cortisol curve describing the cortisol features of interest. For these analyses we specifically used Wakeup, Bedtime, Cortisol awakening response (CAR), Area under the curve (AUC) from 0-16 hours, Early Decline Slope, Late Decline Slope, and Overall Decline Slope.

Table 4: Features of the diurnal cortisol curve. Cortisol concentrations were log-transformed.

	Cortisol Feature	Description
Time points	Wakeup	Average cortisol concentration from wakeup for an individual (Sample 1).
	Bedtime	Average cortisol concentration at bedtime for an individual (Sample 6).
Area	Area under the curve (AUC)	Standardized AUC for the interval 0hr-16hr since wakeup averaged across all days for an individual
Slopes	Cortisol awakening response (CAR)	The average difference in cortisol concentrations between the peak and wakeup measurements (Sample 2 – Sample 1).
	Early Decline Slope (EDSlope)	The slope from 0.5 hours and 2 hours since wakeup pooled across all days for an individual.
	Late Decline Slope (LDSlope)	The slope from 2 hours to 16 hours since wakeup pooled across all days for an individual
	Overall Decline Slope (ODSlope)	The overall decline slope ignoring the peak value from wakeup to bedtime pooled across all days for an individual.

Time point specific cortisol features, Wakeup and Bedtime, were created by taking the average of the log-transformed time point specific cortisol measurement for an individual across all three days of cortisol sample collection. Averages were created across all available days of data, such that Wakeup or Bedtime values were considered missing if there were no available Wakeup or Bedtime measurements for an individual across all three days.

The cortisol awakening response is the slope between Sample 2 and Sample 1. The Cortisol awakening response variable (CAR) is the average of the three days' slopes. If either Sample 1 or Sample 2 for a given day is missing, then that days' CAR is then missing. Based on the cortisol sampling protocol, Sample 2 was supposed to be taken ~30 minutes after wakeup. However, if Sample 2 was taken more than one hour after wakeup, it was considered missing. The average CAR was taken over the days with CAR available, such that average CAR is only considered missing if CAR was not

available for all the three days. Missingness for CAR is mainly due to compliance issues with the timing of Sample 2 collection.

An area under the curve measure was calculated for each day of cortisol availability, and the three days were averaged together to create AUC. For the daily area under the curve calculations, the trapezoid rule was used to estimate the area beneath the cortisol curve for an individual. Dividing the area estimate by the length of the time period for which the estimate was calculated, which was 16 hours, was used to standardize this area estimate. The standardized area measure can be interpreted as the average log(cortisol) concentration during the interval from wakeup to 16 hours later.

Determination of the 16 hour window that was used for the AUC calculations varied depending on the timing of sample collection for an individual participant. If Sample 5 was taken within the 16-hour window, then Sample 5 and Sample 6 were connected by a straight line that was used to predict the log(cortisol) concentration at 16 hours after wakeup. If Sample 5 was taken later than 16 hours since wakeup, then Samples 4 and 5 were connected by a straight line and used to predict the log(cortisol) concentration at 16 hours after wakeup. In the situation where less than 3 samples were collected on a given day, or both of the two samples needed to predict log(cortisol) at 16 hours since wakeup were missing, then the daily area estimate was considered missing. The averaged AUC measure used for analysis was considered missing only if all three daily area estimates were missing.

Individuals who stayed awake for longer hours had a wider range of the cortisol curve captured than those individuals who stayed awake for fewer hours. Specifically for

the computation of the slope features, combining these different bedtime patterns prohibits a true representation of the declining trend by underestimating the extent of decline as a result of including the rise in concentrations seen on the right hand side of the x-axis is Figure 3. Therefore, in order to more accurately compute the declining slopes, samples taken after 16 hours since wake up were excluded. In choosing the cutoff of 16 hours we are not losing a substantial amount of information as only approximately 8% of samples were taken after the 16-hour window.

The Early Decline Slope (EDSlope) is the difference between Sample 2 and Sample 3 pooled across all days for an individual. Piecewise linear model, shown below, was used to estimate the slope between Sample 2 and Sample 3 by pooling across all available days of cortisol collection:

$$y_i = \beta_0 + \beta_1 t_j + \beta_2(t_j - 0.5) + \beta_3(t_j - 2) + \epsilon_j$$

where, y_j is the log(cortisol) concentration and t_j is the time of sample collection. The estimate of slope was $\hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3$,which is in the units of log(cortisol) concentration per hour. EDSlope was only considered missing if the piecewise linear model could not be fit due to design matrix singularity or the value was identified as an outlier. The Rosner Extreme Studentized Deviate method was used to identify and remove outlying observations [81]. There were 10 outliers removed from EDSlope.

The Late Decline Slope (LDSlope) is the difference between 2 hours and 16 hours after wakeup pooled across all available days of cortisol collection. The estimate of the slope was obtained by fitting a simple linear model: $y_j = \beta_0 + \beta_1 t_j + \epsilon_j$

where, y_j is the log(cortisol) concentration, t_j is the time of sample collection, and the slope is $\hat{\beta}_1$ (log(cortisol) concentration per hour). LDSlope could not be defined when there were fewer than two data points available after 2 hours since wakeup across all days for an individual, and therefore was considered missing. The Rosner Extreme Studentized Deviate method was also used for LDSlope to identify and remove outlying observations [81]. There were 10 outliers removed LDSlope

The Overall Decline Slope (ODSlope) is the difference between wakeup and 16 hours after wakeup, pooled across all available days of cortisol collection. For this feature we were interested in the overall difference between where and individual's concentration starts at wakeup and their end of the day measurement, and therefore ignored the influence of other time points, specifically the Peak, which have been captured in the other features. The estimate of the slope was obtained by fitting a simple linear model: $y_j = \beta_0 + \beta_1 t_j + \epsilon_j$

where, y_j is the log cortisol concentration, t_j is the time of sample collection, and the slope is $\hat{\beta}_1$ (log(cortisol) concentration per hour). If after excluding the Peak value there were fewer than two data points for an individual, then ODSlope was defined as missing.

Details on the number of total missing or removed observations are available in Table 5. The CAR has the greatest number of missing observations. Most of this missingness is due to sampling compliance issues with cortisol Sample 2. Based on the sampling protocol Sample 2 was to be taken at ~30 minutes after wakeup. Often, this sample was taken an hour or more after wakeup. At that time interval, the concentration

captured by the sample would no longer be representative of the maximum concentration for that day, as the cortisol curve quickly declines from the peak. Therefore, when Sample 2 was taken too late it was no longer representative of the peak and was defined as missing. With the exception of the CAR, overall missingness for each cortisol feature is at roughly 3% or less.

The distributions of the cortisol features, by ethnic group, are represented in Table 6. The cortisol feature means varied across ethnic groups. Wakeup means ranged from 2.38 to 2.58, Bedtime means ranged from 0.49 to 0.98, CAR means ranged from 0.35 to 0.45, AUC means ranged from 1.46 to 1.64, EDSlope means ranged from -0.53 to -0.40, LDSlope means ranged from -0.13 to -0.10, and ODSlope means ranged from -0.12 to -0.10. The ANOVA procedure was used to assess whether or not the mean for each cortisol feature differed across the ethnic groups. There was a statistically significant difference in means across the ethnic groups for all cortisol features except CAR. Figures 5-7 demonstrate the distributions of the cortisol features, by ethnic group, and their approximation of the normal distribution. Overall, the log-transformed features have a centered, bell-shaped distribution. The Hispanic American features, in general, have longer tail distributions than the other ethnic groups.

Table 5: Number (percentage) of missing cortisol feature observations.

	Cortisol Feature							
	Wakeup	**Bedtime**	**CAR**	**AUC**	**EDSlope**	**LDSlope**	**ODSlope**	**Total**
Total Missing	9 (0.98%)	13 (1.42%)	68 (7.42%)	23 (2.51%)	28 (3.05%)	7 (0.76%)	4 (0.44%)	917
Race								
African Americans	1 (0.42%)	3 (1.27%)	13 (5.49%)	7 (2.95%)	6 (2.53%)	1 (0.42%)	1 (0.42%)	237
European Americans	4 (2.22%)	4 (2.22%)	10 (5.56%)	4 (2.22%)	7 (3.89%)	3 (1.67%)	1 (0.56%)	180
Hispanic Americans	4 (0.80%)	6 (1.20%)	45 (9.00%)	12 (2.40%)	15 (3.00%)	3 (0.60%)	2 (0.40%)	500
Gender								
Male	4 (0.90%)	6 (1.35%)	32 (7.21%)	12 (2.70%)	13 (2.93%)	3 (0.68%)	2 (0.45%)	444
Female	5 (1.06%)	7 (1.48%)	36 (7.61%)	11 (2.33%)	15 (3.17%)	4 (0.85%)	2 (0.42%)	473

Table 6: Distributions of cortisol features by ethnic group.

	African Americans		**European Americans**		**Hispanic Americans**		**ANOVA**
Cortisol Feature	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	p-value
Wakeup	214	2.38 (0.55)	166	2.58 (0.54)	450	2.38 (0.58)	0.0002
Bedtime	212	0.98 (0.74)	166	0.78 (0.77)	448	0.49 (0.84)	<0.0001
CAR	203	0.35 (0.46)	160	0.45 (0.46)	412	0.37 (0.52)	0.17
AUC	209	1.60 (0.42)	166	1.64 (0.43)	442	1.46 (0.51)	<0.0001
EDSlope	209	-0.42 (0.44)	163	-0.53 (0.35)	433	-0.40 (0.44)	0.003
LDSlope	211	-0.10 (0.06)	164	-0.12 (0.06)	447	-0.13 (0.06)	<0.0001
ODSlope	214	-0.10 (0.06)	169	-0.12 (0.07)	452	-0.12 (0.06)	<0.0001

Cortisol concentrations (nmol/L) were log-transformed and averaged across the three days of collection to create each feature. SD = Standard Deviation.

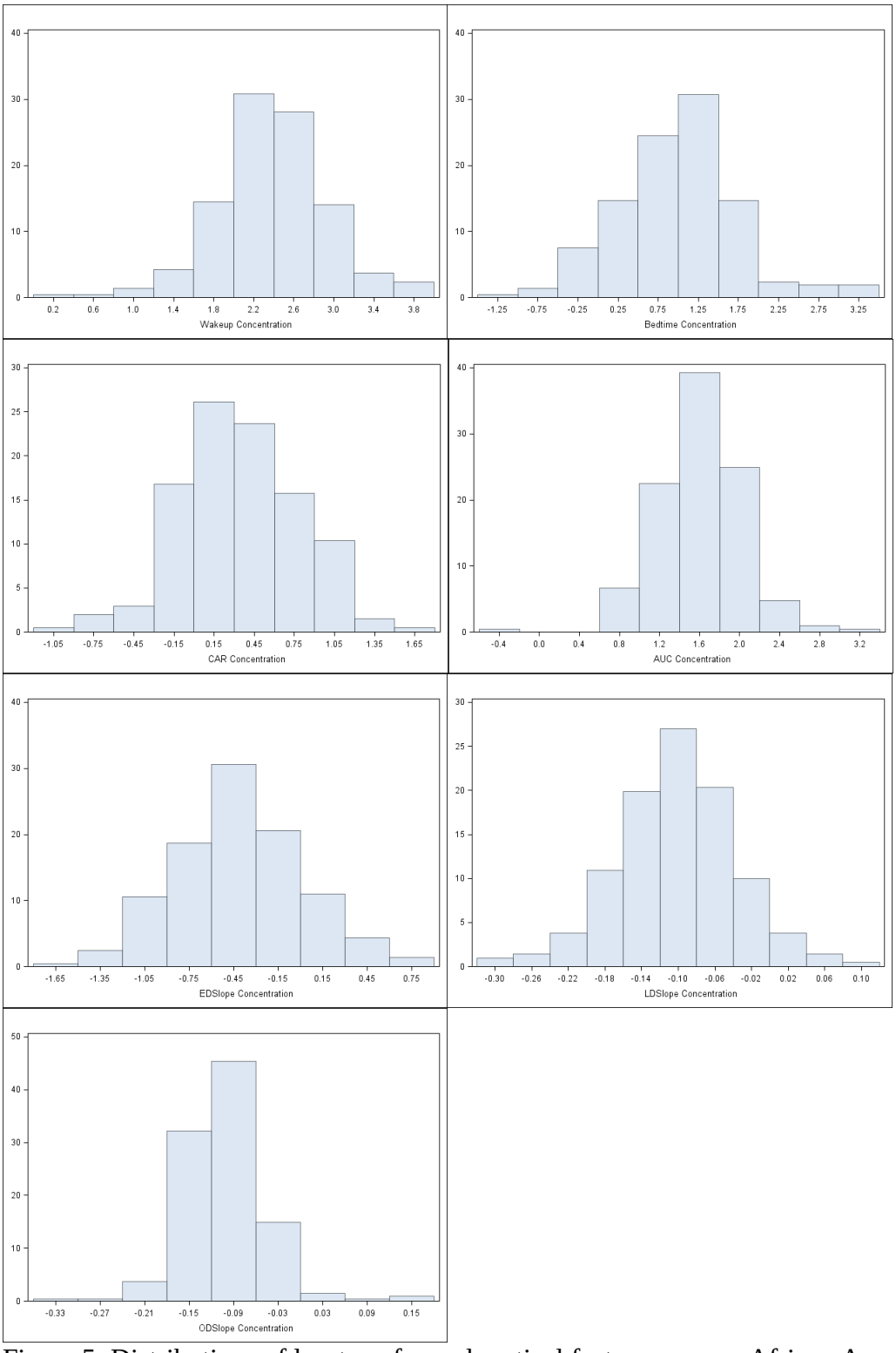


Figure 5: Distributions of log-transformed cortisol features among African Americans.

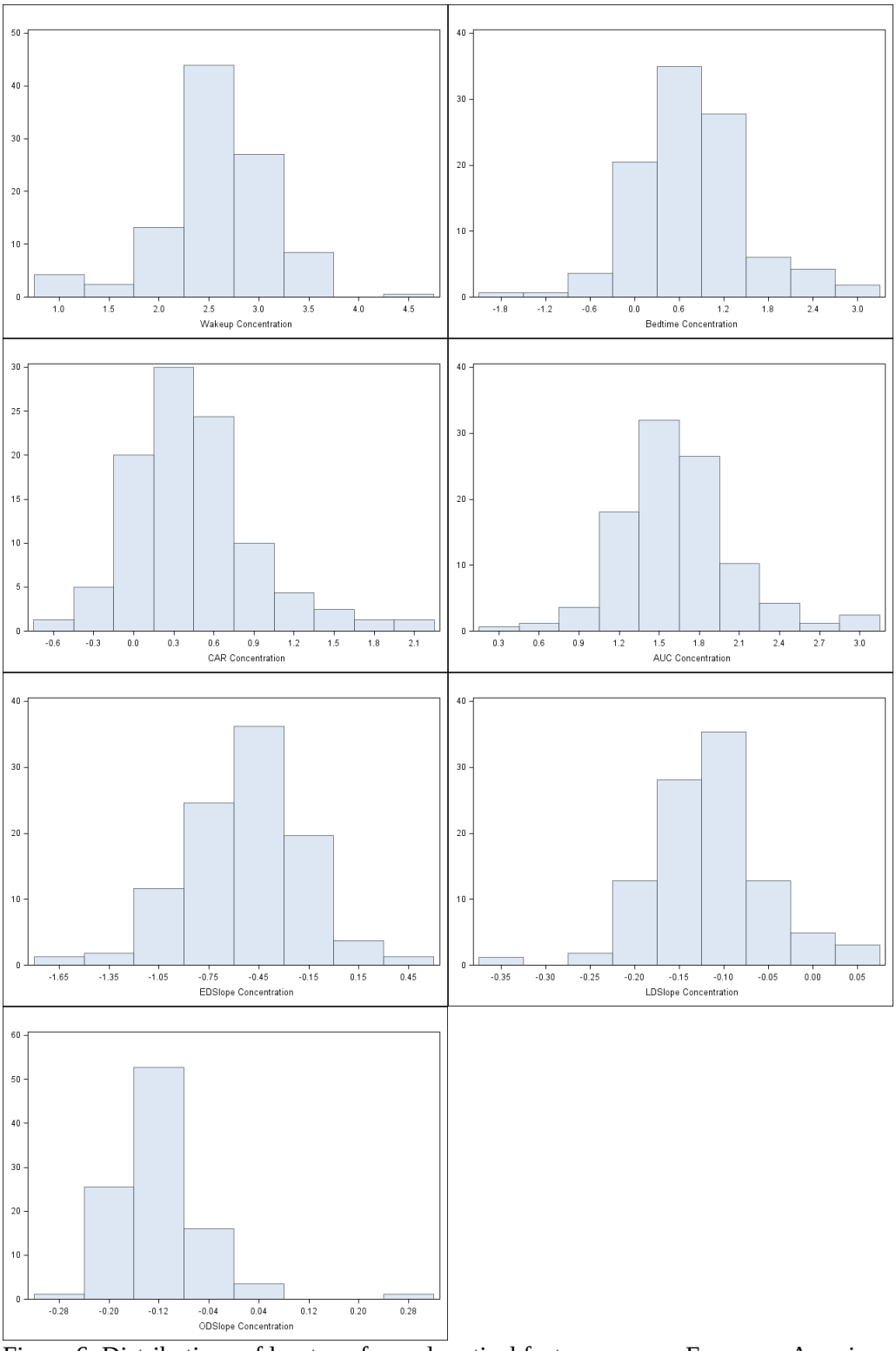


Figure 6: Distributions of log-transformed cortisol features among European Americans.

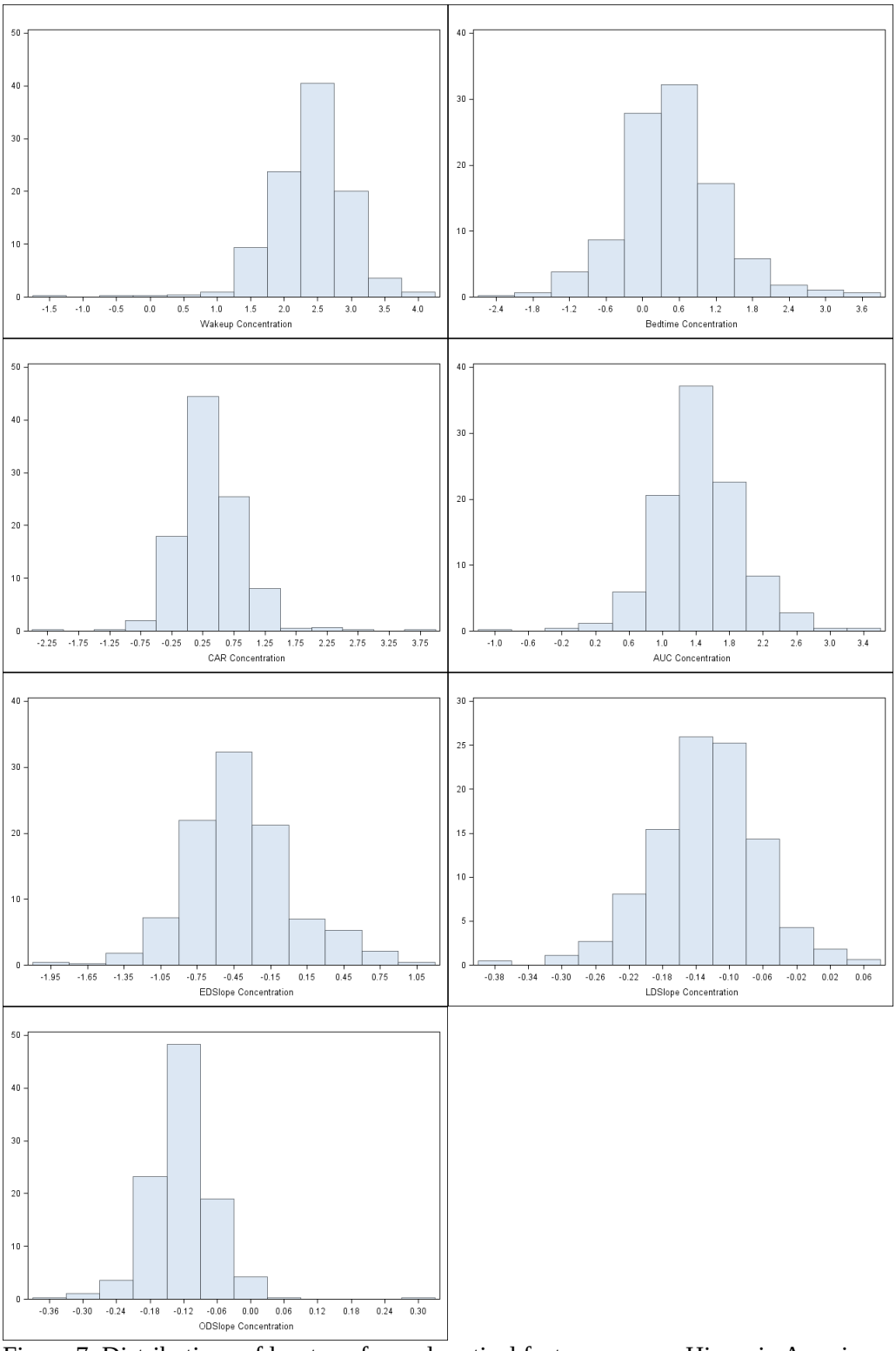


Figure 7: Distributions of log-transformed cortisol features among Hispanic Americans.

Given that the cortisol features are derived from the same data points, it was expected that some of the features would be highly correlated. Pearson's product moment correlation coefficients were used to assess the correlation between features (Table 7). The most highly correlated features were Bedtime with AUC, and LDSlope with ODSlope, which ranged from a correlation of approximately 0.6 to 0.8 depending on ethnic group. The bolded correlations coefficients represent significant ($p<0.05$) correlations between any two cortisol features. In those instances where the correlation coefficient is not significant, generally those cortisol features were derived from different data points. Take the correlation between Bedtime and EDSlope. Unlike the other decline slope features which are anchored by end of day cortisol concentrations, EDSlope is not computed using Bedtime concentrations, which explains why the two features are uncorrelated.

Table 7: Pearson's product moment correlation coefficients between cortisol summary features.

Panel A: Combined ethnic groups							
	Wakeup	**Bedtime**	**CAR**	**AUC**	**EDSlope**	**LDSlope**	**ODSlope**
Wakeup	1.00						
Bedtime	**0.39**	1.00					
CAR	**-0.44**	**-0.14**	1.00				
AUC	**0.49**	**0.74**	-0.01	1.00			
EDSlope	**-0.24**	-0.05	**-0.29**	**0.08**	1.00		
LDSlope	**0.08**	**0.50**	**-0.08**	**0.35**	**-0.39**	1.00	
ODSlope	**-0.34**	**0.43**	**0.18**	**0.39**	0.04	**0.72**	1.00
Panel B: African Americans							
	Wakeup	**Bedtime**	**CAR**	**AUC**	**EDSlope**	**LDSlope**	**ODSlope**
Wakeup	1.00						
Bedtime	**0.38**	1.00					
CAR	**-0.30**	-0.13	1.00				
AUC	**0.51**	**0.66**	-0.07	1.00			
EDSlope	**-0.36**	-0.13	**-0.30**	0.05	1.00		
LDSlope	0.10	**0.46**	-0.08	**0.30**	**-0.38**	1.00	
ODSlope	**-0.41**	**0.36**	0.04	**0.35**	**0.19**	**0.60**	1.00
Panel C: European Americans							
	Wakeup	**Bedtime**	**CAR**	**AUC**	**EDSlope**	**LDSlope**	**ODSlope**
Wakeup	1.00						
Bedtime	**0.23**	1.00					
CAR	**-0.58**	-0.08	1.00				
AUC	**0.39**	**0.66**	0.04	1.00			
EDSlope	**-0.18**	0.01	**-0.20**	**0.19**	1.00		
LDSlope	-0.05	**0.38**	-0.05	**0.28**	**-0.35**	1.00	
ODSlope	**-0.40**	**0.30**	**0.31**	**0.19**	0.02	**0.78**	1.00
Panel D: Hispanic Americans							
	Wakeup	**Bedtime**	**CAR**	**AUC**	**EDSlope**	**LDSlope**	**ODSlope**
Wakeup	1.00						
Bedtime	**0.46**	1.00					
CAR	**-0.47**	**-0.16**	1.00				
AUC	**0.51**	**0.78**	-0.01	1.00			
EDSlope	**-0.18**	-0.02	**-0.30**	0.09	1.00		
LDSlope	**0.11**	**0.52**	-0.08	**0.37**	**-0.42**	1.00	
ODSlope	**-0.29**	**0.47**	**0.21**	**0.48**	-0.02	**0.74**	1.00

Bold = $p<0.05$

Given that the previous literature has shown associations between cortisol features and age, sex, socioeconomic status, and race/ethnicity, each of these characteristics was examined across the seven cortisol features used in this analysis. First the univariate associations between age, sex, and education were examined (Table 8). We used an individual's highest completed education level, which was assessed at the baseline MESA exam, as a proxy for socioeconomic status. With the ethnic groups combined, age at baseline is a significant ($p<0.05$) predictor for five of the seven cortisol features. Sex was a significant predictor for only AUC and EDSlope, with males having lower mean concentrations compared to females. Education was a significant ($p<0.05$) predictor of three cortisol features, Wakeup, Bedtime, and EDSlope.

In examining the associations between the different ethnic groups and the cortisol features (Table 9), the effects for African Americans and Hispanic Americans are presented relative to the European Americans (reference group). Overall, there is a significant difference between European Americans and the minority racial/ethnic groups with the cortisol features. Given the strength of many of the associations between racial/ethnic group and the cortisol features, the analyses for this dissertation will be racial/ethnic group specific.

In examining the multivariable associations of age and sex with the cortisol features stratified by race (Table 10), age is a significant ($p<0.05$) predictor for nearly every feature across the ethnic groups. Sex was less commonly a significant predictor in these multivariable models. Education was not a significant predictor for any of the cortisol features. While previous work has shown that cortisol concentrations vary based

on oral contraceptive use, among the MESA Stress participants there were only seven reports of oral contraceptive or hormone replacement therapy use. Therefore, use of hormones does not substantially contribute as a source of variation.

Table 8: Univariate associations with cortisol features.

Cortisol Feature	N	Age		Sex		Education	
		β	p-value	β	p-value	β	p-value
Wakeup	830	0.009	<0.0001	-0.020	0.62	0.022	0.009
Bedtime	826	0.023	<0.0001	0.027	0.64	0.025	0.04
CAR	775	-0.003	0.08	0.062	0.08	0.003	0.70
AUC	817	0.012	<0.0001	-0.106	0.002	0.014	0.05
EDSlope	805	0.001	0.70	-0.120	<0.0001	-0.013	0.03
LDSlope	822	0.001	0.002	0.002	0.66	0.0008	0.41
ODSlope	835	0.001	0.003	-0.008	0.05	-0.0002	0.82

Table 9: Associations of race with cortisol features.

Race[a]	Wakeup		Bedtime		CAR		AUC	
	β	p-value	β	p-value	β	p-value	β	p-value
AFA	-0.201	0.0006	0.207	0.01	-0.094	0.07	-0.040	0.41
HIS	-0.207	<0.0001	-0.283	0.0001	-0.073	0.11	-0.182	<0.0001
	EDSlope		**LDSlope**		**ODSlope**			
	β	p-value	β	p-value	β	p-value		
AFA	0.112	0.01	0.015	0.02	0.016	0.01		
HIS	0.134	0.0006	-0.012	0.04	-0.007	0.20		

a. European American reference group.

Table 10: Racial/ethnic group stratified multivariable associations of age, sex, and education with cortisol features.

Cortisol Feature	N	Age		Sex		Education		R^2
		β	p-value	β	p-value	β	p-value	
African Americans								
Wakeup	214	0.009	**0.02**	0.017	0.83	0.007	0.74	0.025
Bedtime	212	0.011	**0.04**	0.005	0.96	0.001	0.70	0.021
CAR	203	-0.0009	0.79	0.102	0.12	-0.010	0.56	0.014
AUC	209	0.009	**0.003**	-0.060	0.30	-0.011	0.46	0.058
EDSlope	209	-0.004	0.26	-0.196	**0.001**	0.000002	0.99	0.057
LDSlope	211	0.0006	0.21	0.010	0.26	-0.0001	0.96	0.015
ODSlope	214	0.00006	0.88	-0.005	0.52	-0.002	0.38	0.006
European Americans								
Wakeup	166	0.002	0.58	0.0002	0.99	-0.005	0.81	0.003
Bedtime	166	0.023	**<0.0001**	0.089	0.45	0.011	0.69	0.095
CAR	160	-0.004	0.31	0.069	0.35	0.021	0.25	0.022
AUC	166	0.006	0.07	-0.102	0.13	0.017	0.29	0.041
EDSlope	163	0.006	**0.02**	-0.113	**0.04**	0.001	0.94	0.059
LDSlope	164	0.0003	0.55	-0.013	0.20	-0.0007	0.79	0.013
ODSlope	169	0.0007	0.24	-0.019	0.09	0.0001	0.96	0.027
Hispanic Americans								
Wakeup	450	0.013	**<0.0001**	-0.035	0.52	0.024	0.08	0.050
Bedtime	448	0.028	**<0.0001**	-0.016	0.83	-0.007	0.72	0.107
CAR	412	-0.005	0.08	0.040	0.45	-0.011	0.38	0.011
AUC	442	0.015	**<0.0001**	-0.139	**0.003**	-0.003	0.78	0.102
EDSlope	433	0.001	0.58	-0.091	**0.03**	-0.011	0.31	0.013
LDSlope	447	0.0009	**0.006**	0.001	0.82	-0.001	0.46	0.019
ODSlope	452	0.0008	**0.004**	-0.009	0.11	-0.003	0.05	0.033

Bold = $p<0.05$

Genetic Data

Genotype data included both measured and imputed SNPs available through participation in MESA SHARe (SNP Health Association Resource) project. Under the SHARe project, genome-wide genotyping was obtained using the Affymetrix Genome-Wide Human SNP Array 6.0 platform. Imputation to HapMap was completed at the MESA Genetics Centers using the IMPUTE2 [82] program with the following reference panels: the HapMap Phase I and II, the human genome reference sequence (NCBI Build 36). The HapMap project is based on ethnic specific reference panels, composed of the following groups: Yoruba in Ibadan, Nigeria (abbreviation: YRI), Japanese in Tokyo, Japan (abbreviation: JPT), Han Chinese in Beijing, China (abbreviation: CHB), CEPH (Utah residents with ancestry from northern and western Europe) (abbreviation: CEU). Imputation for African Americans and Hispanic Americans was performed using the CEU+YRI+CHB+JPT reference panels (release #22). Imputation for European Americans was performed using only the CEU reference panel (release #24). All imputed and genotyped SNPs were aligned to the "+" strand of the human genome reference sequence (NCBI Build 36).

In order to account for population structure and admixture within MESA samples, principal components were extracted from genome-wide data from MESA Classic participants, in each ethnic group separately. The multivariable associations of the top 10 principal components for each cortisol feature, stratified by ethnic group, are presented in Tables 11-17. There was limited evidence of association with the principal components in predicting the cortisol features among European Americans. As there was only

evidence for a few principal components on Bedtime, we did not adjust for principal components in European Americans for the genome-wide association analyses. There were a number of significant ($p<0.05$) principal components for African Americans and Hispanic Americans. These associations indicate that there is underlying population structure beyond what is being captured by stratifying by race/ethnicity. Given this evidence, we adjusted for the top 10 principal components for the African Americans and Hispanic Americans.

Table 11: Multivariable associations of the top 10 principal components in each ethnic group on Wakeup.

	Wakeup					
Principal Component	**AFA (n=214)**		**EA (n=166)**		**HIS (n=450)**	
	β (SE)	**p-value**	**β (SE)**	**p-value**	**β (SE)**	**p-value**
PC1	2.87 (1.50)	0.057	-1.97 (1.91)	0.30	-1.19 (1.01)	0.24
PC2	-4.95 (1.59)	**0.002**	1.70 (1.61)	0.29	-1.30 (1.06)	0.22
PC3	-0.05 (1.63)	0.98	-1.99 (3.01)	0.51	0.41 (0.65)	0.53
PC4	1.54 (1.97)	0.44	-0.86 (2.87)	0.77	-1.29 (1.87)	0.49
PC5	-3.54 (2.29)	0.12	3.44 (2.83)	0.23	0.76 (0.92)	0.41
PC6	2.53 (1.91)	0.19	1.49 (2.69)	0.58	0.94 (1.42)	0.51
PC7	-2.56 (1.80)	0.16	4.59 (2.68)	0.09	0.72 (1.94)	0.71
PC8	1.34 (2.37)	0.57	5.20 (5.25)	0.32	0.83 (0.98)	0.39
PC9	-0.95 (2.08)	0.68	3.68 (4.39)	0.40	-0.20 (1.19)	0.86
PC10	-0.23 (2.26)	0.91	4.94 (4.25)	0.25	-0.48 (0.75)	0.52
	Model R^2	0.10	**Model R^2**	0.05	**Model R^2**	0.03

AFA = African Americans. EA = European Americans. HIS = Hispanic Americans. PC = Principal Component. Bold = $p<0.05$

Table 12: Multivariable associations of the top 10 principal components in each ethnic group on Bedtime.

Principal Component	Bedtime					
	AFA (n=212)		EA (n=166)		HIS (n=448)	
	β (SE)	p-value	β (SE)	p-value	β (SE)	p-value
PC1	2.68 (2.05)	0.19	1.07 (2.69)	0.69	2.74 (1.45)	0.06
PC2	-5.92 (2.17)	**0.007**	-0.69 (2.27)	0.76	-1.57 (1.52)	0.30
PC3	1.43 (2.23)	0.52	3.39 (4.10)	0.41	-0.45 (0.93)	0.63
PC4	3.29 (2.69)	0.22	6.72 (4.02)	0.10	-1.21 (2.70)	0.65
PC5	-4.45 (3.11)	0.15	-1.51 (3.97)	0.70	1.26 (1.33)	0.34
PC6	-0.79 (2.60)	0.76	-1.60 (3.78)	0.67	2.34 (2.04)	0.25
PC7	-0.47 (2.47)	0.85	9.95 (3.67)	**0.008**	2.67 (2.78)	0.34
PC8	-2.57 (3.23)	0.42	21.80 (7.40)	**0.004**	0.12 (1.41)	0.93
PC9	1.09 (2.84)	0.70	3.74 (6.17)	0.55	0.16 (1.71)	0.92
PC10	-0.03 (3.10)	0.99	5.02 (5.93)	0.40	-0.82 (1.08)	0.45
	Model R^2	0.07	**Model R^2**	0.09	**Model R^2**	0.03

AFA = African Americans. EA = European Americans. HIS = Hispanic Americans. PC = Principal Component. Bold = $p<0.05$

Table 13: Multivariable associations of the top 10 principal components in each ethnic group on CAR.

Principal Component	CAR					
	AFA (n=203)		EA (n=160)		HIS (n=412)	
	β (SE)	p-value	β (SE)	p-value	β (SE)	p-value
PC1	0.82 (1.30)	0.53	0.20 (1.69)	0.90	-0.08 (0.95)	0.93
PC2	1.56 (1.51)	0.30	-1.00 (1.43)	0.48	0.72 (1.01)	0.47
PC3	1.22 (1.41)	0.39	1.28 (2.66)	0.63	0.20 (0.60)	0.74
PC4	-0.04 (1.69)	0.98	2.46 (2.60)	0.35	1.98 (1.76)	0.26
PC5	2.03 (2.03)	0.32	-3.12 (2.47)	0.21	-0.59 (0.84)	0.48
PC6	-0.75 (1.69)	0.66	-0.07 (2.38)	0.98	-1.82 (1.31)	0.17
PC7	1.67 (1.58)	0.29	-0.46 (2.37)	0.85	1.52 (1.77)	0.39
PC8	-1.33 (2.03)	0.51	-3.86 (4.65)	0.41	-0.29 (0.90)	0.75
PC9	-0.05 (1.85)	0.98	0.96 (3.92)	0.81	0.36 (1.09)	0.74
PC10	-3.49 (1.95)	0.07	-2.04 (3.80)	0.59	0.63 (0.69)	0.36
	Model R^2	0.05	**Model R^2**	0.04	**Model R^2**	0.02

AFA = African Americans. EA = European Americans. HIS = Hispanic Americans. PC = Principal Component. Bold = $p<0.05$

Table 14: Multivariable associations of the top 10 principal components in each ethnic group on AUC.

Principal Component	AUC					
	AFA (n=209)		EA (n=166)		HIS (n=442)	
	β (SE)	p-value	β (SE)	p-value	β (SE)	p-value
PC1	2.19 (1.18)	0.06	-0.62 (1.54)	0.69	1.80 (.89)	**0.04**
PC2	-1.98 (1.24)	0.11	-0.49 (1.30)	0.71	-0.53 (0.93)	0.57
PC3	0.34 (1.28)	0.79	0.67 (2.35)	0.78	-0.43 (0.57)	0.45
PC4	1.22 (1.54)	0.43	2.77 (2.30)	0.23	0.61 (1.66)	0.71
PC5	-2.87 (1.82)	0.12	-1.41 (2.27)	0.54	-0.55 (0.81)	0.49
PC6	-1.35 (1.50)	0.37	-0.68 (2.17)	0.76	1.09 (1.25)	0.39
PC7	-1.27 (1.43)	0.37	1.83 (2.11)	0.39	2.79 (1.69)	0.10
PC8	-1.41 (1.86)	0.45	3.05 (4.24)	0.47	0.39 (0.86)	0.65
PC9	0.80 (1.64)	0.62	1.25 (3.54)	0.72	-0.20 (1.05)	0.85
PC10	0.82 (1.79)	0.65	1.15 (3.40)	0.73	0.25 (0.66)	0.70
	Model R^2	0.06	**Model R^2**	0.02	**Model R^2**	0.03

AFA = African Americans. EA = European Americans. HIS = Hispanic Americans. PC = Principal Component. Bold = $p<0.05$

Table 15: Multivariable associations of the top 10 principal components in each ethnic group on EDSlope.

Principal Component	EDSlope					
	AFA (n=209)		EA (n=163)		HIS (n=433)	
	β (SE)	p-value	β (SE)	p-value	β (SE)	p-value
PC1	-0.91 (1.23)	0.45	0.19 (1.25)	0.88	0.90 (0.77)	0.24
PC2	4.01 (1.30)	**0.002**	-1.75 (1.06)	0.10	0.93 (0.82)	0.26
PC3	-0.20 (1.33)	0.88	-0.18 (1.97)	0.93	-0.99 (0.49)	0.05
PC4	-0.36 (1.61)	0.83	1.75 (1.94)	0.37	0.58 (1.44)	0.69
PC5	0.78 (1.87)	0.68	0.75 (1.84)	0.68	-0.73 (0.70)	0.30
PC6	-2.77 (1.57)	0.08	-3.40 (1.76)	0.05	2.19 (1.09)	**0.04**
PC7	-0.09 (1.47)	0.95	-2.47 (1.71)	0.15	-2.02 (1.46)	0.19
PC8	0.37 (1.93)	0.85	3.28 (3.46)	0.34	-0.44 (0.74)	0.56
PC9	0.63 (1.72)	0.71	-5.66 (2.87)	0.05	-1.00 (0.91)	0.27
PC10	2.08 (1.84)	0.26	-4.11 (2.82)	0.15	0.07 (0.57)	0.91
	Model R^2	0.07	**Model R^2**	0.07	**Model R^2**	0.03

AFA = African Americans. EA = European Americans. HIS = Hispanic Americans. PC = Principal Component. Bold = $p<0.05$

Table 16: Multivariable associations of the top 10 principal components in each ethnic group on LDSlope.

	LDSlope					
Principal Component	**AFA (n=211)**		**EA (n=164)**		**HIS (n=447)**	
	β (SE)	**p-value**	**β (SE)**	**p-value**	**β (SE)**	**p-value**
PC1	-0.05 (0.18)	0.77	-0.17 (0.23)	0.46	0.17 (0.11)	0.12
PC2	-0.35 (0.19)	0.07	0.23 (0.19)	0.23	-0.02 (0.12)	0.87
PC3	0.01 (0.20)	0.95	0.03 (0.36)	0.94	0.01 (0.07)	0.91
PC4	0.06 (0.24)	0.79	-0.28 (0.34)	0.41	-0.20 (0.21)	0.35
PC5	0.23 (0.28)	0.41	-0.50 (0.34)	0.13	-0.04 (0.10)	0.72
PC6	0.15 (0.23)	0.51	0.70 (0.32)	**0.03**	0.01 (0.16)	0.94
PC7	-0.12 (0.22)	0.59	-0.12 (0.31)	0.71	0.31 (0.21)	0.15
PC8	-0.21 (0.29)	0.47	-0.43 (0.63)	0.49	0.00 (0.11)	0.99
PC9	0.05 (0.25)	0.83	0.15 (0.52)	0.78	0.11 (0.13)	0.43
PC10	0.02 (0.28)	0.94	0.41 (0.51)	0.43	0.01 (0.08)	0.92
	Model R^2	0.03	**Model R^2**	0.07	**Model R^2**	0.02

AFA = African Americans. EA = European Americans. HIS = Hispanic Americans. PC = Principal Component. Bold = $p<0.05$

Table 17: Multivariable associations of the top 10 principal components in each ethnic group on ODSlope.

	ODSlope					
Principal Component	**AFA (n=214)**		**EA (n=169)**		**HIS (n=452)**	
	β (SE)	**p-value**	**β (SE)**	**p-value**	**β (SE)**	**p-value**
PC1	-0.04 (0.15)	0.81	0.00 (0.25)	0.99	0.34 (0.10)	**0.0005**
PC2	0.04 (0.16)	0.81	0.11 (0.21)	0.59	-0.05 (0.10)	0.61
PC3	-0.06 (0.17)	0.74	-0.54 (0.39)	0.17	-0.06 (0.06)	0.33
PC4	0.02 (0.20)	0.92	-0.39 (0.38)	0.31	0.08 (0.18)	0.67
PC5	0.12 (0.24)	0.62	-0.42 (0.38)	0.27	-0.15 (0.10)	0.10
PC6	-0.20 (0.20)	0.30	0.02 (0.36)	0.95	0.03 (0.14)	0.83
PC7	0.00 (0.19)	0.99	-0.45 (0.35)	0.20	0.28 (0.19)	0.13
PC8	-0.45 (0.24)	0.07	-0.07 (0.70)	0.92	-0.05 (0.10)	0.60
PC9	0.34 (0.22)	0.11	-0.04 (0.58)	0.94	0.03 (0.12)	0.83
PC10	0.29 (0.23)	0.21	-0.86 (0.56)	0.13	0.09 (0.07)	0.23
	Model R^2	0.05	**Model R^2**	0.06	**Model R^2**	0.05

AFA = African Americans. EA = European Americans. HIS = Hispanic Americans. PC = Principal Component. Bold = $p<0.05$

Chronic Disease Risk Factors

Body Mass Index

A variety of anthropometric measurements, including height, weight, hip circumference, and waist circumference, are measured at each MESA examination using standardized instruments and procedures. For this dissertation research, the anthropometric variable of interest is body mass index (BMI), calculated as weight in kilograms/height in meters squared. We chose BMI as the anthropometric variable of interest since HPA axis dysfunction has been previously linked to abdominal obesity [9]. We did not examine waist circumference in addition to BMI as previous MESA work did not find associations between waist circumference and multiple cortisol features (CAR, overall decline, and AUC) [83]. Given the repeat assessment of BMI at multiple MESA exams, an averaged measure of BMI was created. Averaged BMI ranged from 15 to 55 kg/m^2, with a mean of 29 kg/m^2 (standard deviation of 5.6 kg/m^2) (Figure 8). In examining the distribution of BMI stratified by the different racial/ethnic groups (Figure 9, Table 18), the means are statistically different from each other ($p<0.0001$).

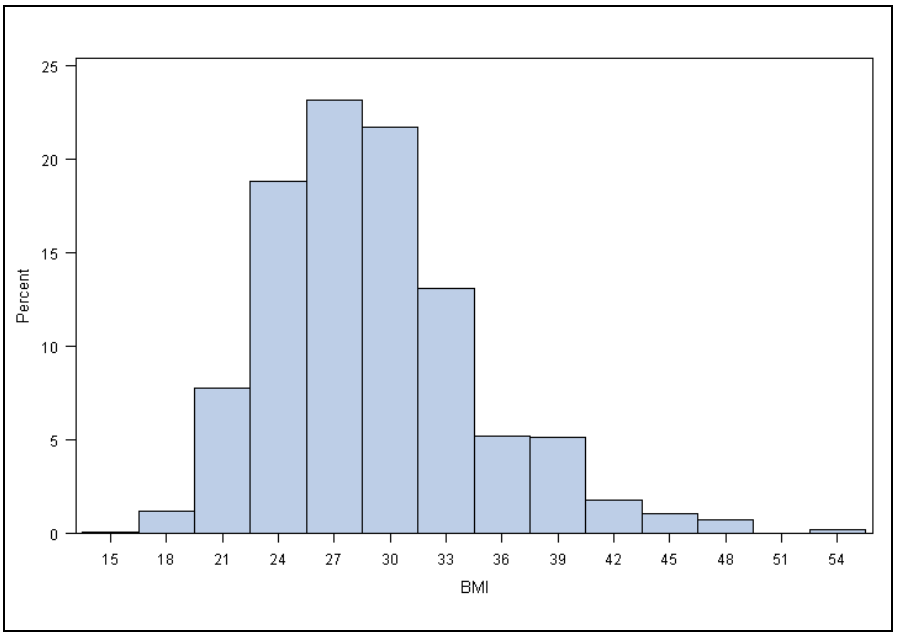


Figure 8: Distribution of averaged BMI in kg/m^2.

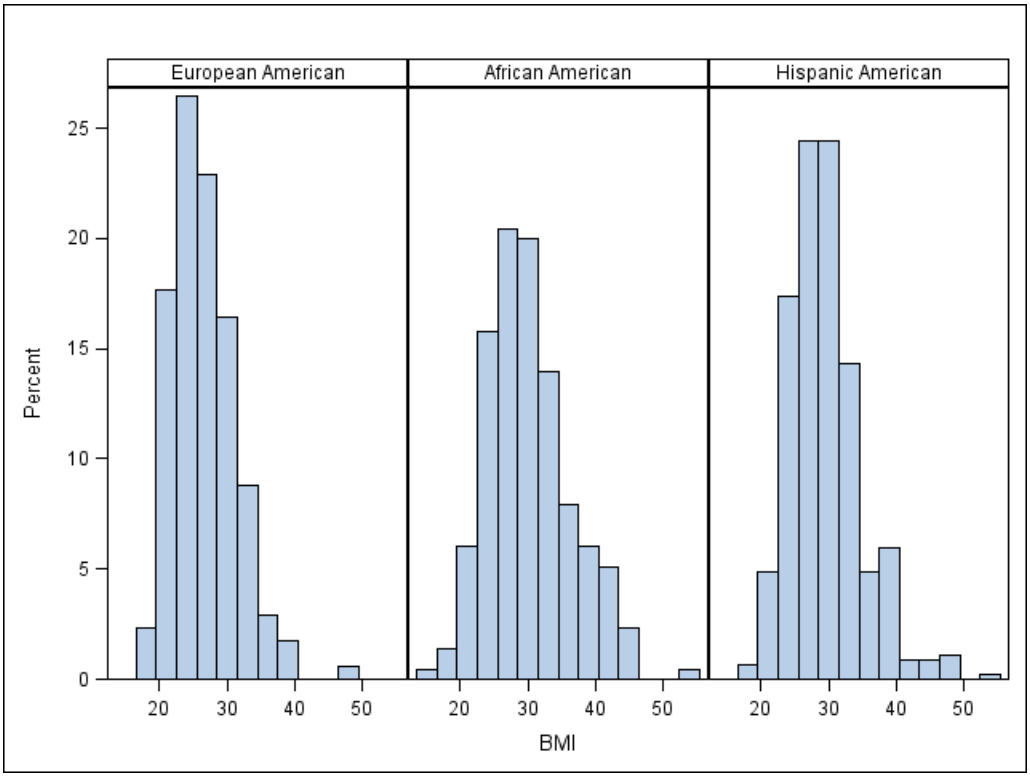


Figure 9: Distribution of averaged BMI in kg/m^2, by racial/ethnic group.

Table 18: Ethnic group specific distributions of chronic disease risk factors.

	European Americans				**African Americans**				**Hispanic Americans**				**ANOVA**
	N	Min	Max	Mean (SD)	N	Min	Max	Mean (SD)	N	Min	Max	Mean (SD)	p-value
BMI	170	18.2	47.0	26.6 (4.5)	215	15.3	54.0	30.1 (6.3)	454	18.7	54.9	29.5 (5.3)	<0.0001
ln(Glucose)	170	4.2	5.1	4.5 (0.1)	213	4.3	5.4	4.6 (0.2)	452	4.3	5.6	4.6 (0.2)	<0.0001
ln(IL-6 + 1)	166	0.5	2.5	1.1 (0.4)	205	0.4	2.5	1.2 (0.5)	434	0.4	2.6	1.3 (0.4)	0.009
ln(TNF-α + 1)	167	0.1	3.5	1.4 (0.6)	214	0.1	3.7	1.4 (0.6)	440	0.1	3.1	1.5 (0.6)	0.19

Fasting Glucose

The MESA Study collects a range of variables capturing metabolic parameters. For this dissertation, fasting glucose is the feature of interest. Fasting glucose concentrations were chosen rather than diabetes status due to the additional information and power in analysis of a continuous trait compared to the dichotomous trait derived from a continuous distribution. Fasting glucose concentrations were measured in mg/dl. Given the repeat assessment of fasting glucose at multiple MESA exams, an averaged measure of fasting glucose concentrations was created. Averaged fasting glucose ranged from 30 to 301 mg/dl, with a mean of 103 mg/dl (standard deviation of 30.2 mg/dl) (Figure 10).

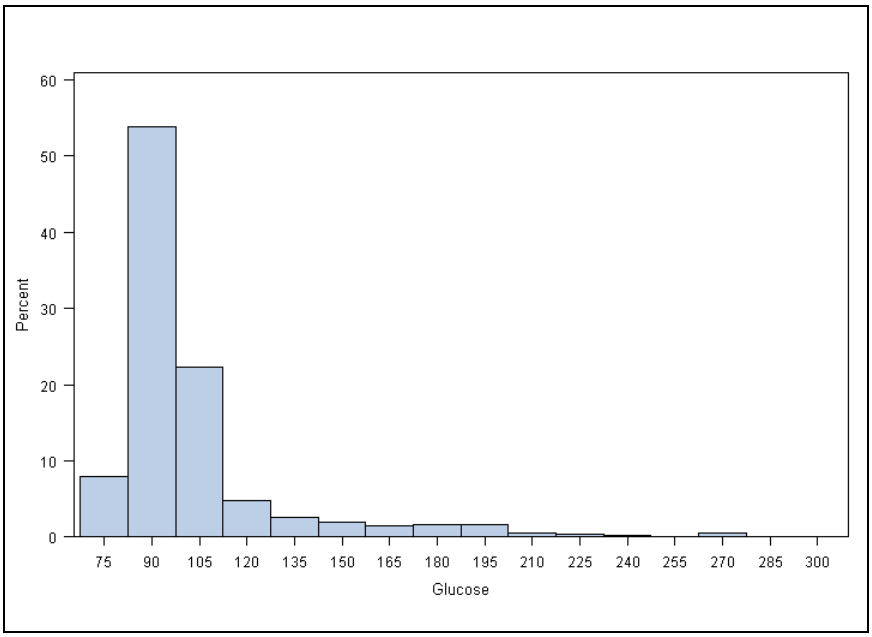


Figure 10: Distribution of averaged fasting glucose concentrations in mg/dl.

Given the extreme right tail for averaged fasting glucose concentrations, we natural log-transformed this variable. After transformation ln(Glucose) ranged from 4.25 to 5.71 mg/dl, with a mean of 4.60 mg/dl (standard deviation of 0.22 mg/dl) (Figure 11). In examining the distribution of ln(Glucose) stratified by the different racial/ethnic groups (Figure 12, Table 18), the means are statistically different from each other ($p<0.0001$).

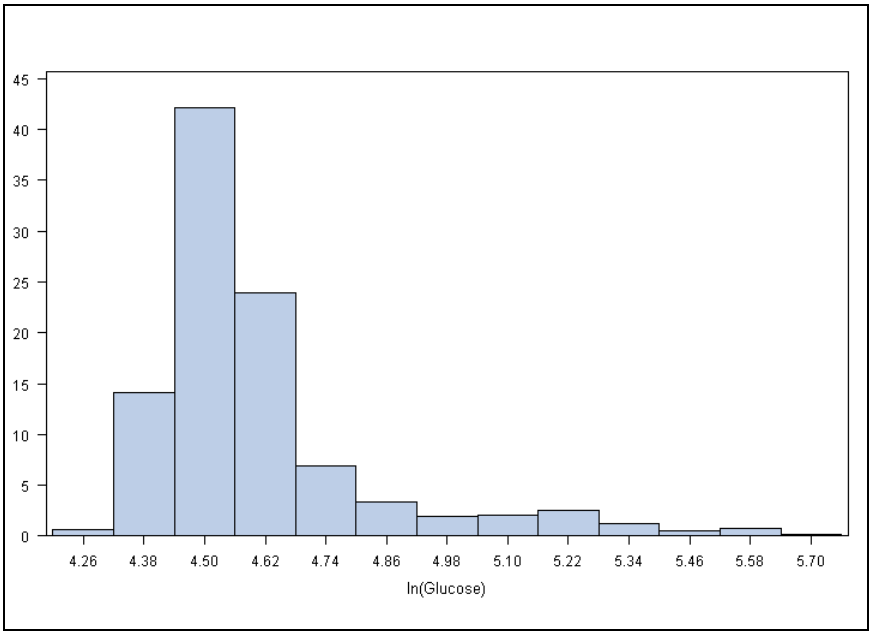


Figure 11: Distribution of natural log-transformed average fasting glucose concentrations.

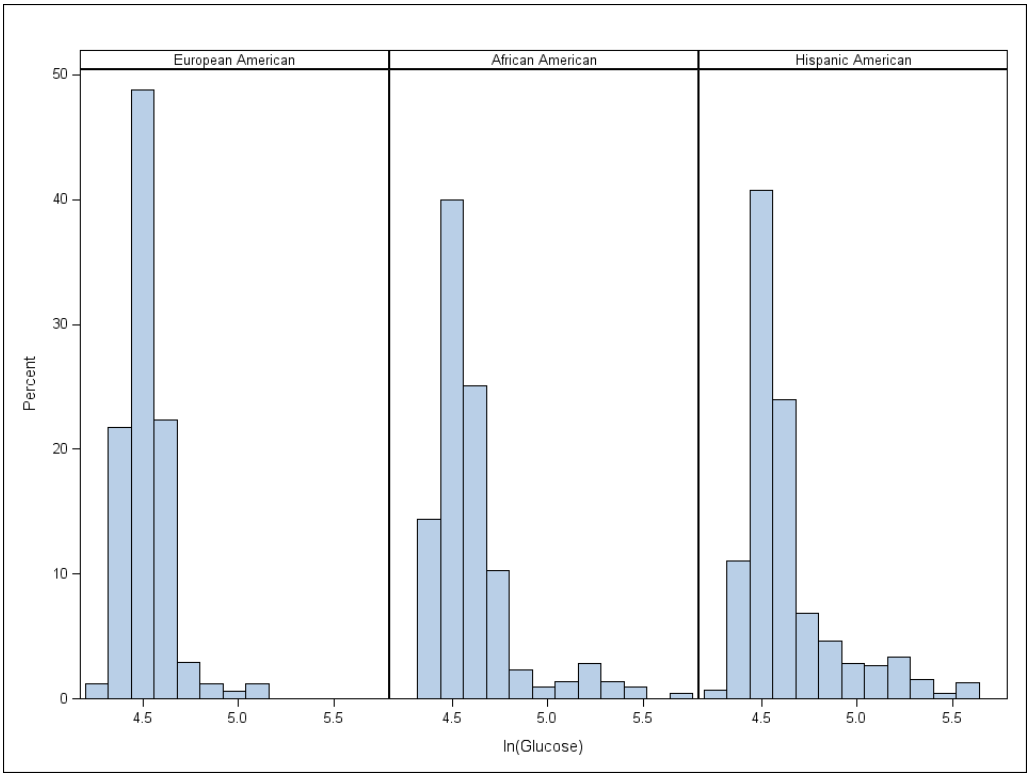


Figure 12: Distribution of ln(Glucose) concentrations, by racial/ethnic group.

Among the African American and Hispanic American histograms in Figure 12, there are long right side tails. In examining the data for outliers, we plotted ln(Glucose) concentrations against age (Figure 13). It was noted that there were a large number of points where ln(Glucose) concentrations were greater than 5.0. In examining for outliers there were 21 observations which had a studentized residual (r) with an absolute value greater than 3. These large residuals corresponded to ln(Glucose) concentrations that ranged from 5.28 to 5.71. In evaluating for high leverage points, there were 37 observations which had leverage estimates greater than 0.0047 (threshold determined by (2k+2)/n).

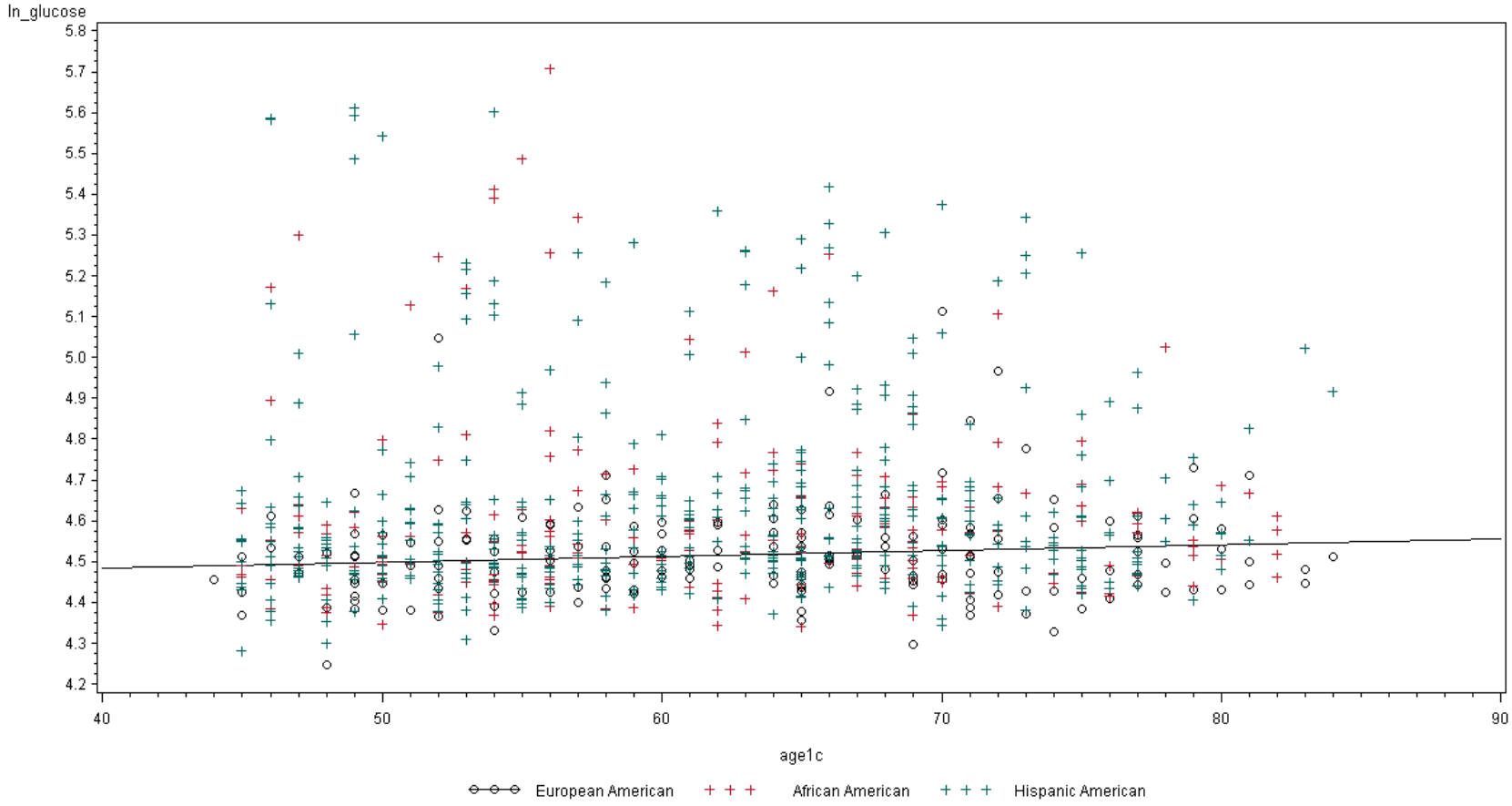


Figure 13: Scatter plot of natural log-transformed average fasting glucose concentrations versus age, by racial/ethnic group.

None of the data points with large residuals were high leverage points. We decided to exclude individuals with ln(Glucose) concentrations that were greater than ±4 standard deviations away from the mean, which excluded two African Americans and two Hispanic Americans.

Inflammatory Measures

The MESA Study has data available on a range of markers of inflammation. This dissertation specifically focuses on two inflammatory measures, interleukin-6 (IL-6) and tumor necrosis factor-α (TNF-α), in an effort to expand on previous work in the MESA Study that showed associations between these measures and cortisol features [12]. Both IL-6 and TNF-α were measured from fasting blood draws at the MESA Exam that corresponds to an individual's cortisol collection.

Serum from the blood draws was frozen then shipped and stored at the Central Blood Analysis Laboratory at the University of Vermont. IL-6 was measured by an IL-6 assay (Quantikine HS Human IL-6 Immunoassay; R&D Systems, Minneapolis, MN), with an average coefficient of variation of 6.3%. TNF-α was measured using the LINCOplex Human Cardiovascular Disease Panel 3 Kit (Millipore Corporation, St. Charles, MO), with an average coefficient of variation of 10.3%. The unit of measurement for both markers of inflammation was pg/ml. Both assays were completed at the Laboratory for Clinical Biochemistry Research at the University of Vermont.

IL-6 concentrations ranged from 0.45 to 12.19 pg/ml, with a mean of 2.81 pg/ml (standard deviation 2.00 pg/ml) (Figure 14). Given the extreme right tail for IL-6

concentrations, we transformed this variable as ln(IL-6 + 1) (Figure 15). ln(IL-6 + 1) concentrations ranged from 0.37 to 2.58 pg/ml, and averaged 1.23 pg/ml (standard deviation 0.43 pg/ml). In examining the distribution of ln(IL-6 + 1) stratified by the different racial/ethnic groups (Figure 16, Table 18), the means are statistically different from each other (p=0.009).

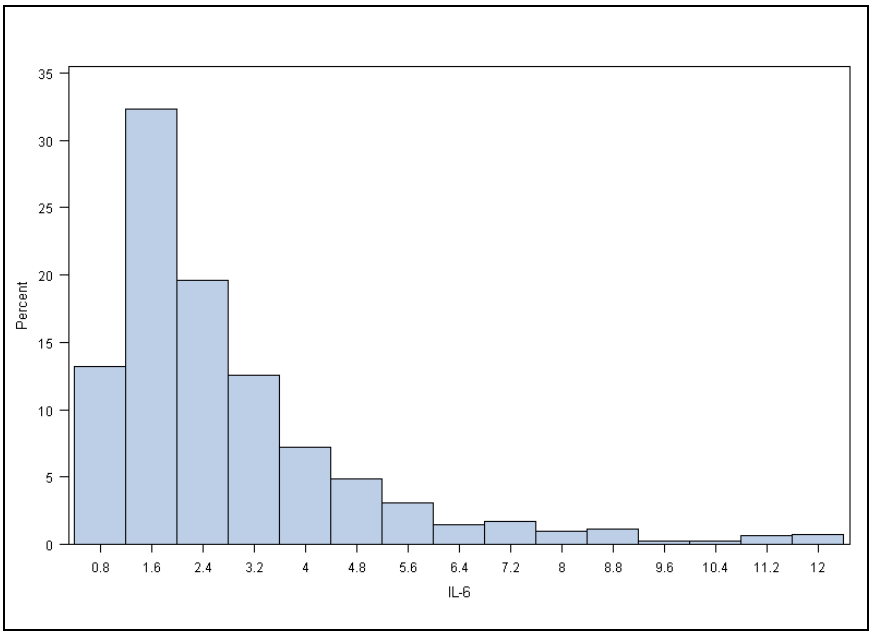


Figure 14: Distribution of IL-6 concentrations in pg/ml.

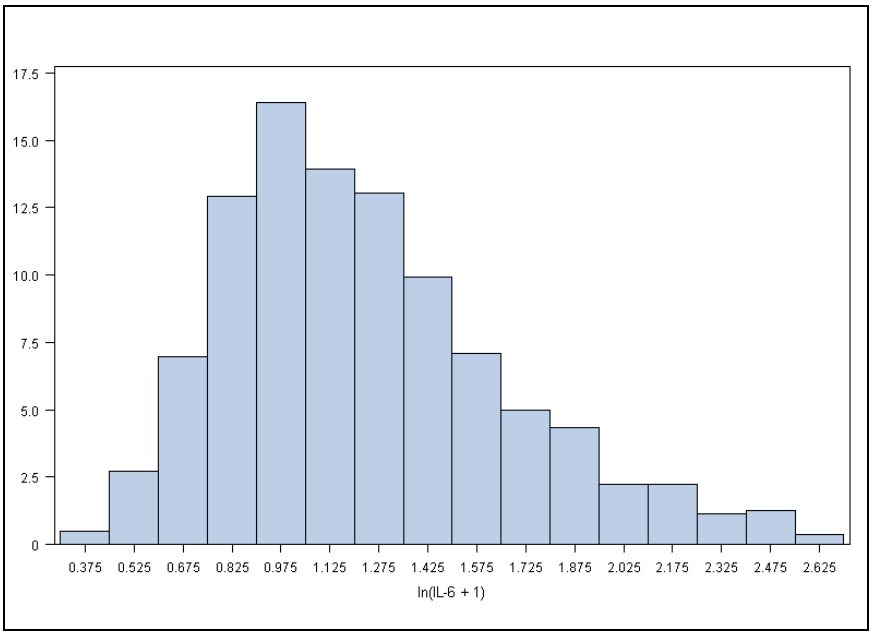


Figure 15: Distribution of ln(IL-6 + 1) in pg/ml.

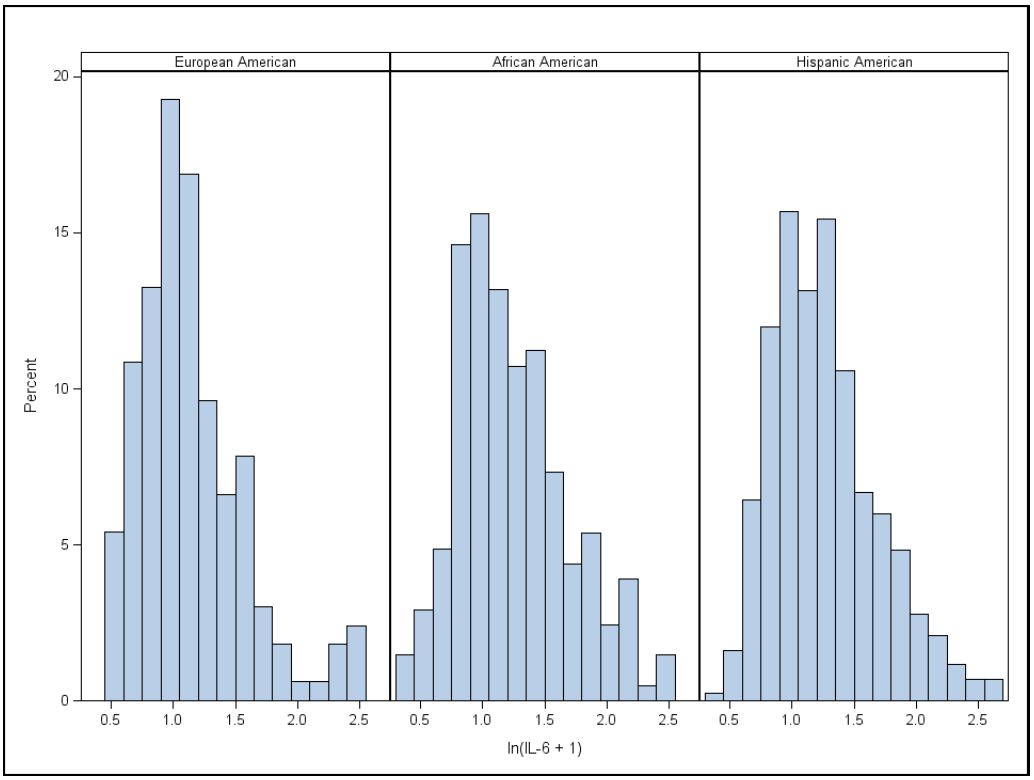


Figure 16: Distribution of ln(IL-6 + 1) concentrations, by racial/ethnic group.

TNF-α concentrations ranged from 0.060 to 140.89 pg/ml, with a mean of 4.38 pg/ml (standard deviation 6.61 pg/ml) (Figure 17). Given the extreme right tail for TNF-α concentrations, the two outlying observations (at concentrations of 90 and 140 pg/ml) were removed and we transformed this variable as ln(TNF-α + 1) (Figure 18). ln(TNF-α + 1) concentrations ranged from 0.058 to 3.68 pg/ml, and averaged 1.46 pg/ml (standard deviation 0.59 pg/ml). In examining the distribution of ln(TNF-α + 1) stratified by the different racial/ethnic groups (Figure 19, Table 18), the means are statistically different from each other (p=0.009).

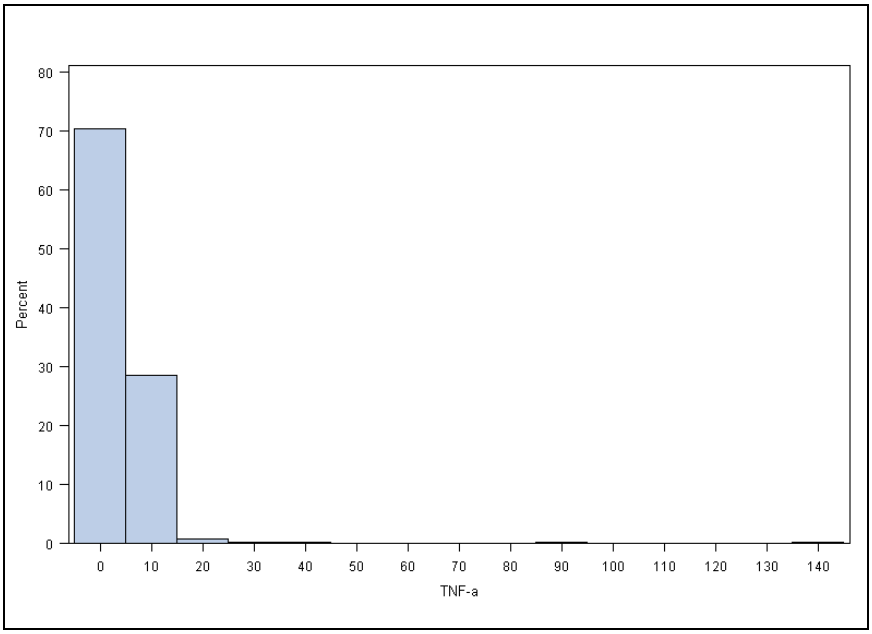


Figure 17: Distribution of TNF-α in pg/ml.

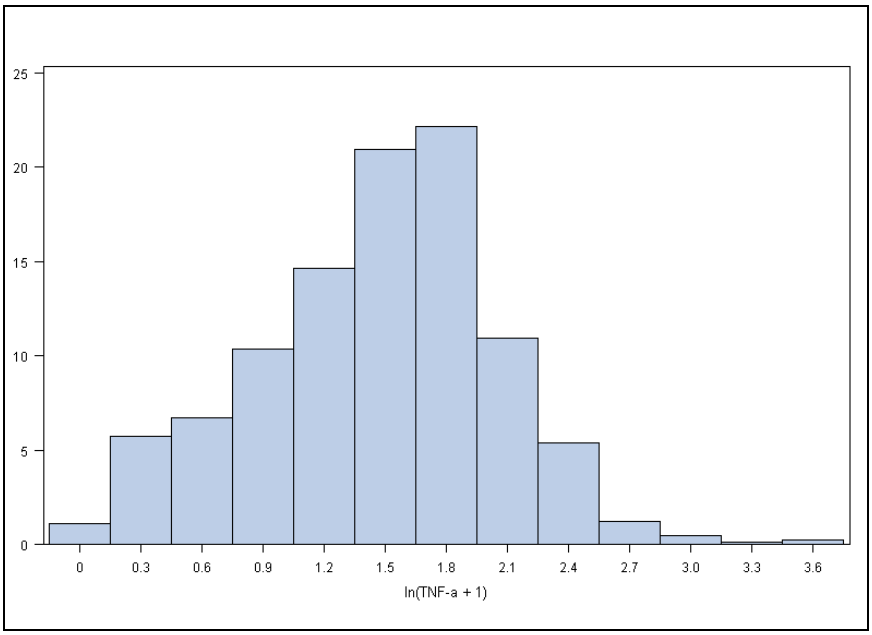


Figure 18: Distribution of ln(TNF-α + 1)in pg/ml.

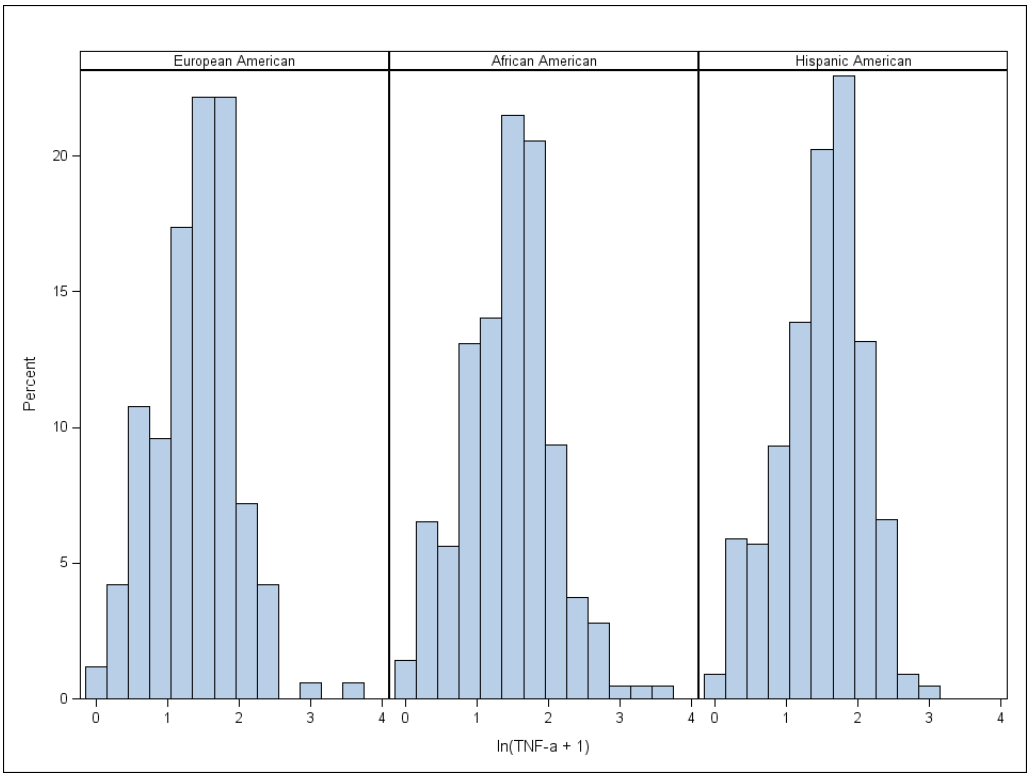


Figure 19: Distribution of ln(TNF-α + 1) concentrations, by racial/ethnic group.

The univariate association of covariates and cortisol summary features on all four chronic disease risk factor outcomes is provided in Table 19. The associations between age, sex, and African American race and the chronic disease risk factor outcomes are not consistent with respect to direction of effect. Nor are the associations consistent between many of the cortisol features and the chronic disease risk factor outcomes with respect to direction of effect. For BMI, every covariate is a significant predictor ($p<0.05$) with the exception of four cortisol features. Fewer of the covariates are significantly predictive of ln(Glucose) and ln(IL-6 +1). Only age was a significant predictor in the univariate associations with ln(TNF-α +1).

Table 19: The univariate associations of covariates and cortisol features on average BMI, ln(Glucose), ln(IL6 + 1), and ln(TNF-α +1).

	BMI		**Glucose**		**IL6**		**TNF-α**	
	β	**p-value**	**β**	**p-value**	**β**	**p-value**	**β**	**p-value**
Age	-0.07	0.0003	0.0008	0.29	0.01	<0.0001	0.01	<0.0001
Sex	-1.44	0.0002	0.04	0.01	-0.03	0.30	-0.03	0.46
AFA	1.41	0.001	-0.007	0.70	0.01	0.68	-0.02	0.66
HIS	0.97	0.01	0.07	<0.0001	0..06	0.05	0.07	0.10
Education	-0.37	<0.0001	-0.02	<0.0001	-0.03	<0.0001	-0.01	0.38
Wakeup	-1.76	<0.0001	-0.03	0.01	-0.06	0.03	-0.03	0.45
Bedtime	-0.29	0.22	-0.004	0.64	0.06	0.002	0.02	0.41
CAR	-0.29	0.48	-0.02	0.16	-0.02	0.50	-0.0004	0.93
AUC	-1.11	0.006	-0.01	0.46	0.10	0.002	0.04	0.34
EDSlope	1.37	0.003	0.06	0.0005	0.15	<0.0001	0.09	0.07
LDSlope	-0.53	0.86	-0.11	0.34	0.63	0.008	0.31	0.34
ODSlope	6.01	0.06	0.10	0.40	1.32	<0.0001	0.47	0.16

In the next chapter, I begin the discussion of the individual studies that comprise this dissertation. I detail the investigation of the gene-level associations between stress responses candidate genes and each of the seven cortisol features. I also present the results from the gene-level meta-analysis, which allows for comparison of the ethnic group specific analyses.

CHAPTER 3

VARIATION IN STRESS RESPONSE GENES IS RELATED TO FEATURES OF DIURNAL CORTISOL CURVES IN THE MULTI-ETHNIC STUDY OF ATHEROSCLEROSIS

Introduction

Cortisol concentrations follow a strong daily pattern. They are high upon awakening, reach a maximum concentration approximately half an hour later, and slowly decrease throughout the rest of the day [1-3]. Additionally, cortisol concentrations increase in response to stressful situations, such as public speaking [32]. Under conditions of chronic stress, prolonged increased concentrations could have detrimental downstream physiological effects.

Several population-based studies have linked daily cortisol patterns to health outcomes, including elevated blood pressure, abdominal obesity, and coronary calcification [8-10]. Cortisol concentrations and various features of the cortisol daily profile have also been linked to diabetes mellitus [11] and markers of inflammation [12].

Despite evidence of associations of various risk factors with cortisol, considerable inter-individual variability in cortisol remains unexplained. This has led to increased interest in examining genetic predictors of cortisol phenotypes [15], as genetic factors could contribute to unexplained variability in cortisol concentrations. Most genetic research on

cortisol to date has focused on candidate gene associations, notably the glucocorticoid receptor gene (*NR3C1*) and the mineralocorticoid receptor gene (*NR3C2*) [21-23].

The increase in cortisol concentrations in response to a stressor [32] occurs through the activation of the hypothalamic-pituitary-adrenal (HPA) axis, which has downstream implications for the cardiovascular system, immune system, and metabolism [30, 31]. The cortisol metabolic pathway suggests several key genes whose variation could affect cortisol levels. We selected genes with downstream implications for either cortisol concentrations or cortisol responsivity. The six stress response genes of interest for this work include a glucocorticoid receptor gene (*NR3C1*), a mineralocorticoid receptor gene (*NR3C2*), the tyrosine hydroxylase gene (*TH*), the alpha-2A-adrenergic receptor gene (*ADRA2A*), the beta-2-adrenergic receptor gene (*ADRB2*), and the serotonin transporter gene (*SLC6A4*), all of which have been suggested to be involved in the physiologic response to psychological stressors [22, 52, 54, 56, 62]. However, few if any population based studies have investigated the associations between polymorphisms in these genes and cortisol levels in multiple ethnic groups.

In this study we investigate how variation in six stress response genes is related to diurnal cortisol features within and across ethnic groups utilizing a gene-level analysis approach, the sequence kernel association test (SKAT) [84]. When comparing single nucleotide polymorphism (SNP)-level results across multiple ethnic groups, difference in ethnic specific linkage disequilibrium structures may result in inconsistent findings. The gene-level analysis bypasses the problem that different tagging SNPs within gene regions may show association across ethnic groups. Since humans are 99% genetically similar,

gene structure (exon and intron organization) is not likely to differ across ethnic groups, making the assessment of entire genes a better analytic approach than individual SNPs. We performed a gene-level analysis for each of the six stress response gene regions (defined as all SNPs within the gene and 5 kilobases (kb) window up- and downstream of each gene) for multiple cortisol features in each ethnic group separately. We also utilize novel meta-analysis methods (MetaSKAT) for summarizing the ethnic specific gene-level results.

Methods

Study Population

The MESA Stress Study is an ancillary study to the Multi-Ethnic Study of Atherosclerosis (MESA). The MESA study is a longitudinal cohort study focused on investigating the early stages of atherosclerosis. Eligible participants were 45-84 years of age and free from history of cardiovascular disease at the baseline examination (2000-2002) [78]. The MESA Stress Study took place in the context of MESA examinations 3 and 4 conducted between 2004 and 2006, and obtained detailed stress hormone data on a subsample of 1002 MESA participants recruited from the New York and Los Angeles Field sites. Participants for the MESA Stress Study were African Americans, European Americans, and Hispanic Americans and were enrolled as they presented for follow-up, until approximately 500 participants were recruited from each location.

Of the 1002 MESA Stress Study participants, after exclusions for 1) raw cortisol data missingness, 2) unavailable genotype or principal component information, 3) no

consent for use of genetic information, and 4) concurrent corticosteroid usage, our resultant sample size was 839 individuals. The ethnic specific distribution of this sample is as follows: 170 European Americans, 215 African Americans, 454 Hispanic Americans.

Cortisol Sample Collection

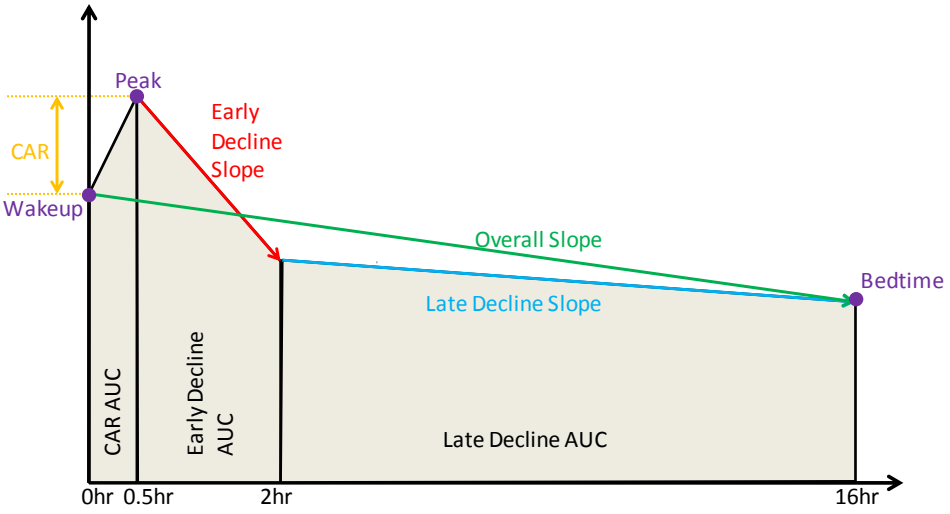
Each MESA Stress Study participant was asked to collect six saliva samples per day at pre-specified times over three consecutive weekdays, for a maximum of 18 samples per participant, using Salivette collection tubes. The samples were collected using the following schedule: sample (1) upon waking and before getting out of bed; (2) 30 minutes later; (3) around 10:00am; (4) around 12:00 noon or before lunch, whichever came first; (5) around 6:00pm or before dinner, whichever came first; (6) just before bed. Because earlier work has shown that the use of a time tracking device improves sample collection compliance [79], each collection tube was equipped with a time tracking device, which recorded the time when the swabs were removed for sample collection.

Cortisol Features

Rather than explore only cortisol concentrations at specific time points, we explored multiple features of the diurnal cortisol cycle (Table 20, Figure 20). Features were selected for investigation because prior work has hypothesized or demonstrated their associations with health risk factors or health outcomes [11,12,83,85]. Features were modeled using all available salivary cortisol data (up to six samples per day collected over three days). Raw cortisol concentrations, measured in nmol/L, were log-transformed to more closely approximate a normal distribution [5, 12, 80].

Table 20: Features of the diurnal cortisol curve. Cortisol concentrations were log-transformed.

	Cortisol Feature	Description
Time points	Wakeup	Average cortisol concentration from wakeup for an individual (Sample 1).
	Bedtime	Average cortisol concentration at bedtime for an individual (Sample 6).
Area	Area under the curve (AUC)	Standardized AUC for the interval 0hr-16hr since wakeup averaged across all days for an individual
Slopes	Cortisol awakening response (CAR)	The average difference in cortisol concentrations between the peak and wakeup measurements (Sample 2 – Sample 1).
	Early Decline Slope (EDSlope)	The slope from 0.5 hours and 2 hours since wakeup pooled across all days for an individual.
	Late Decline Slope (LDSlope)	The slope from 2 hours to 16 hours since wakeup pooled across all days for an individual
	Overall Decline Slope (ODSlope)	The overall decline slope ignoring the peak value from wakeup to bedtime pooled across all days for an individual.



AUC 16hr = CAR AUC + Early Decline AUC + Late Decline AUC

Figure 20: Representation of the diurnal cortisol curve describing the cortisol features of interest. For these analyses we specifically used Wakeup, Bedtime, Cortisol awakening response (CAR), Area under the curve (AUC) from 0-16 hours, Early Decline Slope, Late Decline Slope, and Overall Decline Slope.

Genetic Data

Genotyping data included both measured and imputed SNPs available through participation in MESA SHARe (SNP Health Association Resource) project. Under the SHARe project, genome-wide genotyping was obtained using the Affymetrix Genome-Wide Human SNP Array 6.0 platform. Imputation to HapMap was completed at the MESA Genetics Centers using the IMPUTE2 [82] program with the following reference panels: the HapMap Phase I and II, the human genome reference sequence (NCBI Build 36). The HapMap project is based on ethnic specific reference panels, composed of the following groups: Yoruba in Ibadan, Nigeria (abbreviation: YRI), Japanese in Tokyo, Japan (abbreviation: JPT), Han Chinese in Beijing, China (abbreviation: CHB), CEPH (Utah residents with ancestry from northern and western Europe) (abbreviation: CEU). Imputation for African Americans and Hispanic Americans was performed using the CEU+YRI+CHB+JPT reference panels (release #22). Imputation for European Americans was performed using only the CEU reference panel (release #24). All imputed and genotyped SNPs were aligned to the "+" strand of the human genome reference sequence (NCBI Build 36). Based on the imputed allele probabilities (AA, AB, BB), most like genotypes were assigned as 0, 1, 2 counts of the minor allele. If the probability of AA was greater than the probability of BB, then allele A was considered the effect allele. If the allele frequency of A was < 0.5, it was considered the minor allele; otherwise allele B was considered minor.

Stress Response Genes

We defined the six stress response gene regions as the entire gene, plus a window ±5kb around each gene. Base pair start and end positions for each gene were assigned based off annotation from the UCSC Genome Browser [86]. Starting base pair positions were rounded down to the nearest kb and ending positions were rounded up to the nearest kb. An additional 5kb were then added upstream of the starting positions and downstream of the ending positions. These gene regions were then restricted to SNPs within the entire gene ±5kb window that were common variants (minor allele frequency (MAF) >5%). Due to the small ethnic group sample sizes, a threshold for the MAF of 5% was chosen to limit the influence of unstable frequency estimates being driven by small sample sizes. Specific details on the chromosomal locations of each of the six genes, the overall size of their regions, and the number of ethnic group specific SNPs in each region with a MAF > 5% can be found in the Appendix (Tables A1-A3).

Statistical Analysis

Gene-level analyses

We performed SKAT separately for each gene region, in each ethnic group, for each cortisol feature separately. All analyses were adjusted for baseline age, sex, and education. The result was a score statistic, Q, for each analysis.

The general SKAT model for testing genetic main effects is as follows:

$$Y_i = \alpha_0 + \boldsymbol{\alpha}' \boldsymbol{X}_i + \boldsymbol{\beta}' \boldsymbol{G}_i + \epsilon_i$$

Where Y_i is the outcome corresponding to subject i, α_0 is an intercept term, $\boldsymbol{X_i}$ is a vector of non-genetic covariates, $\boldsymbol{G_i}$ is a vector of genotypes, and measurement error ε_i follows any distribution with mean zero and variance σ^2. $\boldsymbol{\alpha}$ is a vector of regression coefficients for the covariates, and $\boldsymbol{\beta}$ is a vector of regression coefficients for the genotypes. In SKAT one assumes that each of the β_j, j=1,..,p, follows an arbitrary distribution with mean zero and variance $w_j\tau$. Testing H_0: $\tau = 0$ is equivalent to testing H_0: $\boldsymbol{\beta} = 0$.

The SKAT framework allows for the specification of multiple kernel types, which determine how the genotype information is included in the model, as well as multiple weighting functions. The weights w_j can be specified or set to 1 for instances where weighting in unnecessary. SKAT was designed with rare variant analyses in mind and as such allows for the up-weighting of rare variants, under the assumption that common variants are less likely to have large effects. As we are interested in the effect of common variants, there was no need to utilize the up-weighting algorithms available in SKAT. Instead, we executed the gene level analyses implementing the linear kernel.

Meta-analysis

A new methodology available through the SKAT framework allows for the meta-analysis of the gene-level results across groups, MetaSKAT. MetaSKAT allows for the analysis of either cohort level summary statistics or individual level results. We utilized individual-level genotype data since it was available. Given that we have multi-ethnic sample and that minor allele frequencies may vary across groups, we allowed for a heterogeneous genetic effects model that used ethnic group specific minor allele frequencies. The current MetaSKAT kernel is specified to be the linear weighted kernel.

As we were using common variants we set weights.beta=c(1,1) to unweight the kernel. In SKAT, the Q statistic, a variance-component score statistic, is defined as $Q = (\mathbf{y} - \hat{u})'\mathbf{K}(\mathbf{y} - \hat{u})$. In MetaSKAT, the heterogeneous genetic effects model results in a sum of the ethnic group specific Q statistics. Both SKAT and MetaSKAT were packages executed using R (version 2.14.0) [87].

Power Calculations

To estimate power for gene-based association testing, we used the following SKAT power calculation parameters: 100 simulations over a 40kb gene region, and setting the maximum effect to 2, percentage of causal SNPs to 5%, and the frequency of negative interaction effects to 20%. We tested three ethnic groups, six gene regions, and seven cortisol outcomes, for a total of 126 tests. Using an alpha level of 0.05 and a Bonferroni correction, the result is a significance threshold at $p<4x10^{-4}$. Table 21 presents the SKAT output table showing power calculations, based on the criteria above. Given the limited sample size of each ethnic group, we expect roughly 50% power among the Caucasian Americans (n=170), 58% among the African Americans (n=215), and 74% among the Hispanic Americans (n=454).

Table 21: Gene-level power calculations.

Sample Size	Alpha Level			
	0.01	**1e-04**	**4e-04**	**1e-06**
50	0.33	0.20	0.22	0.15
100	0.51	0.32	0.37	0.23
150	0.61	0.45	0.49	0.33
200	0.68	0.53	0.56	0.42
250	0.73	0.58	0.61	0.49
300	0.77	0.62	0.66	0.54
350	0.80	0.66	0.69	0.57
400	0.82	0.68	0.72	0.60
450	0.85	0.71	0.74	0.63
500	0.86	0.73	0.77	0.65

However, given the significant ($p<0.05$) correlation between cortisol features (Table 22), the Bonferoni corrected estimates represent a conservative lower bound as the tests are not independent. Therefore, the significance threshold for both SKAT and MetaSKAT were set to $p<0.05$. P-values that ranged from 0.05-0.10 were considered suggestive.

Table 22: Pearson's product moment correlation coefficients between cortisol summary features, all ethnic groups combined.

	Wakeup	**Bedtime**	**CAR**	**AUC**	**EDSlope**	**LDSlope**	**ODSlope**
Wakeup	1.00						
Bedtime	**0.39**	1.00					
CAR	**-0.44**	**-0.14**	1.00				
AUC	**0.49**	**0.74**	-0.01	1.00			
EDSlope	**-0.24**	-0.05	**-0.29**	**0.08**	1.00		
LDSlope	**0.08**	**0.50**	**-0.08**	**0.35**	**-0.39**	1.00	
ODSlope	**-0.34**	**0.43**	**0.18**	**0.39**	0.04	**0.72**	1.00

Bold = $p<0.05$

Results

Basic demographic information on the Stress Study participants is provided in Table 23. Hispanic Americans represented the largest proportion of participants (52.8%), relative to the African Americans (28.6%) and European Americans (18.6%). The gender distribution was fairly equal (52.4% female). Overall, cortisol feature means varied across ethnic groups (Table 24). There was a statistically significant difference in means across the ethnic groups for all cortisol features except CAR.

Table 23: Characteristics of MESA Stress Study participants.

	Frequency (n=1002)
Site	
Columbia	52.2%
UCLA	47.8%
Age	
45-54	29.9%
55-64	27.7%
65-74	30.3%
75-84	12.1%
Race	
European American	18.6%
African American	28.6%
Hispanic American	52.8%
Gender	
Male	47.6%
Female	52.4%
Education Level	
Less than High School	27.0%
Completed High School	20.2%
Some College	29.7%
Bachelor's or higher	23.2%
Income	
< $20,000	29.3%
$20,000-34,999	27.5%
$35,000-$49,999	16.5%
$50,000 or higher	26.8%
Percent Current Smokers	11.3%
Percent Diabetic	13.5%
Body Mass Index (BMI) >=30	36.7%

Table 24: Distributions of cortisol summary features.

Cortisol Feature	European Americans		African Americans		Hispanic Americans		ANOVA
	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	p-value
Wakeup	166	2.58 (0.54)	214	2.38 (0.55)	450	2.38 (0.58)	0.0002
Bedtime	166	0.78 (0.77)	212	0.98 (0.74)	448	0.49 (0.84)	<0.0001
CAR	160	0.45 (0.46)	203	0.35 (0.46)	412	0.37 (0.52)	0.17
AUC	166	1.64 (0.43)	209	1.60 (0.42)	442	1.46 (0.51)	<0.0001
EDSlope	163	-0.53 (0.35)	209	-0.42 (0.44)	433	-0.40 (0.44)	0.003
LDSlope	164	-0.12 (0.06)	211	-0.10 (0.06)	447	-0.13 (0.06)	<0.0001
ODSlope	169	-0.12 (0.07)	214	-0.10 (0.06)	452	-0.12 (0.06)	<0.0001

Cortisol concentrations (nmol/L) were log-transformed and combined across the three days of collection to create each feature. The mean and standard deviation (SD) of the transformed and summarized features are presented by ethnic group.

In the gene-level analyses of each gene region on each cortisol feature, there were a number of statistically significant associations. *ADRA2A* was a significant (p-value<0.05) predictor of AUC and EDSlope in European Americans and Bedtime in Hispanic Americans (Table 25). It was also a suggestive (p-value<0.1) predictor of AUC in Hispanic Americans and CAR in African Americans. In the meta-analysis across the three ethnic groups, *ADRA2A* was a marginal predictor of four out of the seven cortisol features. *ADRB2* was a significant predictor for Bedtime, CAR, and ODSlope features in the European Americans (Table 26). In the meta-analysis across the three ethnic groups, *ADRB2* was a significant predictor of CAR, an association that was driven by the strength of the association within the European Americans. There was suggestive evidence that *SLC6A4* is predictive of CAR among European Americans (Table 27). This gene region also showed a suggestive association for European Americans and a significant association for Hispanic Americans in predicting EDSlope. There was also a significant meta-analysis across ethnics of *SLC6A4* on EDSlope. *TH* had suggestive associations for EDSlope and LDSlope in the African Americans. Neither *NR3C1* nor *NR3C2* were

predictive of any of the cortisol features. Tables of the gene-level and meta-analysis results for *NR3C1*, *NR3C2*, and *TH* can be found in the Appendix (Tables A4-A6).

Table 25: Gene-level main effect and meta-analysis results for *ADRA2A*.

		SKAT		MetaSKAT
Outcome	**Race**	**Q**	**p-value**	**p-value**
AUC	AFA	75.08	0.94	
	EA	620.63	0.03**	
	HIS	1457.32	0.05*	0.05*
Bedtime	AFA	134.79	0.88	
	EA	267.43	0.22	
	HIS	1584.18	0.04**	0.08*
CAR	AFA	691.99	0.08*	
	EA	266.52	0.21	
	HIS	739.46	0.22	0.09*
EDSlope	AFA	142.07	0.78	
	EA	542.72	0.04**	
	HIS	524.09	0.38	0.22
LDSlope	AFA	69.02	0.95	
	EA	161.94	0.41	
	HIS	579.57	0.35	0.61
ODSlope	AFA	755.55	0.07*	
	EA	146.87	0.45	
	HIS	1041.60	0.13	0.07*
Wakeup	AFA	610.42	0.13	
	EA	26.34	0.92	
	HIS	781.66	0.23	0.22

* $p<0.1$, ** $p<0.05$

Table 26: Gene-level main effect and meta-analysis results for *ADRB2*.

		SKAT		MetaSKAT
Outcome	**Race**	**Q**	**p-value**	**p-value**
AUC	AFA	2981.16	0.33	
	EA	1490.07	0.41	
	HIS	1803.27	0.92	0.76
Bedtime	AFA	3321.44	0.26	
	EA	4224.59	0.02**	
	HIS	830.64	0.99	0.46
CAR	AFA	2885.84	0.34	
	EA	5368.71	0.01**	
	HIS	4557.04	0.37	0.06*
EDSlope	AFA	1996.78	0.69	
	EA	1958.35	0.25	
	HIS	4648.10	0.40	0.43
LDSlope	AFA	959.33	0.99	
	EA	2544.81	0.13	
	HIS	3532.02	0.61	0.70
ODSlope	AFA	3997.07	0.14	
	EA	4885.04	0.01**	
	HIS	3525.88	0.62	0.22
Wakeup	AFA	2560.95	0.49	
	EA	2793.26	0.11	
	HIS	2707.73	0.77	0.55

* $p<0.1$, ** $p<0.05$

Table 27: Gene-level main effect and meta-analysis results for *SLC6A4*.

		SKAT		MetaSKAT
Outcome	**Race**	**Q**	**p-value**	**p-value**
AUC	AFA	1353.92	0.21	
	EA	981.95	0.22	
	HIS	189.82	0.99	0.56
Bedtime	AFA	782.57	0.61	
	EA	626.78	0.38	
	HIS	261.03	0.97	0.89
CAR	AFA	731.20	0.62	
	EA	1675.96	0.07*	
	HIS	262.30	0.97	0.49
EDSlope	AFA	793.66	0.59	
	EA	1924.14	0.05*	
	HIS	5685.99	0.03**	0.01**
LDSlope	AFA	1515.22	0.16	
	EA	424.99	0.53	
	HIS	2300.88	0.24	0.25
ODSlope	AFA	1655.41	0.13	
	EA	504.80	0.47	
	HIS	635.64	0.78	0.56
Wakeup	AFA	738.96	0.65	
	EA	165.94	0.83	
	HIS	677.54	0.75	0.92

* $p<0.1$, ** $p<0.05$

The association between *SLC6A4* and EDSlope had the strongest meta-analysis evidence of significant ($p<0.05$) gene-level associations across ethnic groups. To further investigate the associations of *SLC6A4* on EDSlope in European Americans and Hispanic Americans, we used LocusZoom plots [88] to examine the association between individual SNPs within the gene region and EDSlope (Figures 21-23). In the European Americans the association is being driven by rs2066713, which is in high linkage disequilibrium

with one other SNP. In the Hispanic Americans a different SNP has the strongest association (rs4583306), and is in high linkage disequilibrium with several other SNPs. In the LocusZoom plots there does appear to be similar structure in the overall pattern of association for the suggestive p-values ($-\log_{10}$(p-value)<1) in the European Americans and Hispanic Americans. In contrast, the plot of the SNP associations for the African Americans has a different overall pattern. LocusZoom plots for the other significant and suggestive gene associations on cortisol features are available in the Appendix (Figures A1-A11).

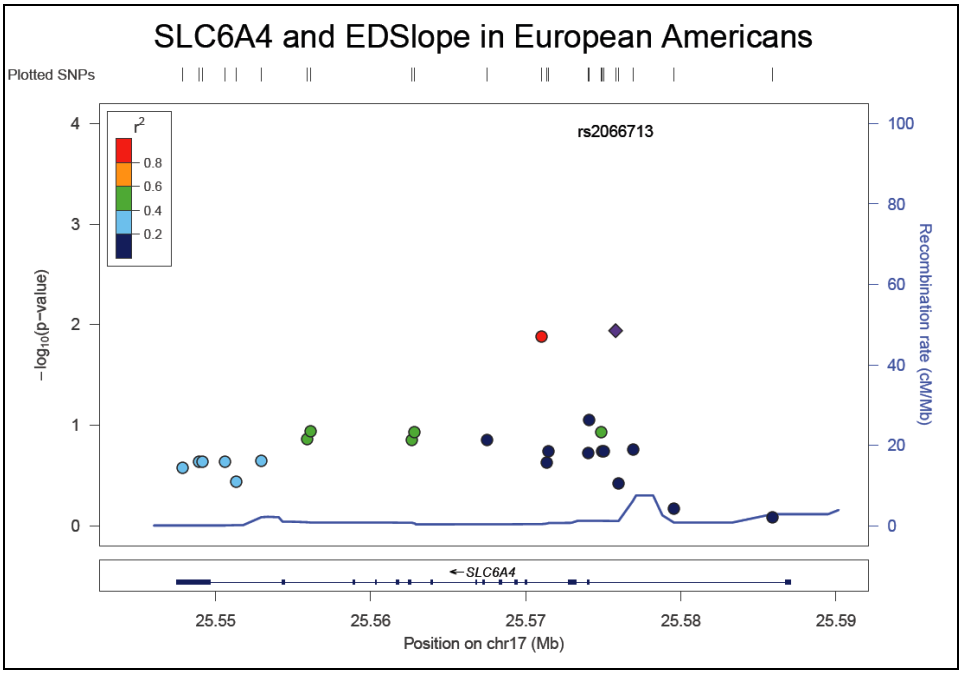


Figure 21: LocusZoom plot of the correlation between loci of the *SLC6A4* gene region among European Americans in predicting EDSlope.

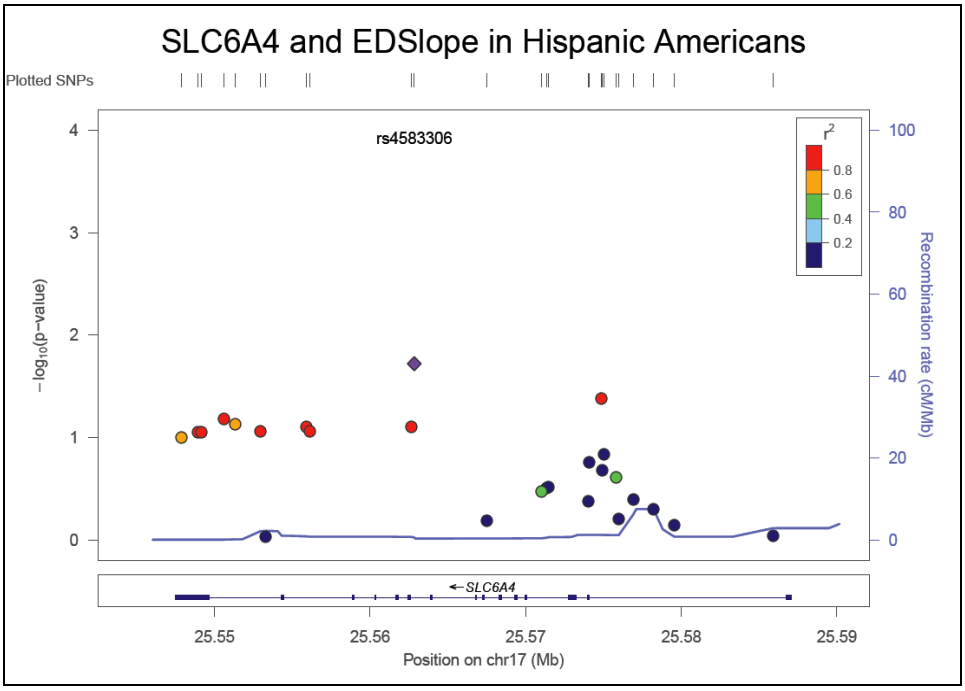


Figure 22: LocusZoom plot of the correlation between loci of the *SLC6A4* gene region among Hispanic Americans in predicting EDSlope.

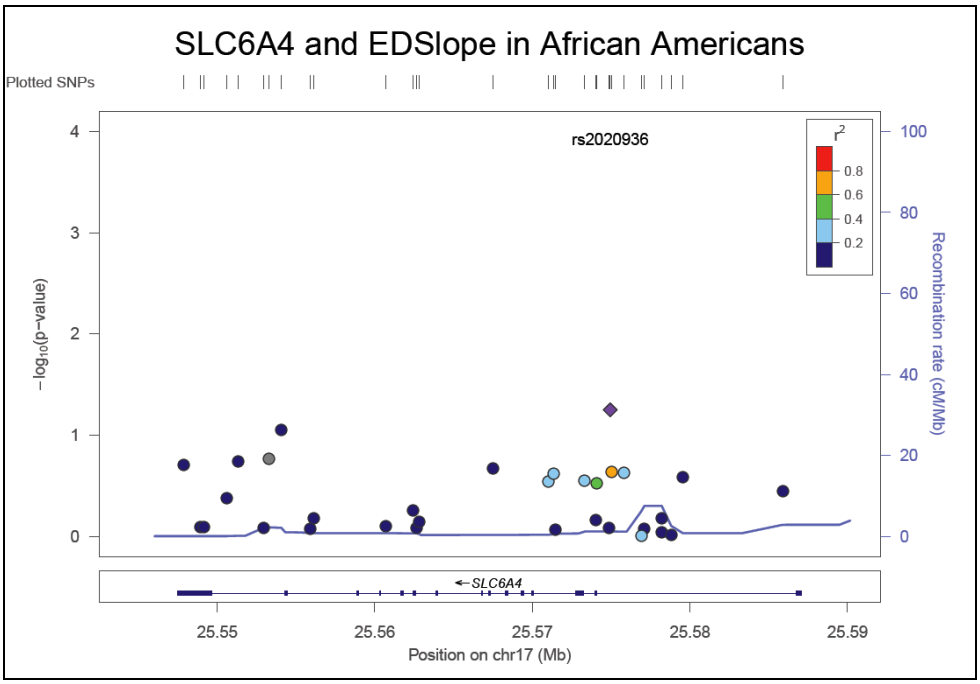


Figure 23: LocusZoom plot of the correlation between loci of the *SLC6A4* gene region among African Americans in predicting EDSlope.

Since SKAT does not provide estimates of specific SNP parameters (i.e. magnitude or direction of effect) we examined the effects of the index SNPs across the ethnic groups (Table 28), which were available from a genome-wide association study of EDSlope. GWAS were performed in each ethnic group separately, using SNPTest genetic analysis software (version 2) [89]. Linear regression was used to estimate the additive genetic effect of each SNP. We used the Frequentist=1 and Method=Expected specifications, which allowed for an additive model of association and use of expected genotype dosages, respectively. The primary model included age and sex as covariates. The top 10 principal components were also included in the model for African Americans and Hispanics after linear modeling indicated evidence of association between background genetic structure represented by the principal components and features of the cortisol curve. There was limited evidence of association in the Europeans, and as such we did not adjust for PCs. Filtering was performed to remove results for SNPs with minor allele frequency (MAF) less than 5% or with imputation quality (Info) < 0.5. The index SNPs were not significant ($p<0.05$) in the other ethnic groups, and rs2020936 in African Americans was only a suggestive association ($p<0.1$).

Table 28: Comparison of *SLC6A4* index SNPs across ethnic groups.

Index SNP	Race	Effect Allele	Frequency	B	p-value
rs2066713	AFA	G	0.76	-0.06	0.23
	CAU*	G	0.63	0.10	0.01
	HIS	G	0.66	0.04	0.24
rs4583306	AFA	G	0.21	0.02	0.71
	CAU	G	0.47	0.06	0.12
	HIS*	G	0.45	0.07	0.02
rs2020936	AFA*	G	0.40	-0.08	0.06
	CAU	G	0.15	0.07	0.18
	HIS	G	0.19	-0.05	0.21

* = Racial group where the index SNP was the strongest association between SLC6A4 and EDSlope.

Discussion

This study investigated the associations between selected stress response genes and cortisol features in a multi-ethnic population by utilizing a gene-level analysis approach, SKAT. We found statistical evidence that variation in established stress response gene regions is related to features of the diurnal cortisol curve, both across and within ethnic groups. Three of the six stress response gene regions revealed indication of across group effects, as evidenced by the suggestive and significant meta-analyses. There was evidence of *ADRA2A* having either marginal or suggestive statistical associations across ethnic groups in meta-analysis of multiple cortisol features: AUC, Bedtime, CAR, and ODSlope. The ethnic group that is driving the main effect associations of these meta-analyses differ by cortisol feature, indicating that the repeated association of *ADRA2A* is not due to low allele frequencies in one ethnic group alone. Minor allele frequencies for the *ADRA2A* SNP-sets in these analyses range from 0.07-0.32 in African

Americans, 0.05-0.27 in European Americans, and 0.07-0.43 in Hispanic Americans. Alpha-2A adrenergic receptors are found primarily in the brain stem and regulate the release of noradrenaline and adrenaline [53]. There is evidence that polymorphisms of this gene are associated with autonomic responses to environmental stressors [54].

Other stress response genes implicated in the gene-region meta-analyses include *ADRB2* for its suggestive association with CAR and *SLC6A4* for its significant association with EDSlope. Stimulation of beta-2 adrenergic receptors by adrenaline leads to vasodilation which counteracts the hypertensive effects of sympathetic activation by stressors [55]. *SLC6A4* may modulate the serotoninergic response to stress [62]. This hypothesis is supported by brain imaging studies showing that carriers of the "s" allele of *SLC6A4* are more responsive to emotional stimuli [63]. A recent study found that *SLC6A4* polymorphisms were associated with CAR [62]. While there was no statistical meta-analysis association between *SLC6A4* and CAR across the three ethnic groups in this study, there was a suggestive main effect association in the European Americans.

The differences in the index SNPs in the associations between EDSlope and *SLC6A4* as well as the correlation patterns with the index SNPs may be a result of underlying differences in the linkage disequilibrium patterns for the *SLC6A4* gene region for the three ethnic groups. Linkage disequilibrium plots were made using the SNP & Variation Suite v7 [90] (Appendix Figures A12-A14). The red colored blocks represent strong linkage disequilibrium, with an $R^2>0.8$. The European Americans and Hispanic Americans show a greater proportion of strong linkage disequilibrium compared to the

African Americans. The weaker correlation between SNPs among African Americans may contribute to the lack of association found in that group.

The SKAT methodology used for these analyses has several advantages over other gene-based association methods (e.g. Cohort Allelic Sum Test (CAST) [91], Weighted Sum Statistic (WSS) [92], C-alpha test [93]). First, SKAT is a more powerful method, even when sample sizes are small (n=500) [84], which is of particular importance given the small ethnic group specific sample sizes for these analyses. Second, SKAT allows for the individual variant effects to vary from mean zero in either direction, and does not assume that all variants have similar direction or magnitude of effect. Thirdly, it allows for the adjustment of covariates. SKAT additionally allows for the assessment of common variants by implementing an unweighted linear kernel, which fit our needs since we are using HapMap imputed genome-wide data.

There are two main limitations to this work. The first is a design limitation due the use of HapMap imputed variants, which are not functional SNPs. However, as the HapMap tagging SNPs may be in linkage disequilibrium with causal SNPs they are still useful for identifying genomic regions of potential interest. Secondly, compliance with cortisol sampling protocols is necessary for estimating reliable cortisol features [79, 94]. Compliance with taking samples within 10 minutes the requested times was greatest for wakeup (68%) and bedtime (75%) collections, and poorest during the middle of the day, ranging from 43%-57%. Stability of the cortisol features is of particular importance for genetic analyses, compared to other MESA cortisol work, as the effect estimates of

individual variants are expected to be modest and large variation in features estimates could mask true associations.

Despite the limitations, this work is novel in the ability to examine the variation in multiple gene regions across ethnic groups in predicting cortisol features, which was possible through the use of the innovative SKAT methodologies as well has the unique, highly detailed cortisol phenotype information. The gene-level analytic approach allows us to address the concern that individual SNPs may not replicate across ethnic groups due to differences in underlying patterns of linkage disequilibrium or to differences in allele frequencies [95-97], by examining a larger analysis unit which is unlikely to differ across populations. Our demonstration of the associations with different loci and correlation patterns for the results on *SLC6A4* in European Americans and Hispanic Americans emphasizes the need for gene-level approaches. The gene-based analyses presented here provide new insight into the relationship between stress response genes and cortisol features.

CHAPTER 4

INTERACTIONS BETWEEN CORTISOL LEVELS AND STRESS RESPONSE GENES IN PREDICTING CHRONIC DISEASE RISK FACTORS IN THE MULTI-ETHNIC STUDY OF ATHEROSCLEROSIS

Introduction

Multiple population-based studies have linked daily cortisol patterns to health outcomes, including elevated blood pressure, abdominal obesity, and coronary calcification [8-10]. Cortisol concentrations and various features of the cortisol daily profile have also been associated with diabetes mellitus [11] and markers of inflammation [12].

Environmental stressors that activate the cortisol-regulating HPA axis have a wide range of physiological and cellular implications. The cascade effect of the HPA axis on many tissues raises the question of whether gene-cortisol interactions play a role in the predisposition to many common chronic diseases.

The cortisol metabolic pathway suggests several key candidate genes whose variation could affect cortisol levels and/or influence the metabolic consequences of cortisol levels. In this paper we focus on six stress response genes: a glucocorticoid receptor gene (*NR3C1*), a mineralocorticoid receptor gene (*NR3C2*), the tyrosine hydroxylase gene (*TH*), the alpha-2A-adrenergic receptor gene (*ADRA2A*), the beta-2-adrenergic receptor gene (*ADRB2*), and the serotonin transporter gene (*SLC6A4*), all of

which have been suggested to be involved in the physiologic response to psychological stressors [22, 52, 54, 56, 62]. In particular, studies have suggested that genetic polymorphisms of the glucocorticoid receptor gene (*NR3C1*) may modify the cardiovascular and metabolic effects of cortisol [22, 24-26].

Few, if any, population-based studies have investigated these interactions in large samples, which may explain some inconsistencies in studies examining the relationship between stress and disease. In addition, genetic differences in linkage disequilibrium structures across ethnic groups may lead to different results across studies. To address this issue, we applied a novel gene-level analysis approach to investigate the influence of gene-by-cortisol interactions on anthropometric, metabolic, and inflammatory traits using the sequence kernel association test (SKAT) [84]. This gene-level analysis bypasses the problem that different SNPs may be associated with different relationships between cortisol and chronic disease the risk factors in different ethnicities because it makes a single gene-based assessment based on the distribution of all SNP-by-Cortisol interactions in the gene region.

Using several measures of the daily cortisol profile, in the Multi-ethnic Study of Atherosclerosis (MESA), we investigated whether genetic polymorphisms of 6 stress-region gene regions influenced the relationship between cortisol features and anthropometric, metabolic, and inflammatory markers.

Methods

Study Population

The MESA Stress Study is an ancillary study to the Multi-Ethnic Study of Atherosclerosis (MESA). The MESA study is a longitudinal cohort study focused on investigating the early stages of atherosclerosis. Eligible participants were 45-84 years of age and free from history of cardiovascular disease at the baseline examination (2000-2002) [78]. The MESA Stress Study took place in the context of MESA examinations 3 and 4 conducted between 2004 and 2006, and obtained detailed stress hormone data on a subsample of 1002 MESA participants recruited from the New York and Los Angeles sites. Participants for the MESA Stress Study were African Americans, European Americans, and Hispanic Americans and were enrolled as they presented for follow-up, until approximately 500 participants were recruited from each location.

Of the 1002 MESA Stress Study participants, after exclusions for 1) raw cortisol data missingness, 2) unavailable genotype or principal component information, 3) no consent for use of genetic information, and 4) concurrent corticosteroid usage, our resultant sample size was 839 individuals. The ethnic specific distribution of this sample is as follows: 170 European Americans, 215 African Americans, and 454 Hispanic Americans.

Cortisol Sample Collection

Each MESA Stress Study participant was asked to collect six saliva samples per day at pre-specified times over three consecutive weekdays, for a maximum of 18 samples, using Salivette collection tubes. The samples were collected using the

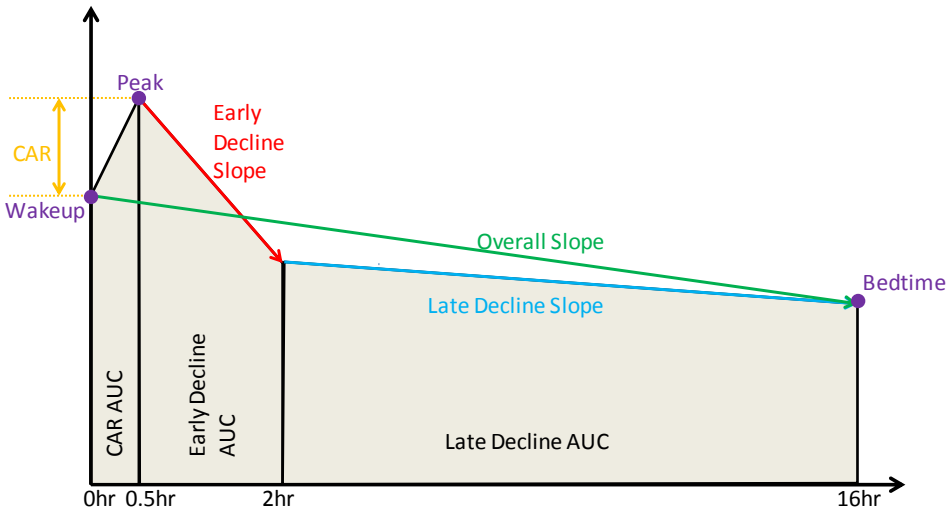
following schedule: sample (1) upon waking and before getting out of bed; (2) 30 minutes later; (3) around 10:00am; (4) around 12:00 noon or before lunch, whichever came first; (5) around 6:00pm or before dinner, whichever came first; (6) just before bed. Because earlier work has shown that the use of a time tracking device improves sample collection compliance [79], each collection tube was equipped with a time tracking device, which recorded the time when the swabs were removed for sample collection.

Cortisol Features

Rather than explore only cortisol concentrations at specific time points, we explored multiple features of the diurnal cortisol cycle (Table 29, Figure 24). Features were selected for investigation because prior work has hypothesized or demonstrated their associations with health risk factors or health outcomes [11, 12, 83, 85]. Features were modeled using all available salivary cortisol data (up to six samples per day collected over three days). Raw cortisol concentrations, measured in nmol/L, were log-transformed to more closely approximate a normal distribution [5, 12, 80].

Table 29: Features of the diurnal cortisol curve. Cortisol concentrations were log-transformed.

	Cortisol Feature	Description
Time points	Wakeup	Average cortisol concentration from wakeup for an individual (Sample 1).
	Bedtime	Average cortisol concentration at bedtime for an individual (Sample 6).
Area	Area under the curve (AUC)	Standardized AUC for the interval 0hr-16hr since wakeup averaged across all days for an individual
Slopes	Cortisol awakening response (CAR)	The average difference in cortisol concentrations between the peak and wakeup measurements (Sample 2 – Sample 1).
	Early Decline Slope (EDSlope)	The slope from 0.5 hours and 2 hours since wakeup pooled across all days for an individual.
	Late Decline Slope (LDSlope)	The slope from 2 hours to 16 hours since wakeup pooled across all days for an individual
	Overall Decline Slope (ODSlope)	The overall decline slope ignoring the peak value from wakeup to bedtime pooled across all days for an individual.



AUC 16hr = CAR AUC + Early Decline AUC + Late Decline AUC

Figure 24: Representation of the diurnal cortisol curve describing our summary features of interest. For these analyses we specifically used Wakeup, Bedtime, Cortisol awakening response (CAR), Area under the curve (AUC) from 0-16 hours, Early Decline Slope, Late Decline Slope, and Overall Decline Slope.

Genetic Data

Genotyping data included both measured and imputed SNPs available through participation in MESA SHARe (SNP Health Association Resource) project. Under the SHARe project, genome-wide genotyping was obtained using the Affymetrix Genome-Wide Human SNP Array 6.0 platform. Imputation to HapMap was completed at the MESA Genetics Centers using the IMPUTE2 [82] program with the following reference panels: the HapMap Phase I and II, the human genome reference sequence (NCBI Build 36). Imputation for African Americans and Hispanic Americans was performed using the CEU+YRI+CHB+JPT reference panels (release #22). Imputation for European Americans was performed using only the CEU reference panel (release #24). All imputed and genotyped SNPs were aligned to the "+" strand of the human genome reference sequence (NCBI Build 36). Based on the imputed allele probabilities (AA, AB, BB), most like genotypes were assigned as 0, 1, 2 counts of the minor allele. If the probability of AA was greater than the probability of BB, then allele A was considered the effect allele. If the allele frequency of A was < 0.5, it was considered the minor allele; otherwise allele B was considered minor.

Stress Response Genes

We defined the six stress response gene regions as the entire gene, plus a window ±5kb around each gene. Base pair start and end positions for each gene were assigned based off annotation from the UCSC Genome Browser [86]. Starting base pair positions were rounded down to the nearest kb and ending positions were rounded up to the nearest kb. An additional 5kb were then added upstream of the starting positions and

downstream of the ending positions. These gene regions were then restricted to SNPs within the entire gene ±5kb window that were common variants (minor allele frequency (MAF) >5%). Due to the small ethnic group specific sample sizes, a threshold for the MAF of 5% was chosen to limit the influence of unstable frequency estimates being driven by small sample sizes. Details on the chromosomal locations of each of the six genes, overall size, and the number of SNPs in each gene region by ethnic group can be found in the Appendix (Tables 1-3).

Outcome variables

Body mass index (BMI) was calculated as weight in kilograms/height in meters squared. Given the repeat assessment of BMI at each MESA exam, average BMI was estimated for each participant. Given that the MESA participants were adults, we assumed that an individual's weight was fairly stable across MESA exams, such that the average BMI was a stable representation of an individual's anthropometric characteristics. Fasting glucose concentrations were assessed by fasting blood draws and measured in mg/dl. Average fasting glucose concentrations across MESA exams was estimated for each participant. This average concentration was then natural log transformed. Within each ethnic group, individuals who were greater than ±4 standard deviations from the mean were excluded. Two markers of inflammation IL-6 and TNF-α were measured from fasting blood samples at the MESA Exam corresponding to an individual's cortisol collection. IL-6 concentrations were natural log transformed. There were two extreme observations for TNF-α which were removed; after which TNF-α concentrations were also natural log transformed.

Statistical Analysis

We performed SKAT separately for each gene, in each ethnic group, for each cortisol summary feature separately. The general SKAT model for testing genetic main effects is as follows:

$$Y_i = \alpha_0 + \boldsymbol{\alpha}' \boldsymbol{X}_i + \boldsymbol{\beta}' \boldsymbol{G}_i + \epsilon_i$$

Where Y_i is the outcome corresponding to subject i, α_0 is an intercept term, $\boldsymbol{X_i}$ is a vector of non-genetic covariates, $\boldsymbol{G_i}$ is a vector of genotypes, and measurement error ϵ_i follows any distribution with mean zero and variance σ^2. $\boldsymbol{\alpha}$ is a vector of regression coefficients for the covariates, and $\boldsymbol{\beta}$ is a vector of regression coefficients for the genotypes. In SKAT one assumes that each of the β_j's, j=1,..,p, follows an arbitrary distribution with mean zero and variance $w_j\tau$. The weights w_j can be specified or set to 1 for instances where weighting in unnecessary. Testing H_0: $\tau = 0$ is equivalent to testing H_0: $\boldsymbol{\beta} = 0$.

A version of SKAT, GESAT [98], allows for the evaluation of gene-by-environment interactions, after adjustment for covariates by the inclusion of an interaction term that represents the matrix of interactions:

$$Y_i = \alpha_0 + \boldsymbol{\alpha_1}' \boldsymbol{X}_i + \boldsymbol{\alpha_2}' \boldsymbol{G}_i + \boldsymbol{\alpha_3}' \boldsymbol{E}_i + \boldsymbol{\beta}' \boldsymbol{S}$$

Where $\mathbf{X}_i$ is a vector of non-genetic covariates, E_i is the environmental factor, $\mathbf{G}_i$ is a vector of genetic markers, and $\mathbf{S}_i$ is a vector of gene-by-environment interaction terms. One assumes that each of the β_j's, j=1,..,p, follows an arbitrary distribution with mean zero and common variance τ^2, and that the β_j's are independent. Testing H_0: $\tau^2 = 0$ is equivalent to testing H_0: $\boldsymbol{\beta} = 0$, which tests whether there is a marker set and environment

interaction. The cortisol features is the environment. Covariates include age, sex, and education at baseline. GESAT was implemented for each gene, for each cortisol feature, in each ethnic group, separately.

Since SKAT does not provide estimates of specific SNP-by-cortisol interaction parameters (i.e. magnitude or direction of effect) we also used traditional least squares regression approaches to estimate SNP-by-cortisol interaction terms when there was evidence of a significant gene-level interaction.

In GESAT, the number of markers in each SNP set cannot exceed the number of individuals, which was problematic for NR3C2, where the number of SNPs with a MAF > 5% in the SNPs sets ranges from 322 to 358 (African Americans, n=215; European Americans, n=170; Hispanic Americans, n=454). To work around this matrix structure limitation, NR3C2 was split into five smaller SNP sets within each ethnic group.

Power Calculations

To estimate power for gene-based association testing, we used the following SKAT power calculation parameters: 100 simulations over a 40kb gene region, and setting the maximum effect to 2, percentage of causal SNPs to 5%, and the frequency of negative interaction effects to 20%. We tested three ethnic groups, four outcomes, six gene regions, and seven cortisol outcomes, for a total of 504 tests. Using an alpha level of 0.05 and a Bonferroni correction, the result is a significance threshold at $p<9x10^{-5}$. Table 30 presents the SKAT output table showing power calculations, based on the criteria above. Given the limited sample size of each ethnic group, we expect roughly 54% power among the Caucasian Americans (n=170), 60% among the African

Americans (n=215), and 76% among the Hispanic Americans (n=454). However, given the significant ($p<0.05$) correlation between cortisol features (Table 31), the Bonferoni corrected estimates represent a conservative lower bound as the tests are not independent. Therefore, the significance threshold for GESAT was set to $p<0.05$.

Table 30: Gene-level power calculations.

Sample Size	**Alpha Level**		
	0.01	**9e-05**	**1e-06**
50	0.40	0.25	0.15
100	0.57	0.39	0.23
150	0.67	0.50	0.33
200	0.73	0.58	0.42
250	0.78	0.63	0.49
300	0.81	0.68	0.54
350	0.83	0.72	0.57
400	0.85	0.74	0.60
450	0.86	0.76	0.63
500	0.88	0.78	0.65

Table 31: Pearson's product moment correlation coefficients between cortisol summary features, all ethnic groups combined.

	Wakeup	**Bedtime**	**CAR**	**AUC**	**EDSlope**	**LDSlope**	**ODSlope**
Wakeup	1.00						
Bedtime	**0.39**	1.00					
CAR	**-0.44**	**-0.14**	1.00				
AUC	**0.49**	**0.74**	-0.01	1.00			
EDSlope	**-0.24**	-0.05	**-0.29**	**0.08**	1.00		
LDSlope	**0.08**	**0.50**	**-0.08**	**0.35**	**-0.39**	1.00	
ODSlope	**-0.34**	**0.43**	**0.18**	**0.39**	0.04	**0.72**	1.00

Bold = $p<0.05$

Results

Basic demographic information on the Stress Study participants is provided in Table 32. Hispanic Americans represented the largest proportion of participants (52.8%), relative to the African Americans (28.6%) and European Americans (18.6%). The gender distribution was fairly equal (52.4% female). Overall, cortisol feature means varied across ethnic groups (Table 33). There was a statistically significant difference in means across the ethnic groups for all cortisol features except CAR.

Table 32: Characteristics of MESA Stress Study participants.

	Frequency (n=1002)
Site	
Columbia	52.2%
UCLA	47.8%
Age	
45-54	29.9%
55-64	27.7%
65-74	30.3%
75-84	12.1%
Race	
European American	18.6%
African American	28.6%
Hispanic American	52.8%
Gender	
Male	47.6%
Female	52.4%
Education Level	
Less than High School	27.0%
Completed High School	20.2%
Some College	29.7%
Bachelor's or higher	23.2%
Income	
< $20,000	29.3%
$20,000-34,999	27.5%
$35,000-$49,999	16.5%
$50,000 or higher	26.8%
Percent Current Smokers	11.3%
Percent Diabetic	13.5%
Body Mass Index (BMI) >=30	36.7%

Table 33: Distributions of cortisol summary features.

Cortisol Feature	European Americans		African Americans		Hispanic Americans		ANOVA
	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	p-value
Wakeup	166	2.58 (0.54)	214	2.38 (0.55)	450	2.38 (0.58)	0.0002
Bedtime	166	0.78 (0.77)	212	0.98 (0.74)	448	0.49 (0.84)	<0.0001
CAR	160	0.45 (0.46)	203	0.35 (0.46)	412	0.37 (0.52)	0.17
AUC	166	1.64 (0.43)	209	1.60 (0.42)	442	1.46 (0.51)	<0.0001
EDSlope	163	-0.53 (0.35)	209	-0.42 (0.44)	433	-0.40 (0.44)	0.003
LDSlope	164	-0.12 (0.06)	211	-0.10 (0.06)	447	-0.13 (0.06)	<0.0001
ODSlope	169	-0.12 (0.07)	214	-0.10 (0.06)	452	-0.12 (0.06)	<0.0001

Cortisol concentrations (nmol/L) were log-transformed and combined across the three days of collection to create each feature. The mean and standard deviation (SD) of the transformed and summarized features are presented by ethnic group.

BMI was available for all individuals with phenotype and genotype information. ln(Glucose) and ln(TNF-α + 1) were available for nearly all individuals, while ln(IL-6 +1) and the greatest missingness (Table 34). In examining the difference in means between ethnic groups (Table 35), the means were significantly different from each other for all of the chronic disease risk factor outcomes except ln(TNF-α + 1). We ran multivariable models to assess the impact of the cortisol features on BMI, ln(Glucose), ln(IL-6 + 1), and ln(TNF-α + 1), adjusting for age, gender, race, and education at baseline (Tables 36-42). After accounting for covariates, every cortisol feature was a significant predictor of at least one outcome, with the exception of CAR.

Table 34: Distribution of outcome variables.

Outcome	N	Mean (SD)	Range
BMI kg/m^2	839	29.0 (5.6)	15.3-54.9
ln(Glucose) mg/dl	835	4.6 (0.2)	4.2-5.6
ln(IL-6 + 1) pg/ml	805	1.2 (0.4)	0.4-2.6
ln(TNF-α + 1) pg/ml	821	1.5 (0.6)	0.1-3.7

Table 35: Ethnic group specific distributions of chronic disease risk factors.

	European Americans				African Americans				Hispanic Americans				ANOVA
	N	**Min**	**Max**	**Mean (SD)**	**N**	**Min**	**Max**	**Mean (SD)**	**N**	**Min**	**Max**	**Mean (SD)**	**p-value**
BMI	170	18.2	47.0	26.6 (4.5)	215	15.3	54.0	30.1 (6.3)	454	18.7	54.9	29.5 (5.3)	<0.0001
ln(Glucose)	170	4.2	5.1	4.5 (0.1)	213	4.3	5.4	4.6 (0.2)	452	4.3	5.6	4.6 (0.2)	<0.0001
ln(IL-6 + 1)	166	0.5	2.5	1.1 (0.4)	205	0.4	2.5	1.2 (0.5)	434	0.4	2.6	1.3 (0.4)	0.009
ln(TNF-α + 1)	167	0.1	3.5	1.4 (0.6)	214	0.1	3.7	1.4 (0.6)	440	0.1	3.1	1.5 (0.6)	0.19

Table 36: Multivariable associations on chronic disease risk factors for AUC.

		Age		Gender		AFA		HIS		Education		AUC	
Outcome	**N**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**
BMI	817	-0.06	0.002	-1.29	0.0007	2.94	<0.0001	1.91	0.001	-0.26	0.007	-0.48	0.24
ln(Glucose)	813	0.001	0.18	0.04	0.008	0.06	0.01	0.08	0.0004	-0.01	0.007	-0.01	0.58
ln(IL-6 + 1)	785	0.01	<0.0001	-0.03	0.36	0.10	0.03	0.11	0.02	-0.02	0.02	0.08	0.02
ln(TNF-α + 1)	801	0.01	<0.0001	-0.03	0.42	0.08	0.18	0.15	0.02	0.01	0.43	0.01	0.82

Table 37: Multivariable associations on chronic disease risk factors for Bedtime.

		Age		Gender		AFA		HIS		Education		Bedtime	
Outcome	**N**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**
BMI	826	-0.06	0.002	-1.39	0.0002	3.05	<0.0001	1.99	0.0006	-0.24	0.01	-0.12	0.61
ln(Glucose)	822	0.001	0.26	0.04	0.01	0.06	0.008	0.08	0.0002	-0.01	0.01	0.0003	0.97
ln(IL-6 + 1)	793	0.01	<0.0001	-0.02	0.48	0.09	0.05	0.10	0.03	-0.02	0.02	0.04	0.07
ln(TNF-α + 1)	810	0.01	<0.0001	-0.03	0.49	0.09	0.15	0.16	0.01	0.01	0.31	-0.05	0.86

Table 38: Multivariable associations on chronic disease risk factors for CAR.

		Age		Gender		AFA		HIS		Education		CAR	
Outcome	**N**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**
BMI	775	-0.08	0.0001	-1.34	0.0005	3.18	<0.0001	2.01	0.0006	-0.24	0.01	-0.41	0.29
ln(Glucose)	771	0.001	0.27	0.04	0.005	0.07	0.003	0.09	0.0002	-0.01	0.03	-0.01	0.33
ln(IL-6 + 1)	745	0.01	<0.0001	-0.01	0.72	0.10	0.04	0.09	0.05	-0.02	0.03	-0.01	0.78
ln(TNF-α + 1)	759	0.01	<0.0001	-0.02	0.71	0.10	0.13	0.15	0.02	0.01	0.60	0.01	0.77

Table 39: Multivariable associations on chronic disease risk factors for EDSlope.

		Age		Gender		AFA		HIS		Education		EDSlope	
Outcome	**N**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**
BMI	805	-0.07	0.0003	-1.51	<0.0001	2.85	<0.0001	1.80	0.002	-0.23	0.02	1.30	0.004
ln(Glucose)	801	0.001	0.24	0.04	0.02	0.06	0.01	0.08	0.0008	-0.01	0.02	0.04	0.01
ln(IL-6 + 1)	773	0.01	<0.0001	-0.04	0.24	0.09	0.06	0.07	0.11	-0.02	0.04	0.14	0.0001
ln(TNF-α + 1)	789	0.01	<0.0001	-0.03	0.54	0.09	0.13	0.14	0.02	0.01	0.48	0.08	0.12

Table 40: Multivariable associations on chronic disease risk factors for LDSlope.

		Age		Gender		AFA		HIS		Education		LDSlope	
Outcome	**N**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**
BMI	822	-0.07	0.0006	-1.26	0.0008	2.99	<0.0001	2.09	0.0003	-0.25	0.008	0.18	0.95
ln(Glucose)	818	0.001	0.26	0.04	0.007	0.07	0.003	0.08	0.0003	-0.01	0.008	-0.08	0.47
ln(IL-6 + 1)	789	0.01	<0.0001	-0.02	0.48	0.09	0.05	0.11	0.02	-0.02	0.03	0.48	0.04
ln(TNF-α + 1)	805	0.01	<0.0001	-0.03	0.43	0.09	0.14	0.17	0.009	0.01	0.30	0.22	0.50

Table 41: Multivariable associations on chronic disease risk factors for ODSlope.

		Age		Gender		AFA		HIS		Education		ODSlope	
Outcome	**N**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**
BMI	835	-0.07	0.0002	-1.39	0.0002	2.93	<0.0001	2.18	0.0001	-0.23	0.01	6.89	0.03
ln(Glucose)	831	0.001	0.29	0.04	0.007	0.06	0.006	0.08	0.0002	-0.01	0.009	0.08	0.49
ln(IL-6 + 1)	801	0.01	<0.0001	-0.04	0.20	0.07	0.10	0.10	0.02	-0.02	0.04	1.18	<0.0001
ln(TNF-α + 1)	817	0.01	<0.0001	-0.04	0.34	0.07	0.24	0.15	0.02	0.01	0.29	0.40	0.24

Table 42: Multivariable associations on chronic disease risk factors for Wakeup.

		Age		Gender		AFA		HIS		Education		Wakeup	
Outcome	**N**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**
BMI	830	-0.05	0.003	-1.27	0.0006	2.95	<0.0001	2.06	0.0003	-0.22	0.02	-1.19	0.0003
ln(Glucose)	826	0.001	0.19	0.04	0.004	0.06	0.004	0.08	0.0002	-0.01	0.02	-0.02	0.07
ln(IL-6 + 1)	796	0.01	<0.0001	-0.02	0.52	0.09	0.05	0.09	0.05	-0.02	0.03	-0.07	0.008
ln(TNF-α + 1)	812	0.01	<0.0001	-0.04	0.43	0.07	0.24	0.14	0.03	0.01	0.27	-0.05	0.17

The results for the GESAT gene-level assessment of gene-by-cortisol interactions are presented in Tables 43-46. Given concerns with multiple testing, we considered instances where there was evidence of gene-level interaction ($p<0.05$) in at least two of the three ethnic groups for any gene-by-cortisol feature combination to provide the strongest evidence of gene-level interaction effects. There were six instances where the interaction p-value was < 0.05 in more than one ethnic group: for ln(Glucose), the interaction between *SLC6A4* and ODSlope; for ln(IL-6 + 1), the interaction between, *ADRB2* and Bedtime, *NR3C2* and Bedtime, and *NR3C2* and EDSlope; for TNF-α, the interaction between *NR3C1* and Bedtime, and the interaction between *NR3C2* and Wakeup. The p-values for NR3C2 are the smallest of the five sub-sectioned NR3C2 analyses.

Table 43: Gene-by-cortisol interactions for BMI.

Cortisol Feature	**Race**	***ADRA2A* p-value**	***ADRB2* p-value**	***NR3C1* p-value**	***NR3C2* p-value**	***SLC6A4* p-value**	***TH* p-value**
AUC	AFA	**0.002**	0.28	0.27	0.31	0.32	0.97
	EA	0.18	0.57	0.62	0.22	**0.001**	0.39
	HIS	0.68	0.62	0.63	0.32	0.09	0.09
Bedtime	AFA	0.14	0.14	**0.03**	0.20	0.36	0.51
	EA	0.67	0.10	0.06	0.52	**0.001**	0.28
	HIS	0.34	0.89	0.86	0.06	0.94	0.41
CAR	AFA	0.88	0.50	0.42	0.23	0.22	0.89
	EA	**0.03**	0.14	0.76	0.06	0.16	0.22
	HIS	0.78	0.05	0.81	0.22	0.97	0.09
EDSlope	AFA	0.41	0.43	**0.03**	0.71	0.16	0.98
	EA	0.77	0.27	0.86	**0.005**	0.47	0.84
	HIS	0.50	0.66	0.22	0.36	0.55	**0.01**
LDSlope	AFA	0.09	0.79	0.06	0.29	**0.02**	**0.03**
	EA	0.53	0.09	0.20	0.45	0.13	0.94
	HIS	0.78	0.98	0.10	0.43	0.63	0.91
ODSlope	AFA	0.28	0.09	0.06	0.20	0.08	0.12
	EA	0.71	0.25	0.69	0.10	0.12	0.86
	HIS	0.93	0.41	0.16	0.42	0.72	0.12
Wakeup	AFA	0.06	0.13	0.21	0.65	0.41	0.34
	EA	0.14	0.45	0.76	**0.01**	0.52	0.46
	HIS	0.26	0.37	0.12	0.24	0.55	0.12

Bold = $p<0.05$

Table 44: Gene-by-cortisol interactions for ln(Glucose).

Cortisol Feature	**Race**	***ADRA2A* p-value**	***ADBR2* p-value**	***NR3C1* p-value**	***NR3C2* p-value**	***SLC6A4* p-value**	***TH* p-value**
AUC	AFA	0.76	0.58	0.93	0.12	0.65	1.00
	EA	0.75	0.87	0.26	0.09	**0.004**	0.13
	HIS	**0.02**	0.35	0.61	0.80	0.86	0.92
Bedtime	AFA	0.98	0.98	0.98	0.12	0.98	0.60
	EA	0.99	0.78	0.23	0.33	**0.04**	**0.01**
	HIS	0.13	0.81	0.93	0.17	0.61	0.96
CAR	AFA	0.83	0.73	0.37	0.27	0.43	0.22
	EA	0.38	0.86	0.57	0.24	0.54	0.07
	HIS	0.06	0.79	0.44	0.53	0.15	0.50
EDSlope	AFA	0.33	0.23	0.21	0.48	0.96	0.16
	EA	0.66	0.47	**0.01**	0.30	0.39	0.45
	HIS	0.63	0.26	0.06	0.66	0.81	0.10
LDSlope	AFA	0.37	0.21	0.18	0.10	**0.003**	**0.01**
	EA	0.20	0.36	0.20	0.27	0.20	0.63
	HIS	0.43	0.52	0.11	**0.03**	0.73	0.44
ODSlope	AFA	0.11	0.40	0.18	0.27	**0.002**	**0.03**
	EA	0.48	0.29	0.62	0.06	**0.01**	0.47
	HIS	0.11	0.58	0.67	0.48	0.93	0.66
Wakeup	AFA	0.08	0.10	0.34	0.51	0.37	0.43
	EA	0.63	0.62	0.92	**0.008**	0.20	0.26
	HIS	0.56	1.00	0.07	0.46	0.80	0.34

Bold = $p<0.05$

Table 45: Gene-by-cortisol interactions for ln(IL-6 + 1).

Cortisol Feature	**Race**	***ADRA2A* p-value**	***ADBR2* p-value**	***NR3C1* p-value**	***NR3C2* p-value**	***SLC6A4* p-value**	***TH* p-value**
AUC	AFA	0.38	**0.006**	0.24	0.07	0.28	0.19
	EA	0.30	0.80	0.22	**0.02**	**0.008**	0.37
	HIS	0.24	0.55	0.31	0.09	0.19	0.83
Bedtime	AFA	0.07	**0.03**	0.81	**0.02**	0.14	0.16
	EA	0.09	0.37	0.07	**0.02**	**0.02**	0.41
	HIS	0.31	**0.03**	0.59	0.18	0.46	0.97
CAR	AFA	0.85	0.15	0.86	**0.02**	0.34	0.69
	EA	0.52	**0.01**	0.20	0.05	**0.02**	0.06
	HIS	0.68	0.16	0.59	0.08	0.14	0.17
EDSlope	AFA	0.14	0.46	0.16	**0.02**	0.37	0.53
	EA	0.19	0.77	0.73	**0.02**	0.38	0.15
	HIS	0.65	0.22	0.11	0.05	0.77	**0.03**
LDSlope	AFA	0.07	0.07	0.95	0.08	**0.03**	0.54
	EA	0.29	0.09	0.23	0.08	0.24	0.48
	HIS	0.24	0.44	0.40	0.40	0.98	0.34
ODSlope	AFA	0.69	0.09	0.23	0.08	0.52	0.65
	EA	0.06	0.23	0.30	**0.04**	0.29	0.05
	HIS	0.23	0.96	0.11	0.31	0.95	0.85
Wakeup	AFA	0.98	0.34	0.25	0.20	0.23	0.26
	EA	0.95	0.12	0.52	**0.04**	0.21	0.62
	HIS	0.54	0.44	0.78	0.10	0.51	0.09

Bold = $p<0.05$

Table 46: Gene-by-cortisol interactions for ln(TNF-a + 1).

Cortisol Feature	**Race**	***ADRA2A* p-value**	***ADBR2* p-value**	***NR3C1* p-value**	***NR3C2* p-value**	***SLC6A4* p-value**	***TH* p-value**
AUC	AFA	0.71	0.41	0.82	0.05	0.28	0.74
	EA	0.45	0.59	0.39	0.42	0.98	0.09
	HIS	0.28	0.45	**0.01**	0.13	0.27	0.95
Bedtime	AFA	0.45	0.95	0.42	0.08	0.39	0.49
	EA	0.39	0.83	**0.01**	**0.04**	0.82	0.25
	HIS	0.86	0.62	**0.01**	0.22	0.60	0.51
CAR	AFA	0.80	0.15	0.16	0.18	0.62	0.64
	EA	0.72	0.81	0.86	0.54	0.21	0.76
	HIS	0.90	0.88	0.65	0.43	0.62	0.20
EDSlope	AFA	0.86	0.79	**0.02**	0.10	0.89	0.64
	EA	0.28	0.43	0.33	0.11	0.08	0.70
	HIS	0.22	0.18	0.86	0.10	0.50	0.37
LDSlope	AFA	0.65	0.31	0.07	0.77	0.93	0.48
	EA	0.81	0.67	0.56	0.23	0.64	0.46
	HIS	0.65	**0.02**	**0.02**	0.11	0.92	0.60
ODSlope	AFA	0.69	0.54	0.63	0.13	0.70	0.22
	EA	0.68	0.97	0.55	0.05	0.27	0.65
	HIS	0.08	0.39	**0.02**	0.44	0.71	0.33
Wakeup	AFA	0.82	0.23	0.59	0.52	0.50	0.73
	EA	0.44	0.37	0.43	**0.03**	0.69	0.45
	HIS	**0.007**	0.50	0.66	**0.04**	0.54	0.46

Bold = $p<0.05$

To further investigate the gene-by-cortisol interactions where there was significant evidence ($p < 0.05$) in two ethnic groups, we examined the individual SNP-by-cortisol interactions in these regions. The interaction between SLC6A4 and ODSlope had the most significant evidence across the ethnic groups, with interaction p-values ≤ 0.01 for African Americans and European Americans. The individual SNP-by-cortisol interaction results for this association are shown in Table 47. SNP level results for the other five instances are in the Appendix (Tables A7– A11).

Table 47: Individual SNP-by-ODSlope interactions in predicting ln(Glucose).

	Cortisol				Coded		SNP Effects		Cortisol Effects		Interaction Effects	
Outcome	**Feature**	**Gene**	**Race**	**SNP**	**Allele**	**Freq**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**
ln(Glucose)	ODSlope	*SLC6A4*	AFA	rs9303628	T	0.37	0.09	0.02	-0.91	0.02	1.02	0.004
				rs3794808	T	0.40	-0.08	0.04	0.70	0.05	-0.99	0.004
				rs2066713	G	0.76	0.10	0.01	-0.65	0.04	1.05	0.004
				rs2020942	T	0.24	0.10	0.01	-0.68	0.04	0.98	0.008
				rs25528	T	0.49	0.11	0.009	-1.02	0.02	0.96	0.01
				rs8076005	G	0.63	0.11	0.006	-0.69	0.05	0.81	0.02
				rs140701	T	0.30	-0.06	0.20	0.38	0.26	-0.89	0.03
				rs2054848	T	0.90	-0.05	0.42	0.25	0.40	-1.08	0.04
			EA	rs140701	T	0.47	-0.07	0.004	0.56	0.02	-0.42	0.02
				rs2020942	T	0.37	0.05	0.06	-0.26	0.18	0.45	0.02
				rs2020939	G	0.52	-0.06	0.01	0.53	0.03	-0.41	0.03
				rs4583306	G	0.47	-0.06	0.007	0.52	0.03	-0.40	0.03
				rs3794808	T	0.47	-0.07	0.003	0.52	0.03	-0.38	0.03
				rs2066713	G	0.63	0.04	0.10	-0.24	0.22	0.41	0.03

Interestingly, four SNPs, rs2020942, rs140701, rs2066713, and rs3794808, had significant interaction p-values in both African Americans and European Americans. The direction of effect for each SNP is also consistent across the ethnic groups.

Some of the individual SNP findings may be driven by linkage disequilibrium patterns within the gene regions. LocusZoom [88]plots were used to evaluate the linkage disequilibrium with top SNPs in the six instances where there was evidence of gene-by-cortisol in two ethnic groups. The plot for the association between *SLC6A4* and ODSlope is shown in Figures 25 and 26. In the African Americans there is not strong evidence of linkage disequilibrium between the index SNP, rs9303628, and other SNPs in *SLC6A4*. In the European Americans, however, there is strong linkage disequilibrium (r^2>0.8) between the index SNP, rs140701, and a number of other SNPs. The plots for the other five instances of gene-by-cortisol interaction are available in the appendix (Figures A15-A23).

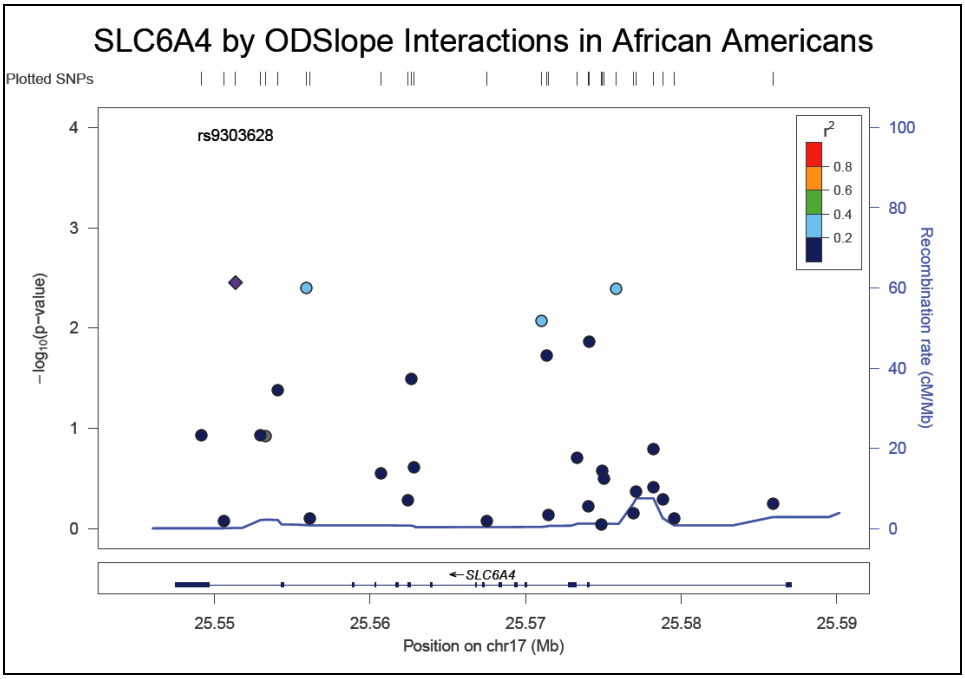


Figure 25: LocusZoom plot of the interaction between loci of the *SLC6A4* gene region and EDSlope among African Americans in predicting ln(Glucose).

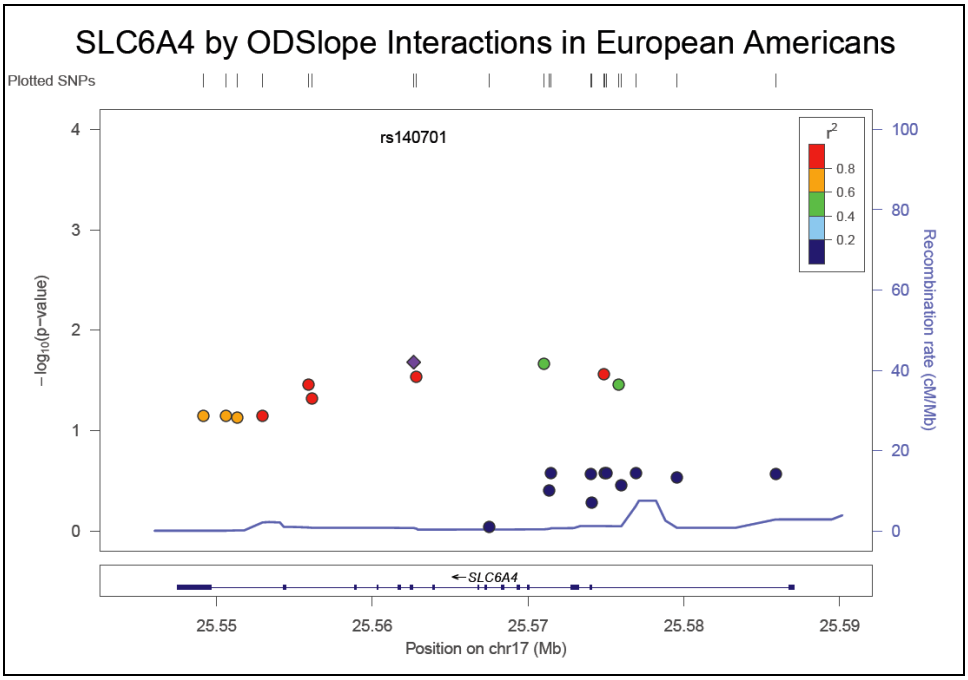


Figure 26: LocusZoom plot of the interaction between loci of the *SLC6A4* gene region and EDSlope among European Americans in predicting ln(Glucose).

The initial individual SNP-by-cortisol interaction models assumed an additive effect of the minor allele. We also ran agnostic models that did not assume an additive effect. Table 48 presents the results for the SNP and interaction effects for the individual SNP-by-ODSlope interactions in predicting ln(Glucose). These models compare the effects of having one copy of the minor allele to zero, as well as the presence of two copies of the minor allele to zero. The SNPs in Table 48 are ordered the same as in Table 47, where they were initially ranked in order of p-value by ethnic group where the interaction p-value was <0.05. The direction of SNP effects for one or two copies of the minor allele are consistent for this set of SNPs. With one exception, the p-values for the interaction between two copies of the minor allele and ODSlope were more significant that the p-values of the interaction with one copy. The tables for the agnostic approach for the other five instances are in the Appendix (Tables A12-A16).

Table 48: Two degree of freedom test of individual SNP-by-ODSlope interactions in predicting ln(Glucose)

	Cortisol				Coded		Minor	SNP Effects		Interaction Effects	
Outcome	**Feature**	**Gene**	**Race**	**SNP**	**Allele**	**Freq**	**Alleles**	**B**	**p-value**	**B**	**p-value**
ln(Glucose)	ODSlope	*SLC6A4*	AFA	rs9303628	T	0.37	1	0.05	0.45	0.87	0.12
							2	0.20	0.01	2.02	0.005
				rs3794808	T	0.40	1	-0.13	0.05	-0.70	0.20
							2	-0.15	0.06	-2.12	0.002
				rs2066713	G	0.76	1	0.03	0.61	0.37	0.51
							2	0.25	0.005	2.73	0.001
				rs2020942	T	0.24	1	0.03	0.62	0.23	0.67
							2	0.26	0.005	2.74	0.001
				rs25528	T	0.49	1	0.09	0.23	0.47	0.45
							2	0.23	0.007	2.08	0.008
				rs8076005	G	0.63	1	0.07	0.25	0.20	0.72
							2	0.23	0.004	1.98	0.006
				rs140701	T	0.30	1	-0.05	0.39	-0.70	0.18
							2	-0.15	0.20	-2.26	0.04
				rs2054848	T	0.90	1	-0.01	0.88	-0.97	0.09
							2	-0.29	0.25	-3.35	0.17
			EA	rs140701	T	0.47	1	-0.10	0.04	-0.57	0.09
							2	-0.14	0.05	-0.89	0.02
				rs2020942	T	0.37	1	0.03	0.04	0.35	0.21
							2	0.10	0.05	0.97	0.02
				rs2020939	G	0.52	1	-0.09	0.05	-0.63	0.07
							2	-0.12	0.05	-0.86	0.02

	rs4583306	G	0.47	1	-0.08	0.04	-0.48	0.15
				2	-0.13	0.05	-0.82	0.03
	rs3794808	T	0.47	1	-0.10	0.05	-0.48	0.16
				2	-0.14	0.05	-0.83	0.02
	rs2066713	G	0.63	1	0.02	0.04	0.28	0.32
				2	0.10	0.05	0.93	0.03

Discussion

This paper examined the key question of whether gene-by-cortisol interactions influence chronic disease risk factors. There was evidence of stress response gene-by-cortisol interaction in predicting BMI, Glucose, IL6, and TNF-α across all seven cortisol features and three ethnic groups. While there were numerous significant ($p<0.05$) individual gene-by-cortisol interactions in one ethnic group, there were six instances when there was a significant interaction among two ethnic groups.

The evidence of interaction between *NR3C1* and cortisol features is particularly interesting. *NR3C1* is a glucocorticoid receptor that is occupied when cortisol concentrations are high [36]. The interactions indicate that variation in *NR3C1* influences the effect of cortisol on the downstream outcomes of glucose and TNF- α concentrations.

Additionally, there were several instances where within an ethnic group there were consistent associations across chronic disease risk factor outcomes. In European Americans, the interaction between *NR3C2* and Wakeup was significant ($p<0.05$) for all four chronic disease risk factor outcomes. Also in European Americans, the interactions between AUC and Bedtime with SLC6A4 were significant predictors of BMI, ln(Glucose), and ln(IL-6 + 1). In African Americans, the interaction between LDSlope and SLC6A4 was also significant for BMI, ln(Glucose), and ln(IL-6 + 1). While previous work has shown that polymorphisms in the promoter region of *SLC6A4* are associated with CAR [62], to my knowledge there is no published information on associations with the other cortisol features examined in this study. As *SLC6A4* is hypothesized to impact the stress response, the interaction between variations in the gene

region and stress hormone levels having downstream effects on chronic disease risk factors is a novel and important finding.

Previous work has examined the relationship between cortisol and chronic disease risk factors. The consequences of HPA axis dysfunction and the extreme hypercortisolism associated with Cushing's syndrome are well known and include glucose intolerance, type II diabetes, and alterations of fat distribution [65]. However, cortisol also has important metabolic consequences at normal physiologic levels, which if extended over long periods as a consequence of chronic stress could result in HPA axis dysregulation that has important downstream effects on glucose metabolism, insulin resistance, and fat deposition [22, 29]. There is some evidence that cortisol and cortisol dysregulation are related to body fat distribution, obesity [29, 66, 67], and diabetes-related outcomes [8, 68-70]. Additionally, there has been increasing evidence that chronic stress can result in elevations of systemic inflammatory markers [73-75]. Recent work has noted associations between CRP (C-reactive protein, a marker of inflammation) and the cortisol awakening response [99], while IL-6 and TNF- α have been associated with several cortisol features [12].

The SKAT methodology used for these analyses has several advantages over other gene-based association methods (e.g. Cohort Allelic Sum Test (CAST) [91], Weighted Sum Statistic (WSS) [92], C-alpha test [93]). First, SKAT is a more powerful method, even when sample sizes are small (n=500) [84], which is of particular importance given the small ethnic group specific sample sizes for these analyses. Second, SKAT allows for the individual variant effects to vary from mean zero in either direction, and does not assume

that all variants have similar direction or magnitude of effect. Thirdly, it allows for the adjustment of covariates. SKAT additionally allows for the assessment of common variants by implementing an unweighted linear kernel, which fit our needs since we are using HapMap imputed genome-wide data.

We followed up the SKAT analyses with individual SNP-by-cortisol associations, using traditional least squares regression approaches. These analyses began with assuming and additive model for the genetic effects. However, assuming additive SNP effects confines the SNP effects such that the difference between zero and one copy of the minor allele is expected to be equivalent to the difference between one and two copies. Therefore, we also ran agnostic models that did not assume perfectly additive effects, but allowed for a comparison of the effect of one copy of the minor allele compared to zero and the effect of two copies of the minor allele compared to zero. If the effects had been perfectly additive, the effect of two copies of the minor allele should have been double that of one copy. We did not see evidence of this perfect minor allele dose response. These analyses should be followed up, looking at each genotype class separately in order to model the effects of each genotype-by-cortisol interaction in predicting the chronic disease risk factor outcomes.

This study has some limitations. First, compliance with cortisol sampling protocols is necessary for estimating reliable cortisol features [79, 94]. Compliance with taking samples within 10 minutes the requested times was greatest for wakeup (68%) and bedtime (75%) collections, and poorest during the middle of the day, ranging from 43%-57%. Second, is a design limitation due the use of HapMap imputed variants, which are

not functional SNPs. Third, these analyses assumed an additive genetic effect, which ignores the possible influence of dominant or epistatic variations.

Fourth, the gene-level analysis method was unable to handle the original matrix dimensions of *NR3C2* given the limited ethnic group specific sample sizes. The SKAT methodology assesses sets of SNPs which do not necessarily have to make up a gene, and as such evaluating smaller sections of a large gene solves a structural problem. As SKAT assesses whether the individual SNP effects vary from a mean of zero in either direction, the cumulative effect of individual SNPs in the sub-set regions may not reflect the overall cumulative effect across *NR3C2*.

Despite the limitations, this work is novel in the ability to examine gene-by-cortisol interactions, considering multiple gene regions, cortisol features, and ethnic groups, in predicting chronic disease risk factors, which was possible through the use of the innovative SKAT methodologies as well has the unique, highly detailed cortisol phenotype information. The gene-level analytic approach allows us to address the concern that individual SNPs may not replicate across ethnic groups due to differences in underlying patterns of linkage disequilibrium or to differences in allele frequencies [95-97]. Future work should expand to other HPA axis genes and also take advantage of new exomic data that would allow evaluation of putative functional variants in these gene regions.

CHAPTER 5

A GENOME-WIDE ASSOCIATION STUDY (GWAS) OF SALIVARY CORTISOL CONCENTRATIONS: THE MULTI-ETHNIC STUDY OF ATHEROSCLEROSIS

Introduction

A growing body of work has examined the contribution of stress and related constructs such as allostatic load to various health outcomes [100, 101]. The hormone cortisol is likely to be a key mediator of the stress response that has implications for various physiologic systems (such as the cardiovascular system, immune system, and metabolism) involved in chronic disease [30, 31]. Consequently several studies have focused on understanding both predictors and consequences of cortisol levels [102, 103].

Cortisol concentrations follow a strong daily pattern. They are high upon awakening, reach a maximum concentration approximately half an hour later, and slowly decrease throughout the rest of the day [1-3]. Additionally, cortisol concentrations increase in response to a stressor [32]. Under conditions of chronic stress, prolonged increased concentrations could have detrimental downstream physiological effects.

The nature of within person variability of cortisol necessitates the use of multiple measures over the day to characterize a given person's daily profile. Cortisol concentrations can be measured from multiple biological samples: urine, blood serum, and saliva. Urine and blood collection methods are difficult in population studies because

of the need for repeated collection, making the ability to measure cortisol concentrations in salivary samples an ideal alternative [40, 41]. Salivary cortisol concentrations have been shown to be highly correlated with blood serum cortisol concentrations, with correlations ranging from 0.71-0.96 [2, 40-43].

An individual's daily cortisol response has been shown to be associated with several demographic factors. Age has been shown to be a significant predictor of cortisol concentrations, where concentrations increase with age [13]. Additionally, there are gender differences with respect to salivary cortisol concentrations, with men having significantly higher mean levels than women [14]. Associations with race/ethnicity and socioeconomic factors have also been reported. Flatter declines later in the day (less steep slopes) have been observed in African Americans compared to European Americans. This pattern of flatter afternoon decline has also been shown in lower socioeconomic status groups relative to higher socioeconomic status groups [5-7]. It has been suggested that chronic stress may explain the flatter declines in these individuals [44, 45].

Several population-based studies have linked daily cortisol patterns to health outcomes, including elevated blood pressure, abdominal obesity, and coronary calcification [8-10]. Cortisol concentrations and various features of the cortisol daily profile have also been associated with diabetes mellitus [11] and markers of inflammation [12].

Despite evidence of associations of various risk factors with cortisol, considerable inter-individual variability in cortisol remains unexplained, which has led to increased interest in examining genetic predictors of cortisol phenotypes [15]. In a combined analysis of cortisol heritability in twin studies, basal cortisol had an estimated heritability of 62%

[20]. While the high heritability estimate for basal cortisol concentrations indicates that there is a genetic component to concentration levels, it is unknown what genetic factors are driving this association.

Alone or in interaction with environmental features, genetic factors could contribute to unexplained variability in cortisol concentrations or cortisol responsivity. Most of the work to date has focused on candidate gene associations of cortisol, notably the glucocorticoid receptor gene (*NR3C1*) and the mineralocorticoid receptor gene (*NR3C2*) [21-23].

To our knowledge only one genome-wide association study (GWAS) has been conducted on salivary cortisol. This study examined the area under the cortisol curve in a group of roughly 1,700 European participants from the Rotterdam Study, which found evidence of associations with FKBP5 [104]. Additionally, there has been one published GWAS study of morning serum cortisol concentrations in a group of approximately 500 Hutterites in the western United States, which found evidence of association for two microsatellite markers, one on chromosome 11 and the other on chromosome 14 [105].

We conducted a genome-wide association study of salivary cortisol concentrations among participants of the Multi-Ethnic Study of Atherosclerosis (MESA) Stress Study to identify novel loci associated with multiple features of the diurnal cortisol curve in European Americans (n=170), African Americans (n=215), and Hispanic Americans (n=454) separately. An important strength of the MESA Stress Study is the availability of multiple measures of cortisol over several days. The richness of this data allows for the evaluation of many different cortisol phenotypes, including the cortisol

awakening response, the slope of declines, and area under the cortisol curve. In addition, the availability of measures over multiple days allows improved characterization of the cortisol features. In this paper we use traditional methods for genome-wide association analysis in each ethnic group and use sample size weighted meta-analysis methods for comparing those results across groups [106].

Methods

Study Population

The MESA Stress Study is an ancillary study to the Multi-Ethnic Study of Atherosclerosis (MESA). The MESA study is a longitudinal cohort study focused on investigating the early stages of atherosclerosis. Eligible participants were 45-84 years of age and free from history of cardiovascular disease at the baseline examination (2000-2002) [78]. The MESA Stress Study took place in the context of MESA examinations 3 and 4 conducted between 2004 and 2006, and obtained detailed stress hormone data on a subsample of 1002 MESA participants recruited from the New York and Los Angeles sites. Participants for the MESA Stress Study were African Americans, European Americans, and Hispanic Americans and were enrolled as they presented for follow-up, until approximately 500 participants were recruited from each location. We used individual covariate data from the MESA examination in which an individual's cortisol data collection occurred.

Of the 1002 MESA Stress Study participants, exclusions for 1) raw cortisol data missingness, 2) unavailable genotype or principal component information, 3) no consent

for use of genetic information, and 4) concurrent corticosteroid usage, resulted in a sample size of 839 individuals. The ethnic specific distribution of this sample is as follows: European Americans (n= 170), African Americans (n= 215), Hispanic Americans (n = 454).

Cortisol Sample Collection

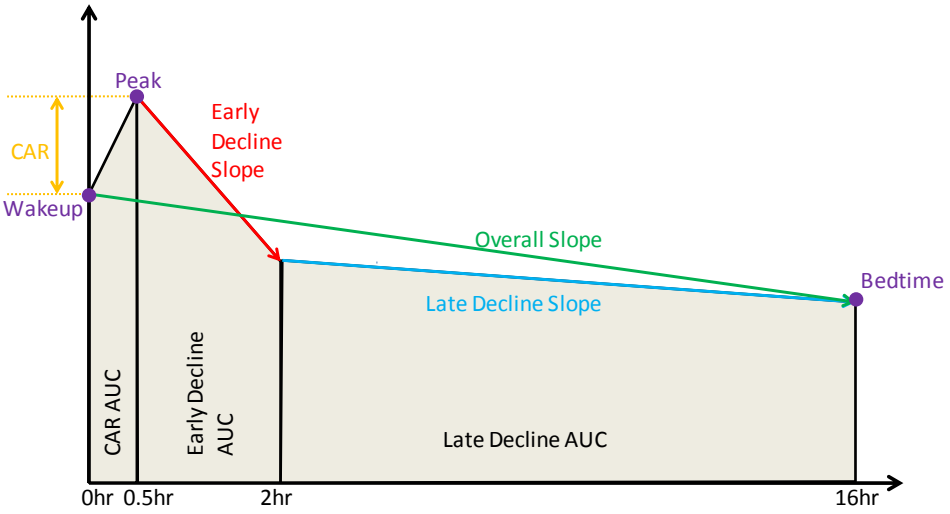
Each MESA Stress Study participant was asked to collect six saliva samples per day at pre-specified times over three consecutive weekdays, for a maximum of 18 samples per participant, using Salivette collection tubes. The samples were collected using the following schedule: sample (1) upon waking and before getting out of bed; (2) 30 minutes later; (3) around 10:00am; (4) around 12:00 noon or before lunch, whichever came first; (5) around 6:00pm or before dinner, whichever came first; (6) just before bed. Because earlier work has shown that the use of a time tracking device improves sample collection compliance [79], each collection tube was equipped with a time tracking device, which recorded the time when the swabs were removed for sample collection.

Cortisol Features

Rather than explore only cortisol concentrations at specific time points, we explored multiple features of the diurnal cortisol cycle (Table 49, Figure 27). Features were selected for investigation because prior work has hypothesized or demonstrated their associations with health risk factors or health outcomes [11, 12, 83, 85]. Features were modeled using all available salivary cortisol data (up to six samples per day collected over three days). Raw cortisol concentrations, measured in nmol/L, were log-transformed to more closely approximate a normal distribution [5, 12, 80].

Table 49: Features of the diurnal cortisol curve. Cortisol concentrations were log-transformed.

	Cortisol Feature	Description
Time points	Wakeup	Average cortisol concentration from wakeup for an individual (Sample 1).
	Bedtime	Average cortisol concentration at bedtime for an individual (Sample 6).
Area	Area under the curve (AUC)	Standardized AUC for the interval 0hr-16hr since wakeup averaged across all days for an individual
Slopes	Cortisol awakening response (CAR)	The average difference in cortisol concentrations between the peak and wakeup measurements (Sample 2 – Sample 1).
	Early Decline Slope (EDSlope)	The slope from 0.5 hours and 2 hours since wakeup pooled across all days for an individual.
	Late Decline Slope (LDSlope)	The slope from 2 hours to 16 hours since wakeup pooled across all days for an individual
	Overall Decline Slope (ODSlope)	The overall decline slope ignoring the peak value from wakeup to bedtime pooled across all days for an individual.



AUC 16hr = CAR AUC + Early Decline AUC + Late Decline AUC

Figure 27: Representation of the diurnal cortisol curve describing our summary features of interest. For these analyses we specifically used Wakeup, Bedtime, Cortisol awakening response (CAR), Area under the curve (AUC) from 0-16 hours, Early Decline Slope, Late Decline Slope, and Overall Decline Slope.

Genetic Data

Genotyping data included both measured and imputed SNPs available through participation in MESA SHARe (SNP Health Association Resource) project. Under the SHARe project, genome-wide genotyping was obtained using the Affymetrix Genome-Wide Human SNP Array 6.0 platform. Imputation to HapMap was completed at the MESA Genetics Centers using the IMPUTE2 [82] program with the following reference panels: the HapMap Phase I and II, the human genome reference sequence (NCBI Build 36). The HapMap project is based on ethnic specific reference panels, composed of the following groups: Yoruba in Ibadan, Nigeria (abbreviation: YRI), Japanese in Tokyo, Japan (abbreviation: JPT), Han Chinese in Beijing, China (abbreviation: CHB), CEPH (Utah residents with ancestry from northern and western Europe) (abbreviation: CEU). Imputation for African Americans and Hispanic Americans was performed using the CEU+YRI+CHB+JPT reference panels (release #22). Imputation for European Americans was performed using only the CEU reference panel (release #24). All imputed and genotyped SNPs were aligned to the "+" strand of the human genome reference sequence (NCBI Build 36). In order to account for population structure and admixture within MESA samples, principal components were extracted from genome-wide data in each ethnic group separately.

Statistical Strategy

Genome-wide association studies (GWAS) were conducted in each ethnic group separately, using SNPTest genetic analysis software (version 2) [89]. Linear regression was used to estimate the additive genetic effect of each single nucleotide polymorphism

(SNP). We used the Frequentist=1 and Method=Expected specifications, which allowed for an additive model of association and use of expected genotype dosages, respectively. The primary model included age and sex as covariates. Ethnic specific principal components were estimated using MESA Classic participants, and outliers were removed. The top 10 principal components were included in the model for African Americans and Hispanics after linear modeling indicated evidence of association between background genetic structure represented by the principal components and features of the cortisol curve. There was limited evidence of association in the European Americans, and as such we did not adjust for PCs. Filtering was performed to remove results for SNPs with minor allele frequency (MAF) less than 5% or with imputation quality (Info) < 0.5. After filtering the ethnic specific number of SNPs considered was 2,516,994 for African Americans, 2,256,299 for European Americans, and 2,258,434 for Hispanic Americans. A p-value$<5x10^{-8}$ was considered the genome-wide significance threshold, and a p-value$<1x10^{-6}$ was considered suggestive of genome-wide significance.

The R (version 2.14.0) [87] package METAL [106] (March 2011 release) was used to conduct a random-effects meta-analysis of the GWAS results from each ethnic group for each cortisol feature. We carried out a sample size weighted analysis based on p-values, where, an overall z-statistic for each SNP is calculated based on the sum of the individual z-statistics from each ethnic group, weighted by the square-root of the number of individuals in each ethnic group.

Results

Basic demographic information on the Stress Study participants is provided in Table 50. Hispanic Americans represented the largest proportion of participants (52.8%), relative to the African Americans (28.6%) and European Americans (18.6%). The gender distribution was fairly equal (52.4% female). Overall, cortisol feature means varied across ethnic groups (Table 51). There was a statistically significant difference in means across the ethnic groups for all cortisol features except CAR.

Table 50: Characteristics of MESA Stress Study participants.

	Frequency (n=1002)
Site	
Columbia	52.2%
UCLA	47.8%
Age	
45-54	29.9%
55-64	27.7%
65-74	30.3%
75-84	12.1%
Race	
European American	18.6%
African American	28.6%
Hispanic American	52.8%
Gender	
Male	47.6%
Female	52.4%
Education Level	
Less than High School	27.0%
Completed High School	20.2%
Some College	29.7%
Bachelor's or higher	23.2%
Income	
< $20,000	29.3%
$20,000-34,999	27.5%
$35,000-$49,999	16.5%
$50,000 or higher	26.8%
Percent Current Smokers	11.3%
Percent Diabetic	13.5%
Body Mass Index (BMI) >=30	36.7%

Table 51: Distributions of cortisol summary features.

Cortisol Feature	European Americans		African Americans		Hispanic Americans		ANOVA
	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	p-value
Wakeup	166	2.58 (0.54)	214	2.38 (0.55)	450	2.38 (0.58)	0.0002
Bedtime	166	0.78 (0.77)	212	0.98 (0.74)	448	0.49 (0.84)	<0.0001
CAR	160	0.45 (0.46)	203	0.35 (0.46)	412	0.37 (0.52)	0.17
AUC	166	1.64 (0.43)	209	1.60 (0.42)	442	1.46 (0.51)	<0.0001
EDSlope	163	-0.53 (0.35)	209	-0.42 (0.44)	433	-0.40 (0.44)	0.003
LDSlope	164	-0.12 (0.06)	211	-0.10 (0.06)	447	-0.13 (0.06)	<0.0001
ODSlope	169	-0.12 (0.07)	214	-0.10 (0.06)	452	-0.12 (0.06)	<0.0001

Cortisol concentrations (nmol/L) were log-transformed and combined across the three days of collection to create each feature. SD = Standard Deviation.

GWAS Results

Graphical representations of the GWAS results in each ethnic group for the seven cortisol features are represented in the Appendix Materials by Q-Q Plots (Appendix Figures A24-A44) and Manhattan Plots (Appendix Figures A45-A51). One locus reached genome-wide significance ($p<5x10^{-8}$) and an additional 17 loci reached a suggestive level of association ($p<1x10^{-6}$). In the African American analysis of the cortisol awakening response (CAR) (Figure 28), two SNPs on chromosome 1 passed the genome-wide significance threshold ($p = 9.42x10^{-10}$ and $p = 1.76x10^{-09}$). As these two SNPs were in high linkage disequilibrium ($r^2 >0.8$), they represent one signal. Table 52 lists the most strongly associated SNPs ($p<1x10^{-6}$) for each feature, ranked in order of significance. When a locus identified multiple SNPs that were in high linkage disequilibrium ($r^2<0.8$), only one SNP from that locus is presented in Table 52.

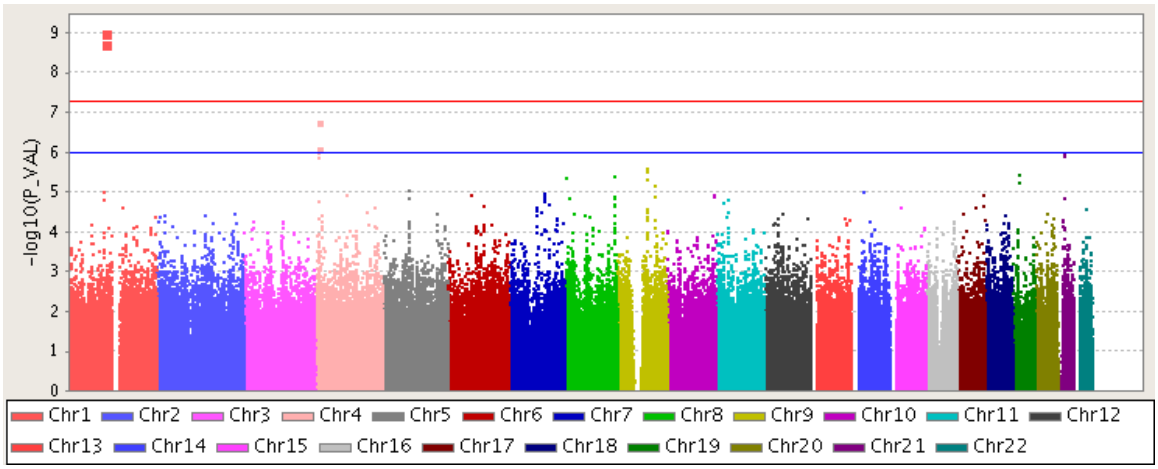


Figure 28: GWAS results for cortisol awakening response (CAR) in African Americans. Individual SNP p-values located above the blue line are suggestive of genome-wide significance $-\log_{10}(p<1x10^{-6})$. Individual SNP p-values located above the red line have reached genome-wide significance $-\log_{10}(p<5x10^{-8})$.

Table 52: Most strongly associated SNPs (at $1x10^{-6}$ or less) for each cortisol summary feature. Only single SNPs in regions where there was strong LD are displayed.

Cortisol Feature	Race	SNP	CHR	POS	Effect Allele	Freq	Beta	p-value	Nearest Gene(s) ±50kb
Wakeup	CAU	rs7701889	5	126043888	G	0.9331	0.566	5.28E-07	----
Wakeup	AFA	rs7217942	17	6087394	T	0.1020	-0.377	5.96E-07	----
Wakeup	HIS	rs1320124	2	240941561	G	0.7973	0.260	8.23E-07	----
Bedtime	AFA	rs10448237	9	36838997	G	0.4490	-0.348	3.56E-07	*PAX5**
Bedtime	CAU	rs715262	16	59384147	G	0.3090	0.430	6.73E-07	----
CAR	**AFA**	**rs290827**	**1**	**97246872**	**G**	**0.3634**	**-0.278**	**9.42E-10**	**----**
CAR	CAU	rs17582686	12	40248012	T	0.8836	-0.419	1.13E-07	*PDZRN4**
CAR	AFA	rs7687462	4	11544862	T	0.6108	-0.236	1.81E-07	----
AUC	HIS	rs10490591	2	146720035	G	0.0956	-0.287	5.65E-07	----
EDSlope	CAU	rs3791682	2	211957951	T	0.6881	0.196	9.47E-07	*ERBB4**
LDSlope	AFA	rs1052199	9	76951677	G	0.5600	0.034	2.73E-07	*OSTF1**
LDSlope	CAU	rs3734068	11	129320230	G	0.0828	-0.067	3.51E-07	*PRDM10**
LDSlope	HIS	rs6557164	6	152044625	T	0.1566	-0.034	7.59E-07	*ERS1*
ODSlope	AFA	rs8086616	18	22049890	G	0.0689	0.049	8.89E-08	*TAF4B*; *PSMA8*
ODSlope	AFA	rs11024350	11	17613262	T	0.0813	0.046	2.99E-07	----
ODSlope	HIS	rs2827248	21	22420514	G	0.8496	0.025	3.48E-07	----
ODSlope	CAU	rs12526290	6	160529342	G	0.0709	-0.056	3.74E-07	*SLC22A1*; *SLC22A2*
ODSlope	AFA	rs2503663	6	92811107	T	0.5157	0.026	6.82E-07	----

CHR: Chromosome. POS: Position. Freq: Effect allele frequency. * = SNP located within gene. Bold = genome-wide significant hit.

Meta-analysis Results

After meta-analysis none of the GWAS SNPs reached genome-wide significance, although there were promising results in four of the seven cortisol summary features at $p < 5x10^{-6}$. Table 53 shows the meta-analysis results for the four features (AUC, Bedtime, CAR, and Wakeup) in which the SNP was available in at least two of the three ethnic groups. These SNPs are located across the genome, on chromosomes 8, 11, 12, and 15. The direction of effect for each meta-analysis locus was consistent across the three ethnic groups. The majority of individual p-values in each group were less than 0.05, indicating a relatively high level of agreement on the association with the cortisol feature, even though the MAF often varied. The sample size for the meta-analysis SNPs is less than the 839 total individuals either due to missingness for the phenotype or SNP missingness.

In order to determine whether the SNP effects were similar across ethnic groups, we performed an assessment of heterogeneity between the three ethnic groups in METAL (Table 54). I^2 is the percentage of effect size variability due to true differences in effects across the ethnic groups [107]. Three of the I^2 values were 0 or close to 0 indicating that variability in effect estimates at that SNP is due to sampling error. Although two of the I^2 values suggested that 32% and 58% of the variability between groups is due to true heterogeneity, p-values were >0.05 indicating no evidence of statistically significant heterogeneity of effects.

Table 53: METAL meta-analysis results.

Cortisol Feature	SNP	Effect Allele	CHR	POS	Mean Freq	Freq by Group	N	Z-score	METAL P-value	GWAS p-value	Direction of Effect	Nearest Gene(s) ±50kb
AUC	rs2111270	A	12	12473003	0.23		761	5.004	5.63E-07		+++	*LOH12CR1**
AFA						0.06				0.0007		
CAU						0.27				0.0014		
HIS						0.29				0.0127		
Bedtime	rs2410611	T	8	19292641	0.29		761	-4.953	7.32E-07		---	*SH2D4A; CSGALNACT1*
AFA						0.19				0.0013		
CAU						0.19				0.7694		
HIS						0.37				1.33E-05		
CAR	rs7174390	T	15	55813252	0.41		761	-4.99	6.05E-07		---	*GCOM1*; *GRINL1A*
AFA						0.28				0.0065		
CAU						0.45				0.0092		
HIS						0.46				0.0009		
Wakeup	rs7929069	A	11	134437775	0.18		756	-4.921	8.59E-07		---	
AFA						0.23				0.0115		
CAU						0.16				0.0008		
HIS						0.16				0.0037		
Wakeup	rs6473381	A	8	83760172	0.34		761	-5.008	5.51E-07		---	
AFA						0.26				0.0067		
CAU						0.39				0.3101		
HIS						0.36				1.48E-05		

CHR: Chromosome. POS: Position. * = SNP located within gene.

Table 54: METAL assessment of heterogeneity.

Cortisol Feature	SNP	Effect Allele	Mean Freq (SE)	Het I^2	Het χ^2 (Df)	Het P-Value
AUC 16	rs2111270	A	0.23 (0.10)	32.5	2.965 (2)	0.227
Bedtime	rs2410611	T	0.29 (0.09)	58.7	4.848 (2)	0.089
CAR	rs7174390	T	0.41 (0.07)	0	0.265 (2)	0.876
Wakeup	rs7929069	A	0.18 (0.03)	0	1.764 (2)	0.414
Wakeup	rs6473381	A	0.34 (0.05)	2.7	2.056 (2)	0.358

Discussion

This study examined the key question of whether genetic factors contribute to inter-individual variation in cortisol profiles by assessing the genome-wide associations of seven cortisol features. We identified novel genetic loci associated with the features of the diurnal cortisol curve in a multi-ethnic study population, with 1 SNP (rs290827) being significantly associated with the CAR in African Americans (β= -0.278, $p=9.43x10^{-10}$). There is no evidence of association for this locus in the other two ethnic groups. Additionally, the direction of effect for this SNP is not consistent for European Americans (β=0.041, p=0.45) and Hispanic Americans (β=0.060, p=0.10) compared to the African American findings. However, for the African Americans the 'G' allele was the minor allele (frequency=0.36), while for the European Americans and the Hispanic Americans the 'G' allele was the major allele (EA frequency=0.63, HIS frequency=0.52).

Beyond the genome-wide significant result in the African Americans, we had many suggestive loci at $p<1x10^{-6}$ across ethnic groups for multiple cortisol features. Additionally, there were six gene regions that were implicated either by the GWAS analyses or the meta-analyses: *PAX5*, *PDZRN4*, *ERBB4*, *OSTF1*, *PRDM10*, and

LOH12CR1. Three of the six genes, *PAX5*, *ERBB4*, and *PRDM10*, are involved in central nervous system development.

PAX5, a B-cell specific transcription factor, has implications for midbrain and cerebellum development [108]. As a B-cell transcription factor, *PAX5* has also been associated with a number of cancers, specifically leukemias and lymphomas [109]. *ERBB4* (also designated as HER4), a cell surface and epidermal growth factor receptor, is involved in the central nervous system through regulating GABA concentrations [110]. *ERBB4* has previously been associated with schizophrenia [111], as well as breast and ovarian cancers [112, 113]. *PRDM10* (also designated as Tristanin), a zinc-finger transcription factor, impacts the central nervous system through dendrite initiation [114]. The remaining genes have less clear functions: *PDZRN4* is involved in ubiquitin-protein ligase activity, *OSTF1* is an osteoclast-stimulating factor, and *LOH12CR1* has a loss of heterozygosity [115]. While the precise function of *LOH12CR1* is not clear, it is located near a tumor suppressor locus that has been associated with leukemia [116, 117].

There have only been two other genome-wide studies of cortisol features. In a salivary cortisol AUC GWAS [104], an association was found with four novel SNPs (rs9470080, rs9394309, rs7748266, rs1360780) in the *FKBP5* gene and cortisol AUC measured in the Rotterdam Study. *FKBP5* is located on chromosome 6. Our analyses were unable to confirm associations at any of these loci. Our p-values for the associations with these loci among MESA Stress Study participants for AUC range from 0.27 to 0.91. In the Rotterdam analysis the four *FKBP5* SNPs, which are in high linkage disequilibrium, have a reported estimate of effect of -0.55, corresponding to a decrease in

AUC. These SNPs in our GWAS had mixed directions of effect (0.02, 0.006, -0.02, -0.03). In a GWAS study of morning serum cortisol measured in a group of Hutterites from Utah [105], two genome-wide significant loci were identified in women, one on chromosome 11 (D11S1981, p=0.000084) and the other on chromosome 14 (D14S74, p=0.000091). Their genome-wide screen used a total of 891 microsatellite markers and 412 intragenic SNPs for analysis. The two identified loci were microsatellite markers and were not available in our HapMap imputed genome-wide dataset; therefore possible replication cannot be assessed.

In examining genetic effects across multiple ethnic groups there are a few concerns. First, differences in population structure make it difficult to compare the effects for single SNPs since they are likely to have differing linkage disequilibrium patterns with functional (unmeasured) variants underlying the association signal [95-97]. Differences in allele frequencies across groups can also make individual SNP replication difficult, since it shifts the power to detect effects of similar size given the same alpha criteria [118, 119]. The differences in underlying genetic architecture and in allele frequencies across populations make it unlikely that SNPs will replicate across ethnic groups. These intricacies may explain why our most significant result did not replicate in the other groups.

A second concern is that complex traits like cortisol are a complex function of both genes and environment that differ in distribution across racial/ethnic groups [120]. Given these differences, new approaches that allow for the assessment of entire genes or a particular set of SNPs in genomic regions is likely to be a better analytic approach.

Since humans are 99% genetically similar, gene structure (exon and intron organization) is not likely to differ across ethnic groups. Therefore, the overall impact of mutations (positive of deleterious) of a gene or gene region is the better unit of inference to compare across ethnic groups. New methods such as the sequence kernel association test (SKAT) [84] have been developed to address this analytic need and will be a part of future studies.

With unique phenotype information also come limitations. First, while the cortisol phenotype data itself is very rich, the restricted sample sizes when stratified by ethnic group specific analysis allows for limited power to detect association. Even if a pooled analysis had been performed, a GWAS on ~800 individuals would be under powered. Compared to other recently published GWAS studies where samples sizes are on the order of tens or hundreds of thousands [121-124], the small sample sizes for these analyses are unlikely to reach the standard genome-wide significance threshold ($p<5x10^{-8}$). However, the novel cortisol features measured and estimated from three consecutive days greatly reduces phenotypic variability compared to a single cortisol measurement, thus increasing power. Additionally, the availability of these detailed cortisol features in a multi-ethnic sample is an unusual opportunity and allowed for a unique genetic epidemiology opportunity in spite of the power limitations.

A second limitation is the use of HapMap imputed variants, which are not functional SNPs. Utilizing rare-variant analyses or focusing on functional SNPs is an important direction for future work [125]. Exome and whole genome sequencing is revealing a greater level of rare variants than had been previously expected [126]. There are

a variety of algorithms available which aid in the selection of functional SNPs by bioinformatic prediction of the impacts of different mutations on protein or mRNA function [127-129].

Thirdly, compliance with cortisol sampling protocols is necessary for estimating reliable cortisol features [79, 94]. Compliance with taking samples at the requested times was greatest for wakeup (Sample 1) and bedtime (Sample 6) collections, and poorest during the middle of the day. Compliance within 10 minutes of the sampling protocol was 68% for Sample 1, 53% for Sample 2, 57% for Sample 3, 43% for Sample 4, 44% for Sample 5, and 75% for Sample 6. The deviations from protocol in the middle of the day would be particularly important in the estimations of CAR, AUC, EDSlope and LDSlope.

For the calculation of the cortisol features, we were particularly concerned with compliance of Sample 2, which was designed to assess the maximum cortisol concentration throughout the day. For the calculation of CAR, if Sample 2 was taken more than 30 minutes late (an hour or more after wakeup) it was defined as missing, since it would no longer be representative of the maximum cortisol concentration given that concentrations decrease rapidly after peaking.

Despite the limitations, this work has a number of notable strengths. First is the richness of the cortisol data. The availability of multiple samples per day and the repeated cortisol measurement across three days allowed for the characterization of multiple cortisol features, including time of day variables, an area measure, and multiple slopes features. Second is the availability of these unique cortisol features in a multi-

ethnic population. Given that previous cortisol GWAS work has been done only in European populations, this work represents the first genome-wide significant finding among African Americans for the cortisol awakening response, as well as suggestive evidence in African Americans in five of the seven cortisol features. This work also presents the first suggestive evidence of genetic associations in Hispanic Americans in four of the seven features. With the continuous advances in genomic technologies and the combined consortia efforts to pool data, the GWAS presented here represent a first step in understanding the genetic architecture of cortisol concentrations.

CHAPTER 6

CONCLUSION

Overall, this dissertation illustrates that genetic analyses across ethnic groups can provide new insights into the role of genes in cortisol features and their relationship with chronic disease risk factors. Chapter 3 examined the relationship between variation in six stress response candidate genes and features of the diurnal cortisol curve. Chapter 4 investigated gene-by-cortisol interactions and their associations with anthropometric, metabolic, and inflammatory outcomes. Chapter 5 examined genome-wide association studies (GWAS) of seven cortisol features and found suggestive evidence of association across cortisol features and ethnic groups. In this conclusion chapter, I review the role of gene-based associations, gene-by-cortisol interactions, and GWAS in better understanding the complex influence of stress on common chronic disease risk factors.

Gene-based Association

Unlike GWAS studies which are hypothesis generating, candidate gene approaches are driven by *a priori* knowledge of individual genes or pathways of interest [130]. Using this hypothesis driven approach restricts the number of tests and

results in increased relative power compared to GWAS studies adjusted for multiple comparisons [131].

We were interested in examining the relationship between six selected stress response candidate genes and the cortisol features. Rather than assessing each SNP in the six gene regions individually, Chapter 4 of the dissertation employs a gene-based association approach to examine the cumulative evidence of multiple markers in a gene region across ethnic groups in predicting cortisol features, which was possible through the use of the innovative SKAT statistical methodologies [84]. This gene-based analytic approach is especially useful in investigating effects across ethnic groups as it allows us to address two concerns that arise in individual SNP based analyses: that individual SNPs may not replicate across ethnic groups due to 1) differences in underlying patterns of linkage disequilibrium and 2) differences in allele frequencies [118,119].

While there are several other gene-based association methods available (e.g. Cohort Allelic Sum Test (CAST) [91], Weighted Sum Statistic (WSS) [92], C-alpha test [93]), SKAT has several advantages. First, SKAT is a more powerful method, even when sample sizes are small (n=500) [84], which is of particular importance given the small ethnic group specific sample sizes for this dissertation. Second, SKAT allows for the individual variant effects to vary from mean zero in either direction, and does not assume that all variants have similar direction or magnitude of effect. Thirdly, it allows for the adjustment of covariates. SKAT additionally allows for the assessment of common variants by implementing an unweighted linear kernel, which fit our needs since we are using HapMap imputed genome-wide data.

In the stress response gene analyses for Chapter 4, three of the six gene regions had significant (p-value < 0.05) associations with cortisol features in at least one ethnic group: *ADRA2A, ADRB2,* and *SLC6A4*. In the meta-analyses across the three ethnic groups, *ADRA2A* was a suggestive predictor (p-value < 0.1) of four out of the seven cortisol features, *ADRB2* was a suggestive predictor of CAR, and *SLC6A4* was a significant predictor (p-value < 0.05) of EDSlope.

There are three ways in which future gene region based work can expand upon the approach used in Chapter 4. First is the expansion to other candidate genes or genomic regions of interest in physiological pathways. This expanded list could include other genes of the HPA axis, such as *FKBP5*, or length polymorphisms, such as the "s" allele and "l" allele of *SLC6A4*. There are also a number of bioinformatics tools designed to visualize biological pathways as a means of selecting candidate genes [132-134] or to select potentially disease related genes [135-139]. As the analyses presented for the gene regions in this dissertation were limited to common tagging SNPs, utilizing rare-variant analyses or focusing on functional SNPs is another important direction for future work. There are a variety of algorithms available which aid in the selection of functional SNPs by predicting the impacts of different SNPs at a given locus [127-129].

A third avenue for future studies would rely on the ability of regional analysis programs, such as SKAT, to incorporate multiple genes or environments into the analytic framework. This option would then allow for assessing the cumulative effect of multiple genes in a physiologic pathway on an outcome, which may be a more reasonable theoretical approach for complex diseases rather than assuming genes are working in

isolation. This integration of multiple gene regions would then also allow for the evaluation of epistasis, which could be operating across the different segments of the stress responsivity pathway.

The gene-based analyses presented in this dissertation are novel in their ability to examine the variation in multiple gene regions across ethnic groups in predicting cortisol features, which was possible through the use of the innovative SKAT methodologies as well has the unique, highly detailed cortisol phenotype information. The gene-based analytic approach allowed us to address the concern that individual SNPs may not replicate across ethnic groups due to differences in underlying patterns of linkage disequilibrium or to differences in allele frequencies [93, 132, 133], by examining a larger analysis unit which is unlikely to differ across populations. The restriction to common variants in these analyses is additionally a unique implementation of the SKAT framework.

Gene-by-Environment Interaction Studies

Complex and chronic disease are inherently due to a mixture of genetic and environmental effects, which may vary across backgrounds. As such, assessment of gene-by-environment interactions is of particular interest for these outcomes. In essence, gene-by-environment interactions assess whether the presence of an environmental factor alters the relationship between a genotype and outcome, or conversely, that the genotype is modifying the relationship between the environment and the outcome. In Chapter 5 of this dissertation we investigated how six stress response genes impact the relationship

between cortisol features and chronic disease risk factors. Cortisol features were the environmental factor, conceptualized as the internal stress environment.

Chapter 5 of this dissertation examined the key question of whether gene-by-cortisol interactions influence chronic disease risk factors, which was implemented using a two-step approach. We first looked for evidence of gene-by-cortisol interaction by implementing an extension to the SKAT framework, which is a variance component score test that assumes the coefficients of the gene-by-environment interaction term to be random effects [98]. Since SKAT does not provide estimates of specific SNP-by-cortisol interaction parameters (i.e. magnitude or direction of effect) we also used traditional least squares regression approaches to estimate SNP-by-cortisol interaction terms when there was evidence of a significant gene-level interaction (interaction p-value < 0.05) in more than one ethnic group for a given gene-by-cortisol relationship.

We found six stress response gene-by-cortisol interactions in multiple ethnic groups in predicting chronic disease outcomes: one for ln(Glucose), three for ln(IL-6 + 1), and two for ln(TNF-α +1). There was not significant evidence of interaction in more than one ethnic group in predicting average BMI. The interaction between *SLC6A4* and ODSlope in predicting average fasting Glucose had the most significant evidence across the ethnic groups, with interaction p-values ≤ 0.01 for African Americans and European Americans. Of the SNPs in the *SLC6A4* region with interaction p-values <0.05, four were identical across the two groups (rs2020942, rs140701, rs2066713, and rs3794808), and their direction of effect was consistent. Of the remaining five instances of gene-by-cortisol interaction in two ethnic groups, one had SNPs (rs1991795 , rs10477394 ,

rs6580582) that were comparable in direction of effect (*ADRB2*-by-Bedtime interaction on ln(IL-6 + 1)), one (*NR3C2*-by-EDSlope on ln(IL-6 + 1)) had a single SNP (rs17024681) in both ethnic groups, but it varied in direction of effect, and three where no SNPs were consistent across the ethnic groups.

The gene-level evidence of interaction with cortisol features for *NR3C1* and *SLC6A4* are particularly interesting. *NR3C1* is a glucocorticoid receptor that is occupied when cortisol concentrations are high [36]. This interaction indicates that when the body is under stress and cortisol concentrations are high, variation in *NR3C1* influences the downstream effect these heighted cortisol levels on metabolic and inflammatory risk factors. While previous work has shown that polymorphisms in the promoter region of *SLC6A4* are associated with CAR [62], to my knowledge there is no published information on associations with a measure of overall decline (ODSlope). As *SLC6A4* is hypothesized to impact the stress response, the interaction between variations in the gene region and stress hormone levels having downstream effects on metabolism is also important. Our findings that the relationship between cortisol features and chronic disease risk factors in some ethnic groups is influenced by variation in stress response genes supports the notion that variations in the HPA axis stress responsivity pathway may be relevant in some groups but not others in explaining variations in chronic disease burden. These findings are only preliminary evidence and should be followed up by future work both due to the small ethnic group specific sample sizes in these interaction analyses and additionally to assess the physiological implications of the interactions.

Gene-by-environment interactions can be difficult to dissect. Given the underlying heterogeneity in both the genes and the environment, it can be difficult to isolate which variable is influencing the other. The hypothesis for Chapter 5 was that polymorphisms in the stress response genes were modifying the association between cortisol features and chronic disease outcomes. The results from these analyses indicate that we found evidence of statistical interaction, but they do not provide direct evidence of biological interaction. Future stress and cortisol work should strive to understand the underlying biological processes of the statistical interactions identified.

Concerns regarding the biological process through which interactions exhibit their effects in gene-by-environment interactions can be addressed in a variety of ways. While the stress response genes for the analyses in this dissertation were selected due to their relationship with cortisol and the stress response, using algorithms for pathway analysis [135-139] to elucidate additional genes along the cortisol metabolic pathway or other pathways involved in stress response may shed light on physiologically relevant mechanisms. Similarly, the evaluation of functional SNPs in the gene-based approach also extends to interaction studies, where the implications of differences in amino acid substitutions may allow for easier interpretation than the evaluation of tagging SNPs in non-coding regions. Also paralleling the future directions for the gene-based analyses, the evidence of gene-by-cortisol interaction may encourage the future investigation of gene-gene interactions and epistasis in understanding the intricate relationships between stress response genes and cortisol features. Additionally, functional biology studies may

be utilized in the evaluation of the physiological implications of any identified interactions (e.g. gene-by-environment, gene-gene, epistatic).

This work makes several unique contributions to the cortisol literature. The genes for this analysis were selected under the construct of HPA axis stress responsivity and they were tested for interaction with a biomarker of the stress response, cortisol. To my knowledge, evaluation of gene interaction with the internal stress environment on downstream chronic disease risk factors has not previously been published. Not only is each individual unique in their genetic and cortisol features, but an individual's ability to respond to or cope with heightened stress as a result of activation of the stress responsivity pathway is likely to also be unique, and therefore may have different influences on chronic disease states. Given the complex and multifaceted nature of these analyses, replication of the gene-by-cortisol interaction findings is necessary. We have approved proposals for replication in the CARDIA Study for the gene-by-cortisol interaction studies of anthropometric, metabolic, and inflammatory factors.

Genome-wide Association Studies of Complex Traits

GWAS takes advantage of high throughput data to assess millions of individual markers, and are often hypothesis generating. While there has been much progress in the GWAS field, a number of challenges remain [140]. For example, individual variants identified by GWAS only explain a small fraction of phenotype associations, and need to be followed up with additional analyses (e.g. sequencing data, animal models, in vitro studies) to confirm potentially causal associations [141]. In addition, studying multi-ethnic

populations using GWAS approaches can be challenging as outcome prevalence estimates differ across groups as do background genetic and environmental factors [97]. In order to address complex differences across ethnic groups in GWAS studies, methods for the consideration of genetic ancestry have been beneficial in admixed populations, such as African Americans and Hispanic Americans [142-144]. Future GWAS work in the MESA Study should examine methods used to consider admixtures; there are several such methods available [145-152].

The power to detect genome-wide associations ($p<5x10^{-8}$) can be a difficult threshold to attain. Compared to other recently published GWAS studies where samples sizes are on the order of tens or hundreds of thousands [121-124], the GWASs in this dissertation are very underpowered given the small ethnic group specific sample sizes (maximum sample size of 454 Hispanic Americans). However, the novel cortisol features measured and estimated from three consecutive days greatly reduces the variability compared to a single cortisol measurement, thus increasing power.

In the analysis of complex traits, given often small sample sizes and modest effect estimates, replication of findings is extremely important [153,154]. The GWAS investigations in this dissertation were unable to replicate the findings of previous cortisol genome-wide association studies [104,105]. In the previous morning serum cortisol GWAS of 504 Hutterites in the western United States [105], two genome-wide significant loci were identified in women, one on chromosome 11 (D11S1981, p=0.000084) and the other on chromosome 14 (D14S74, p=0.000091). Their genome-wide screen used a total of 891 microsatellite markers and 412 intragenic SNPs for analysis. These two loci were not

available in our HapMap imputed genome-wide dataset and therefore possible replication cannot be assessed until 1000Genomes data is available. In the salivary cortisol AUC GWAS of ~1,700Europeans from the Rotterdam Study [104], an association was found with four novel SNPs (rs9470080, rs9394309, rs7748266, rs1360780) in the *FKBP5* gene and cortisol AUC. *FKBP5* is located on chromosome 6. Approximately 2,800 participants from the Whitehall II Study were available for GWAS replication. However, none of the SNPs of interest from the Rotterdam sample replicated in Whitehall II. Our analyses were also unable to confirm associations at any of these loci. Our p-values for the associations with these loci among MESA Stress Study participants for AUC range from 0.27 to 0.91. In the Rotterdam analysis the four *FKBP5* SNPs, which are in high linkage disequilibrium, have a reported estimate of effect of -0.55, corresponding to a decrease in AUC. These SNPs in our GWAS had mixed directions of effect (0.02, 0.006, -0.02, -0.03) compared to those reported by Velders *et al*. This lack of significant replication may be attributable to the underpowered nature of our GWAS sample or differences in gene and/or environmental factors.

Most successful GWAS center on consortia efforts, which allows for increased power with increasing sample size and opportunities for replication. The Cortisol Network (CORNET) is a recently established cortisol consortium, which is comprised mainly of European individuals. There is an agreement with the Cortisol Network (CORNET) to follow up GWAS features of interest from this dissertation in their consortium.

There are also additional opportunities for replication in other ethnic groups. We have an approved proposal for GWAS replication in the Coronary Artery Risk Development in Young Adults (CARDIA) Study, which has cortisol samples available for European Americans and African Americans. While they only collected cortisol samples for one day, they did collect multiple samples across that day, which should allow for calculation of the cortisol features used in this dissertation. Permission has also been obtained to perform replication in the Mid-Life in the U.S. (MIDUS) Study, which has salivary cortisol data available at four time points throughout the day in a sample of African Americans and European Americans. There has not been a replication sample identified for the Hispanic Americans.

The GWAS work presented in this dissertation makes several substantive contributions to the cortisol literature. To my knowledge this is the first population-based, multi-ethnic assessment of the role genetics plays in cortisol profiles. Second, due to the detailed and repeated assessment of cortisol concentrations across multiple days among the MESA Stress Study participants, a range of time point, area, and slopes features were available for examination that have never been considered for GWAS studies. The use of multiple slope features extends upon previous cortisol work which is typically limited to wakeup, bedtime, or cortisol awakening response variables; these slope features are likely not to be available in the larger consortia efforts.

Limitations

Two main limitations of the work presented in this dissertation, power and sample size, have already been addressed. An additional limitation Previous work has shown that compliance with cortisol sampling protocols is necessary for estimating reliable cortisol features [79, 94]. Compliance with taking samples at the requested times was greatest for wakeup (Sample 1) and bedtime (Sample 6) collections, and poorest during the middle of the day. Compliance within 10 minutes sampling protocol was 68% for Sample 1, 53% for Sample 2, 57% for Sample 3, 43% for Sample 4, 44% for Sample 5, and 75% for Sample 6. The deviations from protocol in the middle of the day would be particularly important in the estimations of CAR, AUC, EDSlope and LDSlope.

For the calculation of the cortisol features, there was particular concern with compliance of Sample 2, which was designed to assess the maximum cortisol concentration throughout the day. For the calculation of CAR, if Sample 2 was taken more than 30 minutes late (an hour or more after wakeup) it was defined as missing, since it would no longer be representative of the maximum cortisol concentration given that concentrations decrease rapidly after peaking. Overall, the combining of cortisol features across multiple days reduces the variability in cortisol feature estimation compared to single cortisol measurements. Future cortisol studies need to consider multiple measures per day across multiple days in establishing reliable cortisol estimates.

The MESA has already begun assessing cortisol feature reliability over time as a future extension of this dissertation work. The MESA Stress Study II has recently completed a second round of cortisol sample collection. In this follow-up, cortisol

concentrations were assessed eight times per day over two consecutive days. The change in protocol was designed to aid in the estimation of slope features by having more samples taken throughout the latter portion of the day, and to improve compliance by reducing the burden of sample collection on participants to only two days instead of three. Once these data are available it will be possible to estimate the stability of the cortisol profile within and across individuals over time, and the repeat assessment over two time points may add to feature reliability.

Concluding Remarks

In this dissertation, cortisol was considered an embodiment of external experience. An internal climate that is different across racial/ethnic groups, ages, gender, socioeconomic status, and chronic stress states. Overall, this dissertation illustrates that genetic analyses across ethnic groups can provide new insights into the role of genes in cortisol features and their relationship with chronic disease risk factors.

APPENDIX

Table A1: Chromosomal location of each stress response gene.

Stress Response Gene	Location
Alpha-2A-adrenergic receptor gene (*ADRA2A*)	10q24-q26
Beta-2-adrenergic receptor gene (*ADRB2*)	5q31-q32
Glucocorticoid receptor gene (*NR3C1*)	5q31.3
Mineralocorticoid receptor gene (*NR3C2*)	4q31.1
Serotonin transporter gene (*SLC6A4*)	17q11.1-q12
Tyrosine hydroxylase gene (*TH*)	11p5.5

Table A2: Start and end positions of genes of interest, ±5kb, and over size of gene regions.

Stress Response Gene	Start Position (bp)	End Position (bp)	Overall Size (bases)
ADRA2A	112,821,000	112,836,000	15,000
ADRB2	148,181,000	148,194,000	13,000
NR3C1	142,632,000	142,770,000	138,000
NR3C2	149,214,000	149,589,000	375,000
SLC6A4	25,544,000	25,592,000	48,000
TH	2,136,000	2,155,000	19,000

bp= Chromosomal base pair.

Table A3: Number of SNPs with a minor allele frequency (MAF) of 5% or greater in each gene region for each ethnic group.

	Number of SNPs in each gene region with MAF > 0.05		
Stress Response Gene	African Americans	European Americans	Hispanic Americans
ADRA2A	12	10	11
ADRB2	77	52	62
NR3C1	52	58	56
NR3C2	358	322	327
SLC6A4	30	22	24
TH	16	15	15

Table A4: Gene-level main effect and meta-analysis results for *NR3C1*.

		SKAT		MetaSKAT
Outcome	**Race**	**Q**	**p-value**	**p-value**
AUC	AFA	1259.67	0.41	
	EA	2371.85	0.22	
	HIS	3958.20	0.32	0.31
Bedtime	AFA	1542.37	0.32	
	EA	1384.40	0.46	
	HIS	2938.33	0.48	0.52
CAR	AFA	987.60	0.51	
	EA	349.75	0.91	
	HIS	448.75	0.98	0.98
EDSlope	AFA	402.14	0.91	
	EA	290.52	0.94	
	HIS	5266.97	0.19	0.58
LDSlope	AFA	668.66	0.74	
	EA	165.66	0.98	
	HIS	6503.15	0.13	0.33
ODSlope	AFA	655.27	0.75	
	EA	2246.58	0.25	
	HIS	6590.86	0.13	0.18
Wakeup	AFA	747.40	0.69	
	EA	1964.35	0.30	
	HIS	5233.70	0.21	0.30

Table A5: Gene-level main effect and meta-analysis results for *NR3C2*.

		SKAT		MetaSKAT
Outcome	**Race**	**Q**	**p-value**	**p-value**
AUC	AFA	7461.46	0.84	
	EA	7933.53	0.59	
	HIS	23224.90	0.39	0.66
Bedtime	AFA	9086.91	0.65	
	EA	12372.80	0.19	
	HIS	23331.31	0.40	0.37
CAR	AFA	8208.19	0.72	
	EA	9515.75	0.39	
	HIS	17985.57	0.58	0.66
EDSlope	AFA	6265.41	0.94	
	EA	6846.74	0.71	
	HIS	9028.81	0.99	0.99
LDSlope	AFA	10173.25	0.50	
	EA	9639.74	0.39	
	HIS	16030.00	0.78	0.74
ODSlope	AFA	8485.26	0.74	
	EA	11348.85	0.27	
	HIS	25992.34	0.30	0.37
Wakeup	AFA	10352.27	0.49	
	EA	8841.67	0.49	
	HIS	14351.73	0.86	0.83

Table A6: Gene-level main effect and meta-analysis results for *TH*.

		SKAT		MetaSKAT
Outcome	**Race**	**Q**	**p-value**	**p-value**
AUC	AFA	304.63	0.66	
	EA	147.07	0.78	
	HIS	1041.61	0.35	0.64
Bedtime	AFA	300.57	0.67	
	EA	161.15	0.75	
	HIS	1143.65	0.31	0.56
CAR	AFA	318.50	0.6	
	EA	494.36	0.22	
	HIS	1274.97	0.22	0.26
EDSlope	AFA	882.09	0.06*	
	EA	726.70	0.12	
	HIS	329.19	0.89	0.27
LDSlope	AFA	809.60	0.08*	
	EA	409.65	0.31	
	HIS	428.25	0.81	0.51
ODSlope	AFA	342.74	0.59	
	EA	349.34	0.40	
	HIS	1412.79	0.22	0.29
Wakeup	AFA	284.32	0.72	
	EA	525.84	0.22	
	HIS	1631.95	0.16	0.20

* $p<0.1$

Table A7: Individual SNP-by-Bedtime interactions in *ADRB2* in predicting ln(IL-6 + 1), with an interaction p-value < 0.05, in order of significance by ethnic group.

	Cortisol				Coded		SNP Effects		Cortisol Effects		Interaction Effects	
Outcome	**Feature**	**Gene**	**Race**	**SNP**	**Allele**	**Freq**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**
ln(IL-6 + 1)	Bedtime	*ADRB2*	AFA	rs1991795	T	0.34	0.15	0.02	0.10	0.11	-0.15	0.005
				rs10477394	T	0.35	0.15	0.03	0.11	0.11	-0.15	0.006
				rs10053209	G	0.44	0.16	0.02	0.09	0.16	-0.15	0.01
				rs10463408	G	0.40	0.15	0.03	0.08	0.21	-0.14	0.01
				rs10064479	T	0.60	0.15	0.03	0.08	0.21	-0.14	0.01
				rs877741	T	0.40	0.15	0.03	0.08	0.22	-0.14	0.01
				rs877743	G	0.60	0.15	0.03	0.08	0.22	-0.14	0.01
				rs246503	G	0.14	0.25	0.03	0.02	0.65	-0.19	0.03
				rs6580582	T	0.31	0.12	0.10	0.06	0.32	-0.12	0.03
				rs246502	T	0.14	0.25	0.03	0.02	0.66	-0.18	0.04
			HIS	rs1991795	T	0.57	0.04	0.17	0.12	0.002	-0.08	0.01
				rs10477394	T	0.57	0.04	0.15	0.12	0.002	-0.08	0.01
				rs1347110	G	0.66	0.04	0.22	0.10	0.004	-0.07	0.02
				rs6580582	T	0.53	0.05	0.10	0.12	0.005	-0.07	0.03
				rs30297	G	0.92	0.01	0.93	0.06	0.02	-0.20	0.03

Table A8: Individual SNP-by-Bedtime interactions in *NR3C2* in predicting ln(IL-6 + 1), with an interaction p-value < 0.05, in order of significance by ethnic group.

	Cortisol				Coded		SNP Effects		Cortisol Effects		Interaction Effects	
Outcome	**Feature**	**Gene**	**Race**	**SNP**	**Allele**	**Freq**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**
ln(IL-6 + 1)	Bedtime	*NR3C2*	AFA	rs11724292	C	0.06	0.31	0.06	0.00	0.97	-0.36	0.006
				rs3846312	G	0.78	-0.11	0.19	-0.13	0.01	0.19	0.01
				rs4835131	G	0.30	-0.12	0.10	-0.14	0.01	0.16	0.01
				rs6850597	G	0.30	-0.12	0.10	-0.14	0.01	0.16	0.01
				rs2356374	G	0.72	-0.12	0.10	-0.14	0.02	0.16	0.01
				rs9762822	T	0.82	-0.13	0.14	-0.11	0.03	0.18	0.02
				rs4579099	T	0.27	-0.13	0.07	-0.12	0.02	0.15	0.02
				rs3846310	G	0.89	-0.06	0.59	-0.09	0.04	0.22	0.02
				rs3846322	T	0.83	0.21	0.02	0.02	0.72	-0.18	0.02
				rs7698917	T	0.81	-0.15	0.07	-0.11	0.03	0.17	0.02
				rs3846320	G	0.88	-0.17	0.14	-0.08	0.08	0.18	0.02
				rs13118475	T	0.28	0.06	0.39	-0.11	0.03	0.14	0.03
				rs2883930	G	0.48	-0.14	0.06	-0.17	0.03	0.12	0.04
				rs10018805	C	0.86	-0.21	0.05	-0.10	0.06	0.15	0.04
			EA	rs6535583	T	0.27	-0.09	0.17	-0.03	0.62	0.17	0.01
				rs6857011	T	0.27	-0.09	0.17	-0.03	0.62	0.17	0.01
				rs6856424	G	0.73	-0.09	0.17	-0.03	0.62	0.17	0.01
				rs12499208	T	0.43	0.17	0.01	0.17	0.01	-0.15	0.01
				rs1879827	T	0.75	0.14	0.07	0.15	0.003	-0.20	0.02
				rs6535580	G	0.75	0.15	0.06	0.15	0.003	-0.20	0.02
				rs6535581	G	0.75	0.15	0.06	0.15	0.003	-0.20	0.02

	rs10031194	T	0.25	0.15	0.06	0.15	0.003	-0.20	0.02
	rs17483687	C	0.14	-0.10	0.22	0.03	0.51	0.17	0.03
	rs1403142	G	0.57	0.15	0.03	0.19	0.003	-0.13	0.03
	rs1403143	T	0.57	0.15	0.03	0.19	0.003	-0.13	0.03
	rs12506077	T	0.57	0.15	0.03	0.19	0.003	-0.13	0.03
	rs7693171	T	0.43	0.15	0.03	0.19	0.003	-0.13	0.03
	rs1040288	G	0.43	0.15	0.03	0.19	0.003	-0.13	0.03
	rs7687754	G	0.57	0.15	0.03	0.19	0.003	-0.13	0.03
	rs7665528	G	0.57	0.15	0.03	0.19	0.003	-0.13	0.03
	rs4835128	T	0.43	0.15	0.03	0.19	0.003	-0.13	0.03
	rs4835488	T	0.57	0.15	0.03	0.19	0.003	-0.13	0.03
	rs6855032	G	0.43	0.15	0.03	0.19	0.003	-0.13	0.03
	rs1879829	T	0.56	0.02	0.79	0.20	0.002	-0.14	0.04
	rs2293162	T	0.06	0.28	0.07	0.12	0.01	-0.29	0.04

Table A9: Individual SNP-by-EDSlope interactions in *NR3C2* in predicting ln(IL-6 + 1), with an interaction p-value < 0.05, in order of significance by ethnic group.

	Cortisol				Coded		SNP Effects		Cortisol Effects		Interaction Effects	
Outcome	**Feature**	**Gene**	**Race**	**SNP**	**Allele**	**Freq**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**
ln(IL-6 + 1)	EDSlope	*NR3C2*	AFA	rs1490453	G	0.62	0.22	0.0003	0.02	0.85	0.27	0.003
				rs7688969	C	0.62	0.22	0.0005	0.04	0.73	0.26	0.004
				rs17024681	C	0.75	0.30	2.85E-05	0.09	0.30	0.28	0.004
			EA	rs13109933	T	0.50	0.28	0.0006	-0.22	0.14	0.36	0.003
				rs1996025	T	0.91	0.53	0.001	0.00	1.00	0.70	0.005
				rs6831212	T	0.56	0.26	0.002	-0.19	0.17	0.35	0.006
				rs1512341	T	0.56	0.26	0.002	-0.19	0.17	0.35	0.006
				rs1512327	G	0.56	0.26	0.002	-0.19	0.17	0.35	0.006
				rs17582031	T	0.12	-0.32	0.01	0.29	0.02	-0.47	0.01
				rs7698307	T	0.12	-0.32	0.01	0.29	0.02	-0.47	0.01
				rs6840422	G	0.88	-0.32	0.01	0.29	0.02	-0.47	0.01
				rs2048546	C	0.88	-0.32	0.01	0.29	0.02	-0.47	0.01
				rs17485033	G	0.88	-0.32	0.01	0.29	0.02	-0.47	0.01
				rs16998733	T	0.12	-0.32	0.01	0.29	0.02	-0.47	0.01
				rs13123626	T	0.67	-0.20	0.03	0.38	0.009	-0.33	0.02
				rs1994624	G	0.62	-0.21	0.02	0.40	0.009	-0.33	0.02
				rs2137331	T	0.44	0.25	0.003	-0.16	0.26	0.31	0.02
				rs13133379	T	0.22	-0.25	0.05	0.23	0.04	-0.44	0.03
				rs17484839	T	0.88	-0.29	0.02	0.25	0.03	-0.40	0.04
				rs17024681	C	0.88	-0.29	0.02	0.25	0.03	-0.40	0.04
				rs17484873	T	0.89	-0.29	0.02	0.25	0.03	-0.40	0.04

	rs6834935	T	0.88	-0.29	0.02	0.25	0.03	-0.40	0.04
	rs2063555	C	0.12	-0.28	0.03	0.24	0.04	-0.39	0.04
	rs7686433	G	0.87	-0.31	0.02	0.25	0.03	-0.37	0.04
	rs10519963	G	0.88	-0.31	0.02	0.25	0.03	-0.37	0.04
	rs17484783	T	0.88	-0.31	0.02	0.25	0.03	-0.37	0.04

Table A10: Individual SNP-by-Bedtime interactions in *NR3C1* in predicting ln(TNF-α + 1), with an interaction p-value < 0.05, in order of significance by ethnic group.

	Cortisol				Coded		SNP Effects		Cortisol Effects		Interaction Effects	
Outcome	**Feature**	**Gene**	**Race**	**SNP**	**Allele**	**Freq**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**
ln(TNF-α + 1)	Bedtime	*NR3C1*	EA	rs10482689	T	0.18	0.31	0.01	0.17	0.07	-0.39	0.001
				rs10482642	G	0.82	0.31	0.01	0.17	0.07	-0.39	0.001
				rs10515521	G	0.82	0.31	0.01	0.17	0.07	-0.39	0.001
				rs17339831	G	0.82	0.31	0.01	0.17	0.07	-0.39	0.001
				rs11740792	G	0.18	0.31	0.01	0.17	0.07	-0.39	0.001
				rs10482633	T	0.82	0.31	0.01	0.17	0.07	-0.39	0.001
				rs4128428	T	0.82	0.31	0.01	0.17	0.07	-0.39	0.001
				rs258750	G	0.29	0.21	0.04	0.19	0.08	-0.31	0.002
				rs190488	T	0.71	0.19	0.05	0.18	0.08	-0.30	0.003
				rs258813	G	0.71	0.19	0.05	0.18	0.08	-0.30	0.003
				rs852977	G	0.29	0.19	0.05	0.18	0.08	-0.30	0.003
				rs860457	T	0.71	0.19	0.05	0.18	0.08	-0.30	0.003
				rs852982	G	0.71	0.19	0.05	0.18	0.08	-0.30	0.003
				rs2963149	T	0.29	0.19	0.05	0.18	0.08	-0.30	0.003
				rs2918417	T	0.29	0.19	0.05	0.18	0.08	-0.30	0.003
				rs2918416	T	0.29	0.19	0.05	0.18	0.08	-0.30	0.003
				rs1866388	G	0.29	0.19	0.05	0.18	0.08	-0.30	0.004
				rs10052957	G	0.71	0.19	0.05	0.18	0.08	-0.30	0.004
				rs17287758	G	0.82	0.29	0.03	0.13	0.07	-0.34	0.008
			HIS	rs258763	T	0.60	0.14	0.004	0.10	0.04	-0.14	0.004
				rs10041520	T	0.42	0.13	0.008	0.10	0.06	-0.12	0.009

	rs4634384	T	0.57	0.13	0.008	0.10	0.06	-0.12	0.009
	rs6877893	G	0.43	0.14	0.005	0.10	0.06	-0.12	0.009
	rs852980	G	0.42	0.13	0.006	0.10	0.06	-0.12	0.01
	rs33383	T	0.58	0.13	0.006	0.10	0.06	-0.12	0.01
	rs33388	T	0.58	0.13	0.008	0.10	0.07	-0.12	0.01
	rs10482682	T	0.28	0.07	0.22	0.06	0.14	-0.13	0.01
	rs17209237	G	0.20	0.12	0.05	0.05	0.22	-0.16	0.01
	rs17287745	G	0.31	0.09	0.10	0.07	0.13	-0.13	0.01
	rs258747	G	0.42	0.13	0.007	0.09	0.08	-0.12	0.01
	rs10482634	G	0.18	0.09	0.17	0.04	0.28	-0.16	0.02
	rs17399352	T	0.82	0.09	0.17	0.04	0.28	-0.16	0.02
	rs17209251	G	0.18	0.11	0.09	0.05	0.25	-0.16	0.02
	rs4986593	G	0.18	0.11	0.10	0.04	0.28	-0.16	0.02
	rs10482655	T	0.82	0.09	0.17	0.04	0.29	-0.16	0.02
	rs11750172	G	0.18	0.09	0.18	0.04	0.29	-0.16	0.02
	rs17339455	T	0.82	0.09	0.18	0.04	0.29	-0.16	0.02
	rs11745958	T	0.18	0.09	0.18	0.04	0.29	-0.16	0.02
	rs9324916	G	0.82	0.09	0.18	0.04	0.29	-0.16	0.02
	rs17209258	G	0.17	0.10	0.13	0.04	0.29	-0.15	0.02

Table A11: Individual SNP-by-Wakeup interactions in *NR3C2* in predicting ln(TNF-α + 1), with an interaction p-value < 0.05, in order of significance by ethnic group.

	Cortisol				Coded		SNP Effects		Cortisol Effects		Interaction Effects	
Outcome	**Feature**	**Gene**	**Race**	**SNP**	**Allele**	**Freq**	**B**	**p-value**	**B**	**p-value**	**B**	**p-value**
ln(TNF-α + 1)	Wakeup	*NR3C2*	EA	rs3843410	T	0.76	1.22	0.006	0.01	0.92	-0.42	0.01
				rs4835491	G	0.76	1.18	0.007	0.01	0.94	-0.41	0.01
				rs12499208	T	0.43	0.82	0.006	0.08	0.50	-0.28	0.01
				rs1403142	G	0.57	0.80	0.008	0.07	0.55	-0.27	0.02
				rs1403143	T	0.57	0.80	0.008	0.07	0.55	-0.27	0.02
				rs12506077	T	0.57	0.80	0.008	0.07	0.55	-0.27	0.02
				rs7693171	T	0.43	0.80	0.008	0.07	0.55	-0.27	0.02
				rs1040288	G	0.43	0.80	0.008	0.07	0.55	-0.27	0.02
				rs7687754	G	0.57	0.80	0.008	0.07	0.55	-0.27	0.02
				rs7665528	G	0.57	0.80	0.008	0.07	0.55	-0.27	0.02
				rs4835128	T	0.43	0.80	0.008	0.07	0.55	-0.27	0.02
				rs4835488	T	0.57	0.80	0.008	0.07	0.55	-0.27	0.02
				rs6855032	G	0.43	0.80	0.008	0.07	0.55	-0.27	0.02
				rs10050229	G	0.33	-0.55	0.08	-0.31	0.006	0.28	0.02
				rs4835493	T	0.23	0.98	0.03	0.02	0.83	-0.36	0.03
				rs2272089	T	0.78	1.06	0.02	-0.02	0.82	-0.37	0.03
				rs6844155	G	0.78	1.06	0.02	-0.02	0.82	-0.37	0.03
				rs7694064	G	0.78	1.06	0.02	-0.02	0.82	-0.37	0.03
				rs7694200	G	0.78	1.06	0.02	-0.02	0.82	-0.37	0.03
				rs7694706	G	0.22	1.06	0.02	-0.02	0.82	-0.37	0.03
				rs3843411	T	0.23	0.97	0.03	0.01	0.93	-0.35	0.03
				rs3857079	T	0.70	0.91	0.03	0.00	0.96	-0.32	0.03

	HIS	rs3916013	T	0.89	0.69	0.02	-0.02	0.76	-0.30	0.01
		rs13116347	G	0.11	0.43	0.07	0.001	0.99	-0.24	0.01
		rs2356374	G	0.94	0.64	0.09	-0.05	0.35	-0.35	0.02
		rs6850597	G	0.06	0.63	0.09	-0.04	0.38	-0.34	0.02
		rs17484357	G	0.12	0.56	0.03	-0.006	0.91	-0.23	0.03
		rs7691250	G	0.12	0.53	0.04	-0.008	0.89	-0.23	0.03
		rs10010766	T	0.40	0.40	0.03	0.07	0.41	-0.16	0.03
		rs4579099	T	0.06	0.61	0.13	-0.05	0.33	-0.33	0.04
		rs4835131	G	0.06	0.57	0.12	-0.05	0.37	-0.31	0.04

Table A12: Two degree of freedom test of individual SNP-by-Bedtime interactions in predicting ln(IL-6 + 1). SNPs ordered by significance in the additive model for the individual SNP-by-Bedtime interactions in Table A7.

	Cortisol				Coded		Minor	SNP Effects		Interaction Effects	
Outcome	**Feature**	**Gene**	**Race**	**SNP**	**Allele**	**Freq**	**Alleles**	**B**	**p-value**	**B**	**p-value**
ln(IL-6 + 1)	Bedtime	*ADRB2*	AFA	rs1991795	T	0.34	1	0.15	0.21	-0.13	0.19
							2	0.29	0.03	-0.31	0.005
				rs10477394	T	0.35	1	0.15	0.20	-0.15	0.13
							2	0.30	0.03	-0.31	0.006
				rs10053209	G	0.44	1	0.20	0.09	-0.16	0.10
							2	0.31	0.03	-0.30	0.01
				rs10463408	G	0.40	1	0.17	0.14	-0.14	0.14
							2	0.28	0.05	-0.28	0.02
				rs10064479	T	0.60	1	0.17	0.14	-0.14	0.14
							2	0.28	0.05	-0.28	0.02
				rs877741	T	0.40	1	0.17	0.14	-0.14	0.15
							2	0.28	0.05	-0.28	0.02
				rs877743	G	0.60	1	0.17	0.14	-0.14	0.15
							2	0.28	0.05	-0.28	0.02
				rs246503	G	0.14	1	0.25	0.02	-0.18	0.04
							2	-0.44	0.91	0.20	0.94
				rs6580582	T	0.31	1	0.05	0.67	-0.04	0.70
							2	0.27	0.10	-0.27	0.02
				rs246502	T	0.14	1	0.25	0.02	-0.18	0.05
							2	-0.44	0.91	0.20	0.94
				rs17108773	G	0.42	1	-0.08	0.48	0.09	0.35
							2	-0.31	0.04	0.24	0.05

	HIS	rs1991795	T	0.57	1	0.04	0.50	-0.10	0.06
					2	0.09	0.15	-0.15	0.02
		rs10477394	T	0.57	1	0.04	0.47	-0.10	0.06
					2	0.09	0.13	-0.15	0.02
		rs1347110	G	0.66	1	0.02	0.70	-0.08	0.11
					2	0.09	0.16	-0.14	0.05
		rs6580582	T	0.53	1	0.05	0.36	-0.09	0.16
					2	0.10	0.10	-0.15	0.02
		rs30297	G	0.92	1	0.01	0.93	0.06	0.02
					2	0.01	1E-05	-0.20	0.03
		rs17778143	T	0.65	1	0.02	0.74	-0.07	0.18
					2	0.09	0.16	-0.12	0.07
		rs11742884	T	0.66	1	0.02	0.74	-0.07	0.18
					2	0.09	0.17	-0.12	0.07

Table A13: Two degree of freedom test of individual SNP-by-Bedtime interactions in predicting ln(IL-6 + 1). SNPs ordered by significance in the additive model for the individual SNP-by-Bedtime interactions in Table A8.

	Cortisol				Coded		Minor	SNP Effects		Interaction Effects	
Outcome	**Feature**	**Gene**	**Race**	**SNP**	**Allele**	**Freq**	**Alleles**	**B**	**p-value**	**B**	**p-value**
lm(IL-6 + 1)	Bedtime	*NR3C2*	AFA	rs11724292	C	0.06	1	0.31	0.06	0.002	0.97
							2	0.01	0.02	-0.36	0.006
				rs3846312	G	0.78	1	-0.20	0.07	0.26	0.004
							2	-0.05	0.83	0.21	0.34
				rs4835131	G	0.30	1	-0.11	0.32	0.17	0.06
							2	-0.24	0.13	0.30	0.03
				rs6850597	G	0.30	1	-0.11	0.32	0.17	0.06
							2	-0.24	0.13	0.30	0.03
				rs2356374	G	0.72	1	-0.10	0.35	0.15	0.09
							2	-0.25	0.12	0.32	0.03
				rs9762822	T	0.82	1	-0.20	0.08	0.26	0.00
							2	-0.05	0.82	-0.27	0.38
				rs4579099	T	0.27	1	-0.12	0.27	0.14	0.12
							2	-0.27	0.09	0.30	0.04
				rs3846310	G	0.89	1	-0.11	0.37	0.25	0.01
							2	2.27	0.11	-4.57	0.15
				rs3846322	T	0.83	1	0.21	0.07	-0.20	0.05
							2	0.46	0.06	-0.22	0.44
				rs7698917	T	0.81	1	-0.26	0.02	0.25	0.01
							2	-0.05	0.81	0.00	0.99
				rs3846320	G	0.88	1	-0.15	0.29	0.14	0.26
							2	-0.27	0.54	0.38	0.09

	rs13118475	T	0.28	1	0.16	0.15	0.07	0.42
				2	0.05	0.74	0.31	0.03
	rs2883930	G	0.48	1	-0.13	0.29	0.15	0.15
				2	-0.29	0.06	0.25	0.04
	rs10018805	C	0.86	1	-0.17	0.14	0.07	0.44
				2	-0.03	0.98	0.29	0.46
	rs4835133	G	0.19	1	0.14	0.23	-0.15	0.12
				2	0.45	0.08	-0.22	0.45
EA	rs6535583	T	0.27	1	-0.21	0.03	0.30	0.0009
				2	-0.11	0.42	0.21	0.17
	rs6857011	T	0.27	1	-0.21	0.03	0.30	0.0009
				2	-0.11	0.42	0.21	0.17
	rs6856424	G	0.73	1	-0.21	0.03	0.30	0.00
				2	-0.11	0.42	0.21	0.17
	rs12499208	T	0.43	1	0.10	0.31	-0.12	0.19
				2	0.37	0.008	-0.32	0.01
	rs1879827	T	0.75	1	0.12	0.27	-0.20	0.04
				2	0.30	0.17	-0.37	0.18
	rs6535580	G	0.75	1	0.12	0.26	-0.19	0.05
				2	0.31	0.17	-0.37	0.19
	rs6535581	G	0.75	1	0.12	0.26	-0.19	0.05
				2	0.31	0.17	-0.37	0.19
	rs10031194	T	0.25	1	0.12	0.26	-0.19	0.05
				2	0.31	0.17	-0.37	0.19
	rs17483687	C	0.14	1	-0.20	0.06	0.21	0.03
				2	0.06	0.80	0.22	0.31

rs1403142	G	0.57	1	0.05	0.58	-0.04	0.65
			2	0.37	0.01	-0.32	0.01
rs1403143	T	0.57	1	0.05	0.58	-0.04	0.65
			2	0.37	0.01	-0.32	0.01
rs12506077	T	0.57	1	0.05	0.58	-0.04	0.65
			2	0.37	0.01	-0.32	0.01
rs7693171	T	0.43	1	0.05	0.58	-0.04	0.65
			2	0.37	0.01	-0.32	0.01
rs1040288	G	0.43	1	0.05	0.58	-0.04	0.65
			2	0.37	0.01	-0.32	0.01
rs7687754	G	0.57	1	0.05	0.58	-0.04	0.65
			2	0.37	0.01	-0.32	0.01
rs7665528	G	0.57	1	0.05	0.58	-0.04	0.65
			2	0.37	0.01	-0.32	0.01
rs4835128	T	0.43	1	0.05	0.58	-0.04	0.65
			2	0.37	0.01	-0.32	0.01
rs4835488	T	0.57	1	0.05	0.58	-0.04	0.65
			2	0.37	0.01	-0.32	0.01
rs6855032	G	0.43	1	0.05	0.58	-0.04	0.65
			2	0.37	0.01	-0.32	0.01
rs1879829	T	0.56	1	-0.10	0.32	-0.05	0.56
			2	0.12	0.40	-0.37	0.01
rs2293162	T	0.06	1	0.28	0.07	0.12	0.009
			2	0.01	0.02	-0.29	0.04
rs1879828	T	0.45	1	-0.09	0.36	-0.06	0.53
			2	0.14	0.33	-0.34	0.02

Table A14: Two degree of freedom test of individual SNP-by-EDSlope interactions in predicting ln(IL-6 + 1). SNPs ordered by significance in the additive model for the individual SNP-by-Bedtime interactions in Table A9.

	Cortisol				Coded		Minor	SNP Effects		Interaction Effects	
Outcome	**Feature**	**Gene**	**Race**	**SNP**	**Allele**	**Freq**	**Alleles**	**B**	**p-value**	**B**	**p-value**
ln(IL-6 + 1)	EDSlope	*NR3C2*	AFA	rs1490453	G	0.62	1	0.24	0.009	0.24	0.11
							2	0.43	0.001	0.53	0.005
				rs7688969	C	0.62	1	0.23	0.01	0.26	0.09
							2	0.42	0.001	0.52	0.005
				rs17024681	C	0.75	1	0.23	0.007	0.29	0.04
							2	0.84	4.8E-05	0.76	0.002
			CAU	rs13109933	T	0.50	1	0.38	0.007	0.24	0.26
							2	0.56	0.0005	0.72	0.003
				rs1996025	T	0.91	1	0.53	0.001	-0.0004	1.00
							2	0.01	0.002	0.70	0.005
				rs6831212	T	0.56	1	0.45	0.001	0.47	0.02
							2	0.51	0.003	0.72	0.005
				rs1512341	T	0.56	1	0.45	0.001	0.47	0.02
							2	0.51	0.003	0.72	0.005
				rs1512327	G	0.56	1	0.45	0.001	0.47	0.02
							2	0.51	0.003	0.72	0.005
				rs17582031	T	0.12	1	-0.33	0.013	0.29	0.02
							2	-0.20	0.63	-0.47	0.01
				rs7698307	T	0.12	1	-0.33	0.01	0.29	0.02
							2	-0.20	0.63	-0.47	0.01
				rs6840422	G	0.88	1	-0.33	0.01	0.29	0.02

			2	-0.20	0.63	-0.47	0.01
rs2048546	C	0.88	1	-0.33	0.01	0.29	0.02
			2	-0.20	0.63	-0.47	0.01
rs17485033	G	0.88	1	-0.33	0.01	0.29	0.02
			2	-0.20	0.63	-0.47	0.01
rs16998733	T	0.12	1	-0.33	0.01	0.29	0.02
			2	-0.20	0.63	-0.47	0.01
rs13123626	T	0.67	1	-0.20	0.12	-0.41	0.05
			2	-0.45	0.03	-0.66	0.03
rs1994624	G	0.62	1	-0.20	0.12	-0.38	0.07
			2	-0.45	0.03	-0.69	0.02
rs2137331	T	0.44	1	0.41	0.003	0.34	0.10
			2	0.50	0.003	0.66	0.01
rs13133379	T	0.22	1	-0.26	0.12	-0.44	0.06
			2	-0.48	0.27	-0.77	0.59
rs17484839	T	0.88	1	-0.30	0.03	0.25	0.03
			2	-0.20	0.64	-0.41	0.04
rs17024681	C	0.88	1	-0.30	0.03	0.25	0.03
			2	-0.20	0.64	-0.41	0.04
rs17484873	T	0.89	1	-0.30	0.03	0.25	0.03
			2	-0.20	0.64	-0.41	0.04
rs6834935	T	0.88	1	-0.30	0.03	0.25	0.03
			2	-0.20	0.64	-0.41	0.04
rs2063555	C	0.12	1	-0.29	0.03	0.24	0.04
			2	-0.19	0.65	-0.40	0.04
rs7686433	G	0.87	1	-0.33	0.01	-0.47	0.02

				2	0.006	0.99	0.57	0.59
	rs10519963	G	0.88	1	-0.33	0.01	-0.47	0.02
				2	0.006	0.99	0.57	0.59
	rs17484783	T	0.88	1	-0.33	0.01	-0.47	0.02
				2	0.006	0.99	0.57	0.59

Table A15: Two degree of freedom test of individual SNP-by-Bedtime interactions in predicting ln(TNF-α + 1). SNPs ordered by significance in the additive model for the individual SNP-by-Bedtime interactions in Table A10.

	Cortisol				Coded		Minor	SNP Effects		Interaction Effects	
Outcome	**Feature**	**Gene**	**Race**	**SNP**	**Allele**	**Freq**	**Alleles**	**B**	**p-value**	**B**	**p-value**
ln(TNF-α + 1)	Bedtime	*NR3C1*	EA	rs10482689	T	0.18	1	0.28	0.04	-0.39	0.001
							2	-0.80	0.68	1.55	0.59
				rs10482642	G	0.82	1	0.28	0.04	-0.39	0.001
							2	-0.80	0.68	1.55	0.59
				rs10515521	G	0.82	1	0.28	0.04	-0.39	0.001
							2	-0.80	0.68	1.55	0.59
				rs17339831	G	0.82	1	0.28	0.04	-0.39	0.001
							2	-0.80	0.68	1.55	0.59
				rs11740792	G	0.18	1	0.28	0.04	-0.39	0.001
							2	-0.80	0.68	1.55	0.59
				rs10482633	T	0.82	1	0.28	0.04	-0.39	0.001
							2	-0.80	0.68	1.55	0.59
				rs4128428	T	0.82	1	0.28	0.04	-0.39	0.001
							2	-0.80	0.68	1.55	0.59
				rs258750	G	0.29	1	0.33	0.02	-0.35	0.004
							2	0.29	0.24	-0.63	0.04
				rs190488	T	0.71	1	0.31	0.03	-0.33	0.006
							2	0.28	0.26	-0.62	0.04
				rs258813	G	0.71	1	0.31	0.03	-0.33	0.006
							2	0.28	0.26	-0.62	0.04
				rs852977	G	0.29	1	0.31	0.03	-0.33	0.006
							2	0.28	0.26	-0.62	0.04

	rs860457	T	0.71	1	0.31	0.03	-0.33	0.006
				2	0.28	0.26	-0.62	0.04
	rs852982	G	0.71	1	0.31	0.03	-0.33	0.006
				2	0.28	0.26	-0.62	0.04
	rs2963149	T	0.29	1	0.31	0.03	-0.33	0.006
				2	0.28	0.26	-0.62	0.04
	rs2918417	T	0.29	1	0.31	0.03	-0.33	0.006
				2	0.28	0.26	-0.62	0.04
	rs2918416	T	0.29	1	0.31	0.03	-0.33	0.006
				2	0.28	0.26	-0.62	0.04
	rs1866388	G	0.29	1	0.30	0.03	-0.35	0.005
				2	0.26	0.30	-0.54	0.11
	rs10052957	G	0.71	1	0.30	0.03	-0.35	0.005
				2	0.26	0.30	-0.54	0.11
	rs17287758	G	0.82	1	0.24	0.08	-0.34	0.008
				2	2.23	0.62	-2.59	0.68
HIS	rs258763	T	0.60	1	0.18	0.01	-0.14	0.06
				2	0.26	0.01	-0.26	0.007
	rs10041520	T	0.42	1	0.16	0.02	-0.10	0.20
				2	0.24	0.02	-0.25	0.01
	rs4634384	T	0.57	1	0.16	0.02	-0.10	0.20
				2	0.24	0.02	-0.25	0.01
	rs6877893	G	0.43	1	0.18	0.01	-0.09	0.21
				2	0.25	0.01	-0.25	0.01
	rs852980	G	0.42	1	0.17	0.02	-0.09	0.23

			2	0.25	0.02	-0.24	0.01
rs33383	T	0.58	1	0.17	0.02	-0.09	0.23
			2	0.25	0.02	-0.24	0.01
rs33388	T	0.58	1	0.16	0.02	-0.11	0.15
			2	0.24	0.02	-0.24	0.01
rs10482682	T	0.28	1	0.09	0.16	-0.14	0.05
			2	0.06	0.65	-0.21	0.08
rs17209237	G	0.20	1	0.15	0.03	-0.13	0.07
			2	0.09	0.73	-0.34	0.21
rs17287745	G	0.31	1	0.12	0.07	-0.13	0.07
			2	0.10	0.46	-0.22	0.07
rs258747	G	0.42	1	0.17	0.02	-0.11	0.15
			2	0.24	0.02	-0.23	0.02
rs10482634	G	0.18	1	0.11	0.10	-0.14	0.07
			2	0.01	0.97	-0.29	0.29
rs17399352	T	0.82	1	0.11	0.10	-0.14	0.07
			2	0.01	0.97	-0.29	0.29
rs17209251	G	0.18	1	0.14	0.05	-0.13	0.07
			2	0.02	0.95	-0.30	0.29
rs4986593	G	0.18	1	0.14	0.05	-0.13	0.09
			2	0.02	0.95	-0.29	0.29
rs10482655	T	0.82	1	0.12	0.10	-0.14	0.08
			2	0.01	0.97	-0.29	0.29
rs11750172	G	0.18	1	0.11	0.10	-0.13	0.08
			2	0.01	0.97	-0.29	0.29
rs17339455	T	0.82	1	0.11	0.10	-0.13	0.08

				2	0.01	0.97	-0.29	0.29
	rs11745958	T	0.18	1	0.11	0.10	-0.13	0.08
				2	0.01	0.97	-0.29	0.29
	rs9324916	G	0.82	1	0.11	0.10	-0.13	0.08
				2	0.01	0.97	-0.29	0.29
	rs17209258	G	0.17	1	0.13	0.07	-0.13	0.10
				2	-0.01	0.96	-0.28	0.32

Table A16: Two degree of freedom test of individual SNP-by-Bedtime interactions in predicting ln(TNF-α + 1). SNPs ordered by significance in the additive model for the individual SNP-by-Bedtime interactions in Table A12.

	Cortisol				Coded		Minor	SNP Effects		Interaction Effects	
Outcome	**Feature**	**Gene**	**Race**	**SNP**	**Allele**	**Freq**	**Alleles**	**B**	**p-value**	**B**	**p-value**
ln(TNF-α + 1)	Wakeup	*NR3C2*	EA	rs3843410	T	0.76	1	1.53	0.002	-0.54	0.004
							2	0.14	0.94	-0.02	0.97
				rs4835491	G	0.76	1	1.39	0.006	-0.48	0.01
							2	0.97	0.56	-0.34	0.56
				rs12499208	T	0.43	1	0.60	0.21	-0.22	0.23
							2	1.74	0.006	-0.59	0.01
				rs1403142	G	0.57	1	0.54	0.26	-0.19	0.29
							2	1.72	0.007	-0.58	0.02
				rs1403143	T	0.57	1	0.54	0.26	-0.19	0.29
							2	1.72	0.007	-0.58	0.02
				rs12506077	T	0.57	1	0.54	0.26	-0.19	0.29
							2	1.72	0.007	-0.58	0.02
				rs7693171	T	0.43	1	0.54	0.26	-0.19	0.29
							2	1.72	0.007	-0.58	0.02
				rs1040288	G	0.43	1	0.54	0.26	-0.19	0.29
							2	1.72	0.007	-0.58	0.02
				rs7687754	G	0.57	1	0.54	0.26	-0.19	0.29
							2	1.72	0.007	-0.58	0.02
				rs7665528	G	0.57	1	0.54	0.26	-0.19	0.29
							2	1.72	0.007	-0.58	0.02
				rs4835128	T	0.43	1	0.54	0.26	-0.19	0.29
							2	1.72	0.007	-0.58	0.02

	rs4835488	T	0.57	1	0.54	0.26	-0.19	0.29
				2	1.72	0.007	-0.58	0.02
	rs6855032	G	0.43	1	0.54	0.26	-0.19	0.29
				2	1.72	0.007	-0.58	0.02
	rs10050229	G	0.33	1	-0.69	0.15	0.33	0.07
				2	-1.02	0.15	0.54	0.05
	rs4835493	T	0.23	1	1.10	0.02	-0.40	0.02
				2	0.16	0.94	-0.13	0.86
	rs2272089	T	0.78	1	1.39	0.007	-0.50	0.01
				2	-0.29	0.88	0.15	0.83
	rs6844155	G	0.78	1	1.39	0.007	-0.50	0.01
				2	-0.29	0.88	0.15	0.83
	rs7694064	G	0.78	1	1.39	0.007	-0.50	0.01
				2	-0.29	0.88	0.15	0.83
	rs7694200	G	0.78	1	1.39	0.007	-0.50	0.01
				2	-0.29	0.88	0.15	0.83
	rs7694706	G	0.22	1	1.39	0.007	-0.50	0.01
				2	-0.29	0.88	0.15	0.83
	rs3843411	T	0.23	1	1.09	0.02	-0.38	0.03
				2	0.14	0.95	-0.12	0.87
	rs3857079	T	0.70	1	1.12	0.02	-0.38	0.04
				2	0.34	0.82	-0.14	0.78
HIS	rs3916013	T	0.89	1	0.81	0.009	-0.34	0.007
				2	-0.02	0.99	-0.06	0.94
	rs13116347	G	0.11	1	0.39	0.13	-0.21	0.05

				2	0.79	0.60	-0.57	0.32
	rs2356374	G	0.94	1	0.81	0.08	-0.45	0.03
				2	3.57	0.15	-1.37	0.10
	rs6850597	G	0.06	1	0.74	0.10	-0.41	0.03
				2	3.57	0.15	-1.37	0.10
	rs17484357	G	0.12	1	0.65	0.02	-0.26	0.02
				2	0.53	0.80	-0.39	0.72
	rs7691250	G	0.12	1	0.61	0.02	-0.25	0.02
				2	0.52	0.81	-0.38	0.73
	rs10010766	T	0.40	1	0.33	0.25	-0.12	0.31
				2	0.81	0.03	-0.34	0.03
	rs4579099	T	0.06	1	0.76	0.12	-0.43	0.05
				2	3.57	0.15	-1.36	0.10
	rs4835131	G	0.06	1	0.61	0.17	-0.34	0.07
				2	3.57	0.15	-1.37	0.10

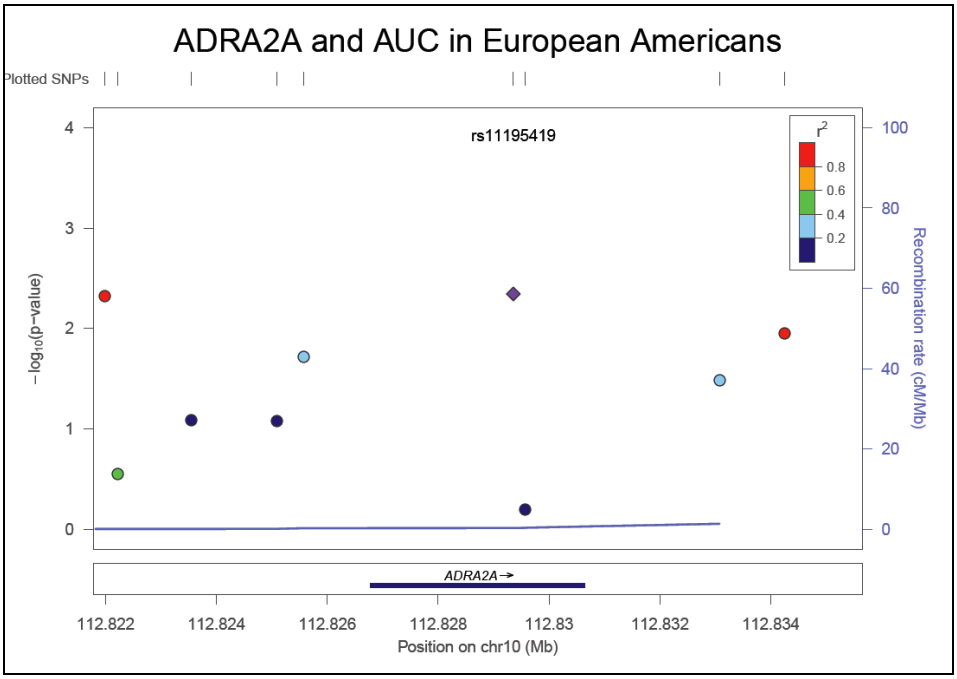


Figure A1: LocusZoom plot of the association between loci of the *ADRA2A* gene region among European Americans in predicting AUC.

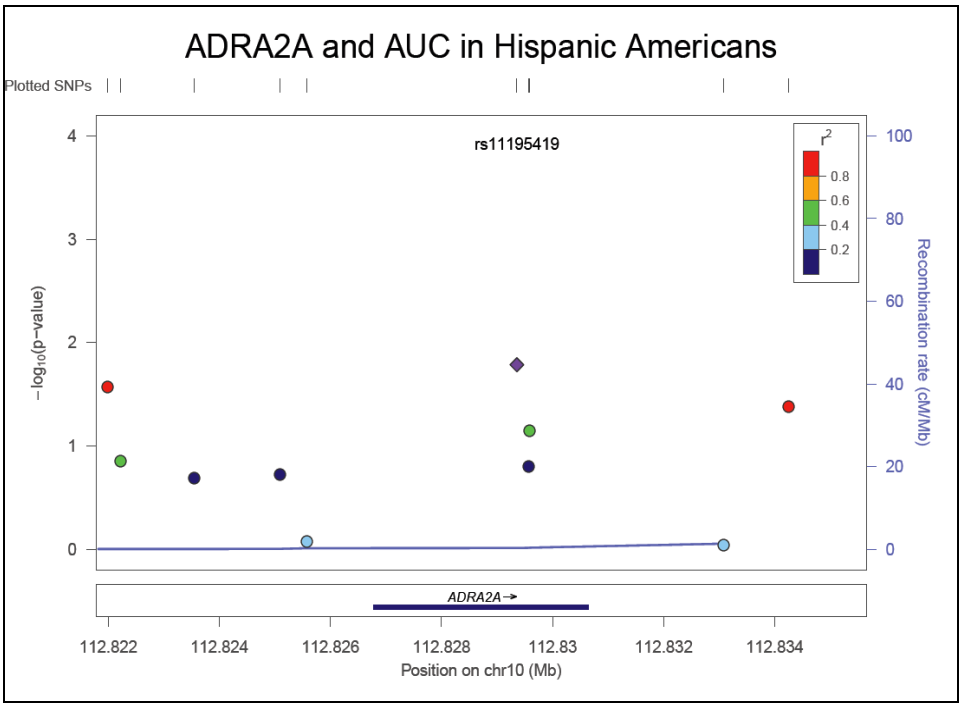


Figure A2: LocusZoom plot of the association between loci of the *ADRA2A* gene region among Hispanic Americans in predicting AUC.

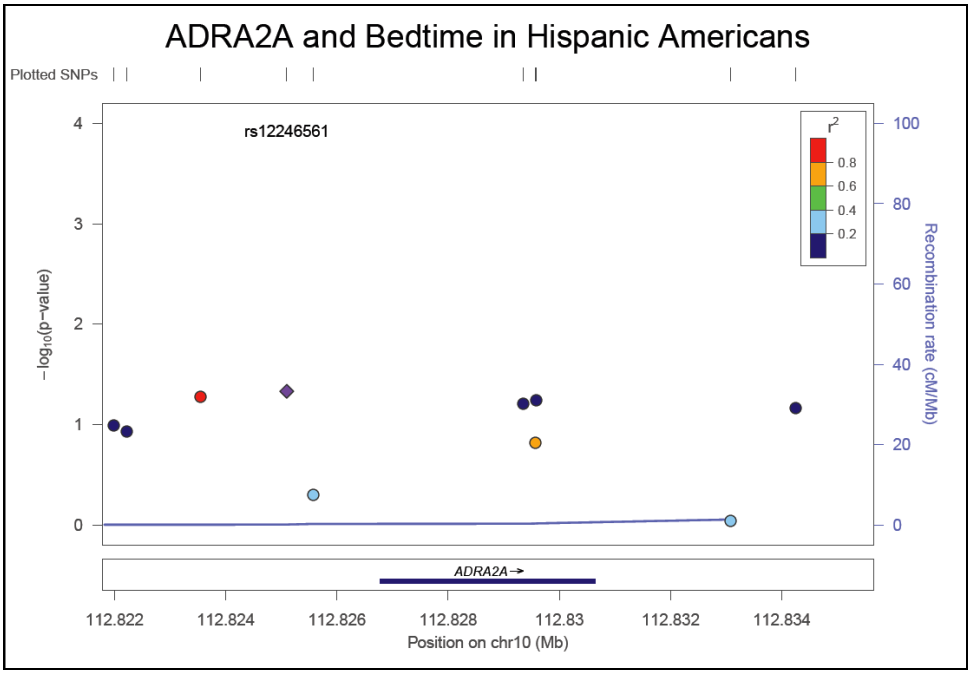


Figure A3: LocusZoom plot of the association between loci of the *ADRA2A* gene region among Hispanic Americans in predicting Bedtime.

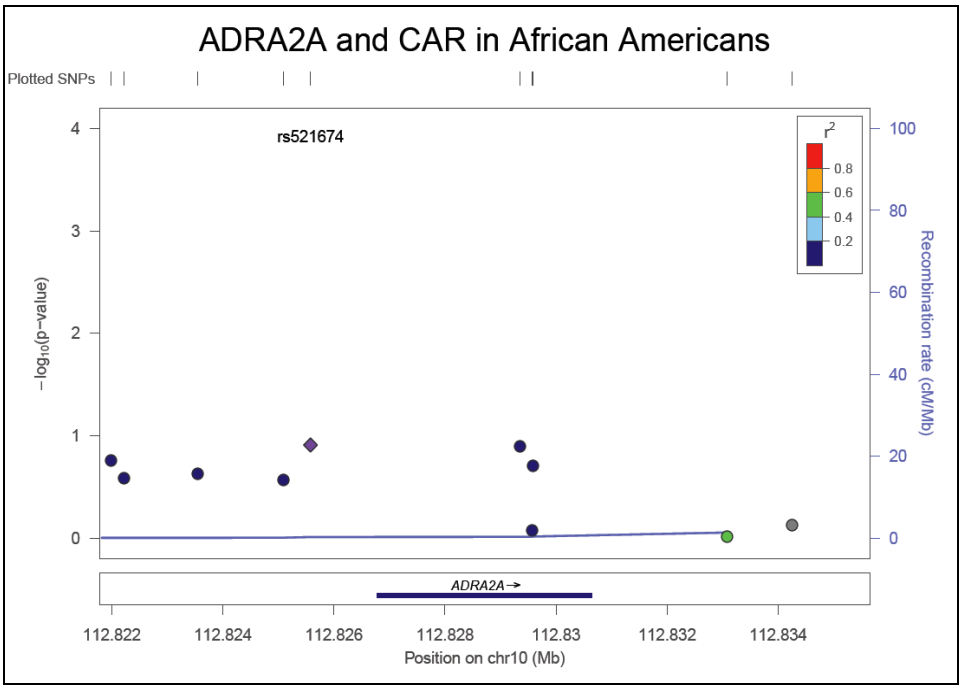


Figure A4: LocusZoom plot of the association between loci of the *ADRA2A* gene region among African Americans in predicting CAR.

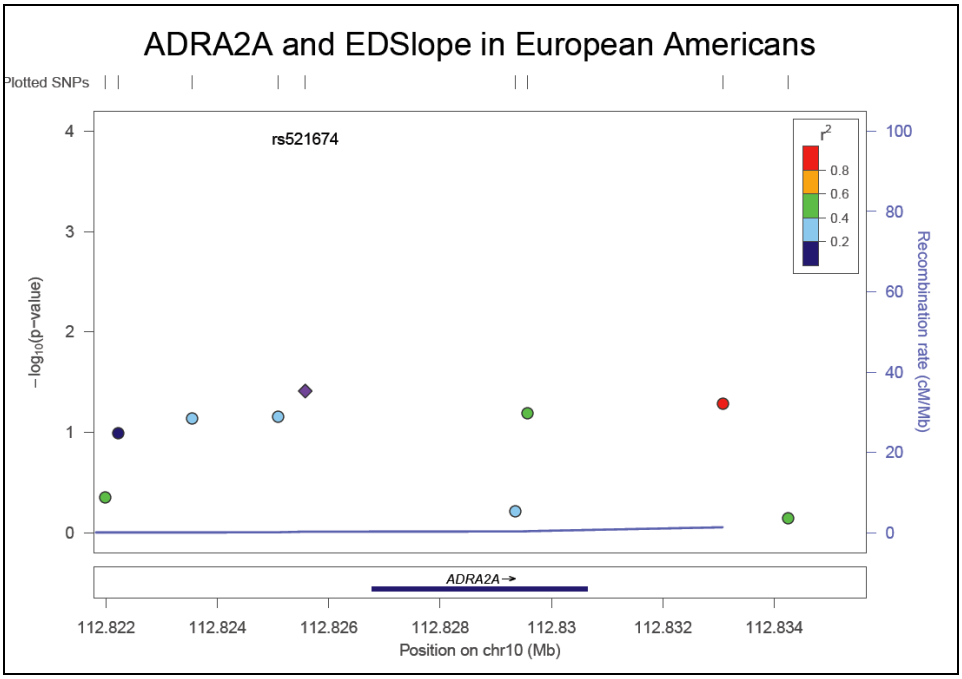


Figure A5: LocusZoom plot of the association between loci of the *ADRA2A* gene region among European Americans in predicting EDSlope.

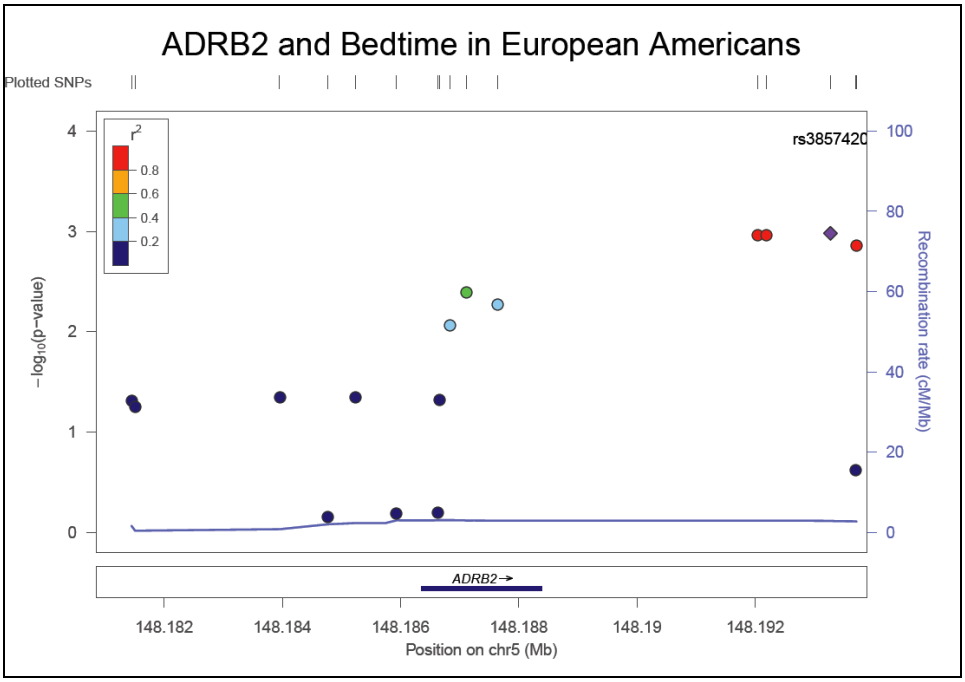


Figure A6: LocusZoom plot of the association between loci of the *ADRB2* gene region among European Americans in predicting Bedtime.

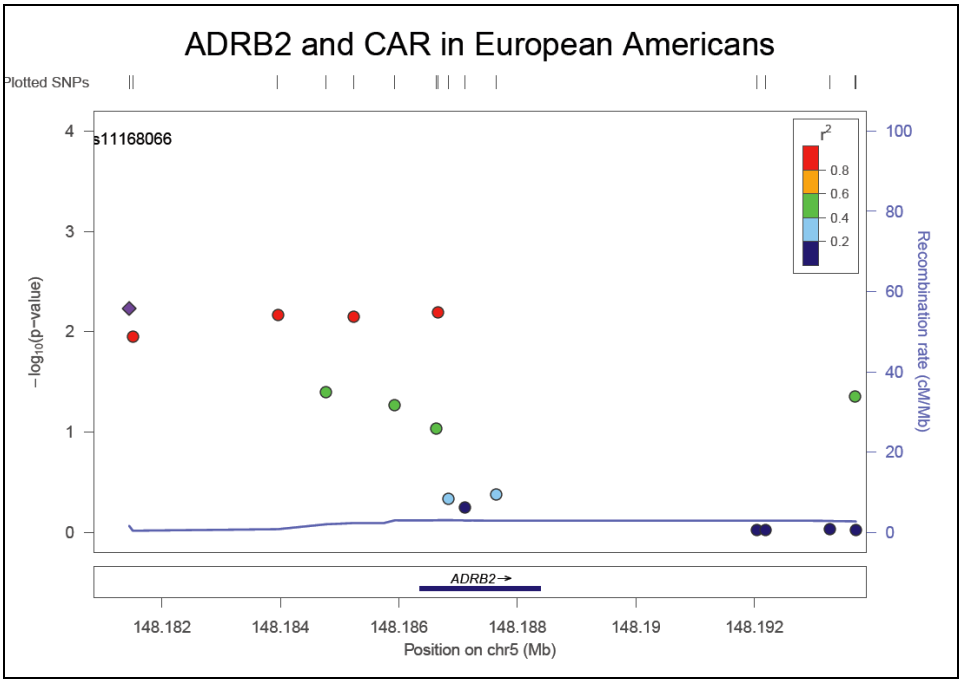


Figure A7: LocusZoom plot of the association between loci of the *ADRB2* gene region among European Americans in predicting CAR.

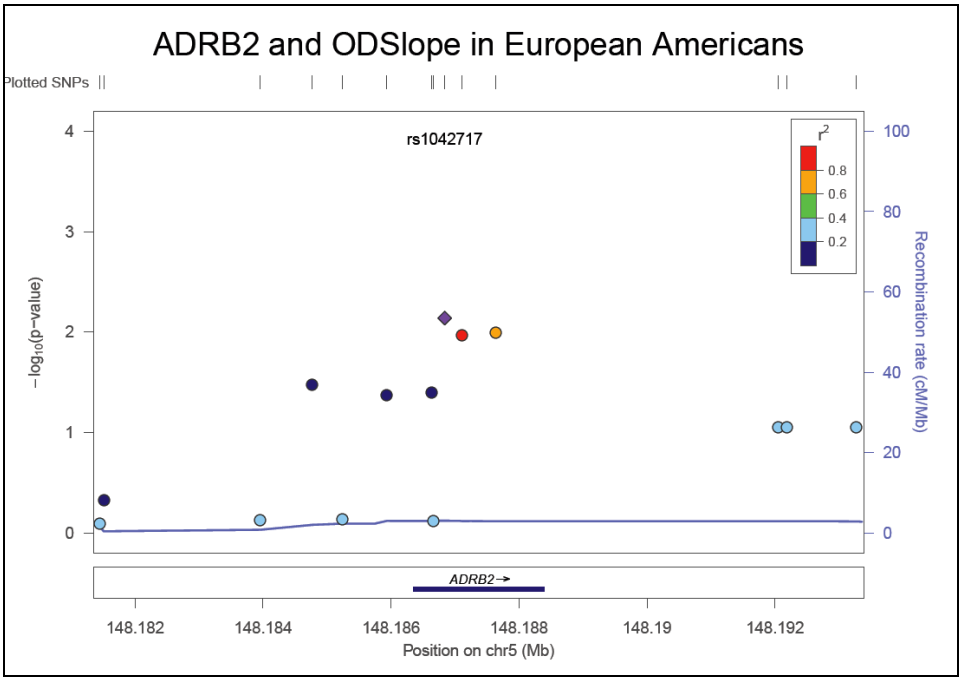


Figure A8: LocusZoom plot of the association between loci of the *ADRB2* gene region among European Americans in predicting ODSlope.

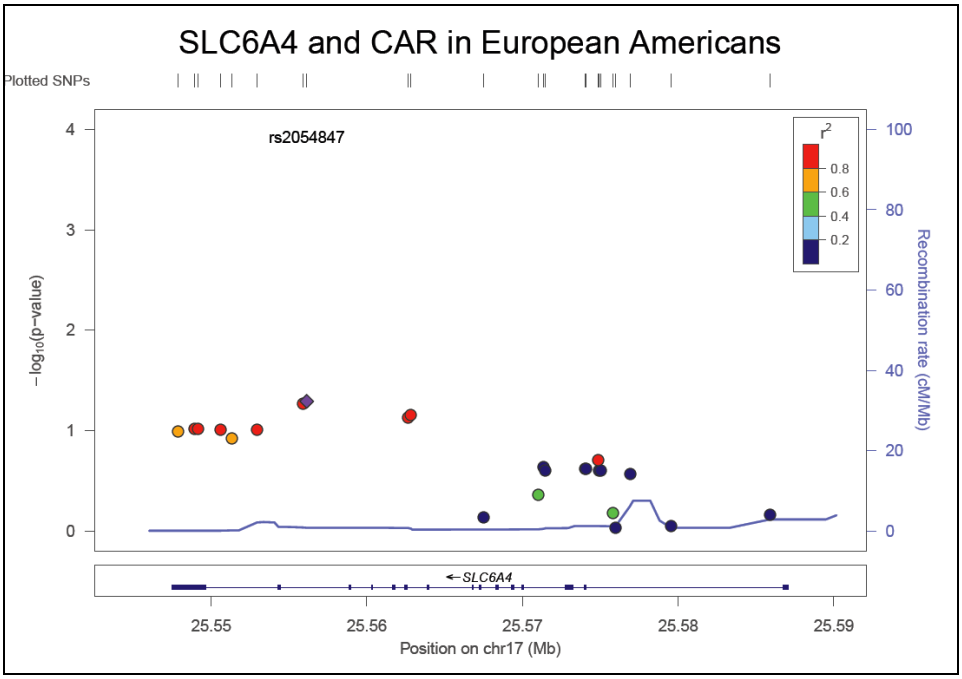


Figure A9: LocusZoom plot of the association between loci of the *SLC6A4* gene region among European Americans in predicting CAR.

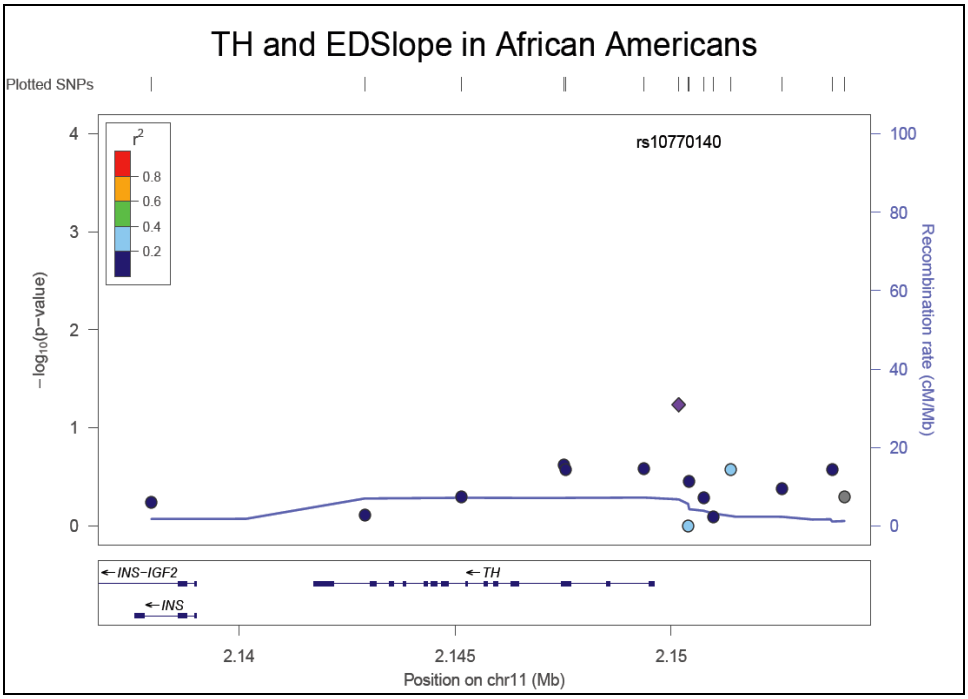


Figure A10: LocusZoom plot of the association between loci of the *TH* gene region among African Americans in predicting EDSlope.

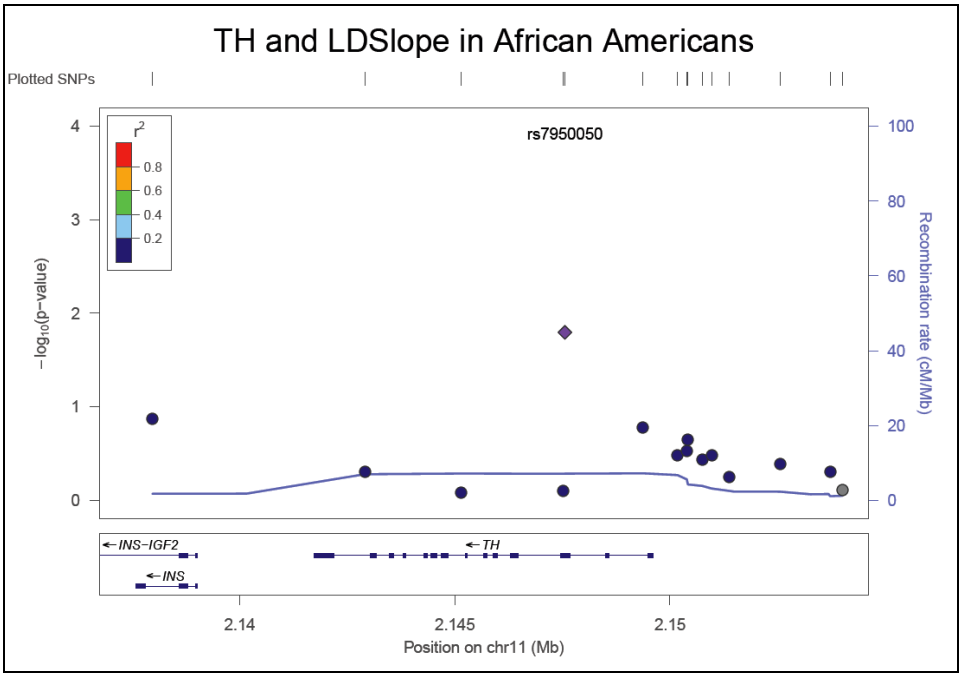


Figure A11: LocusZoom plot of the association between loci of the *TH* gene region among African Americans in predicting LDSlope.

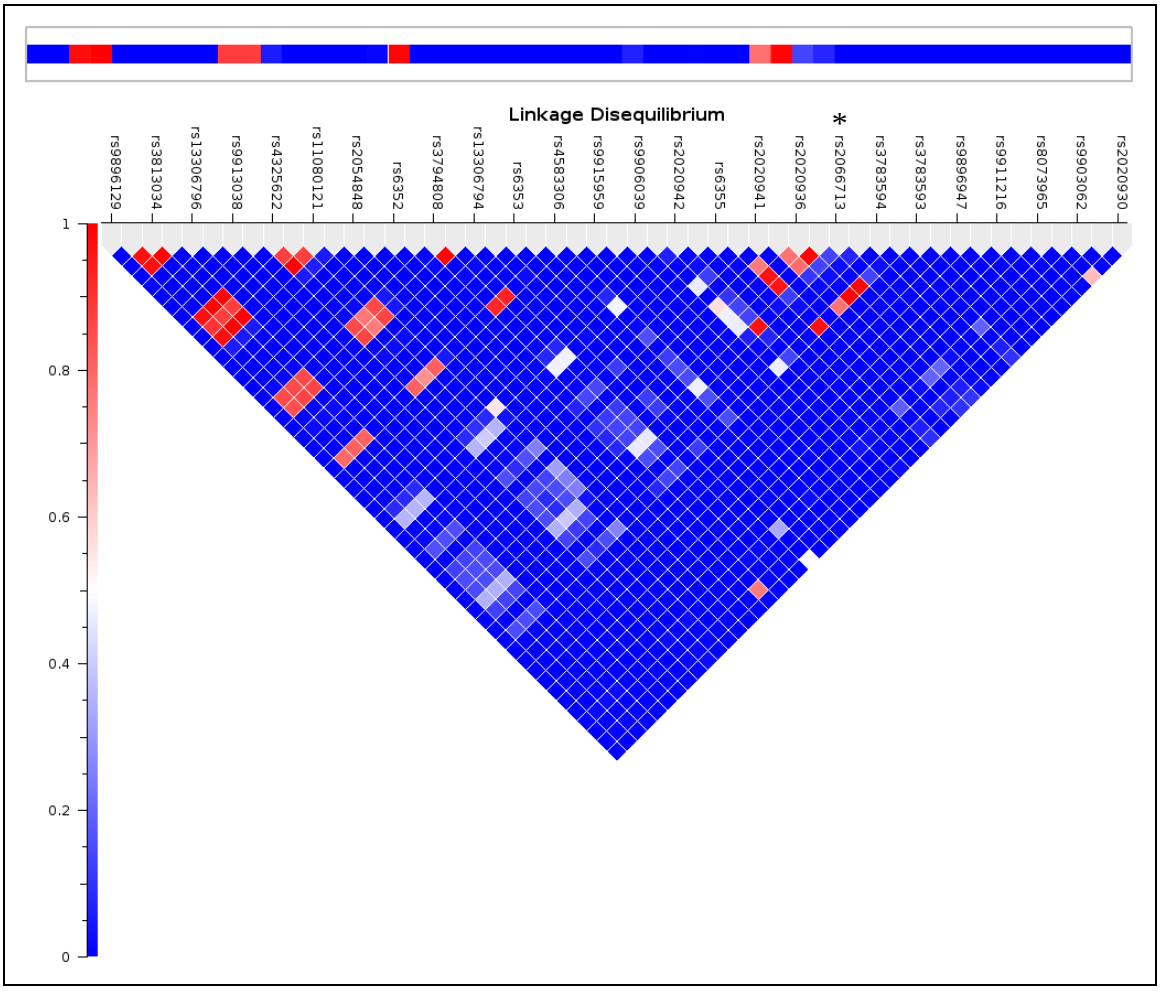


Figure A12: Linkage disequilibrium plot for *SLC6A4* in European Americans. * indicates the index SNP.

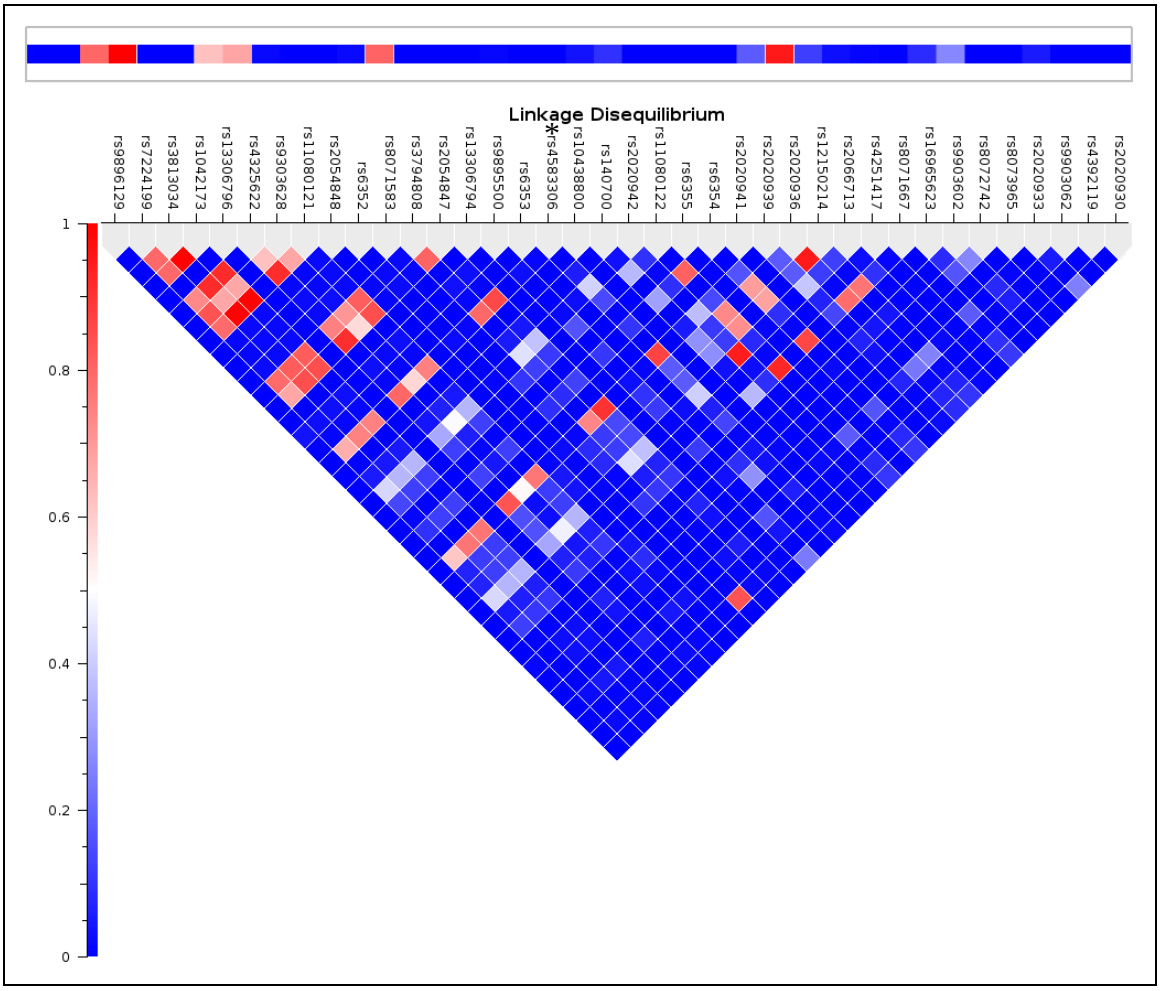


Figure A13: Linkage disequilibrium plot for *SLC6A4* in Hispanic Americans. * indicates the index SNP.

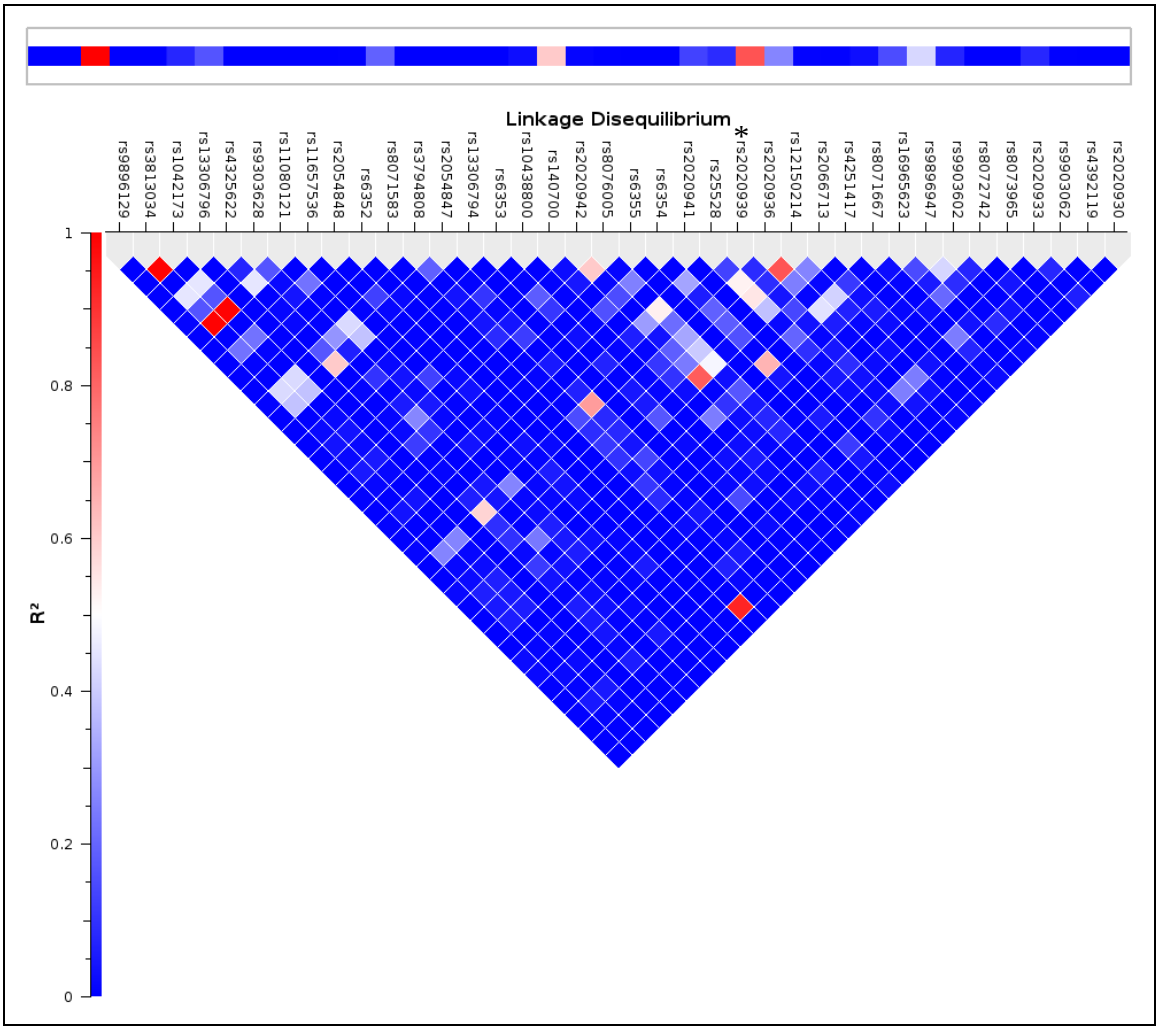


Figure A14: Linkage disequilibrium plot for *SLC6A4* in African Americans. * indicates the index SNP.

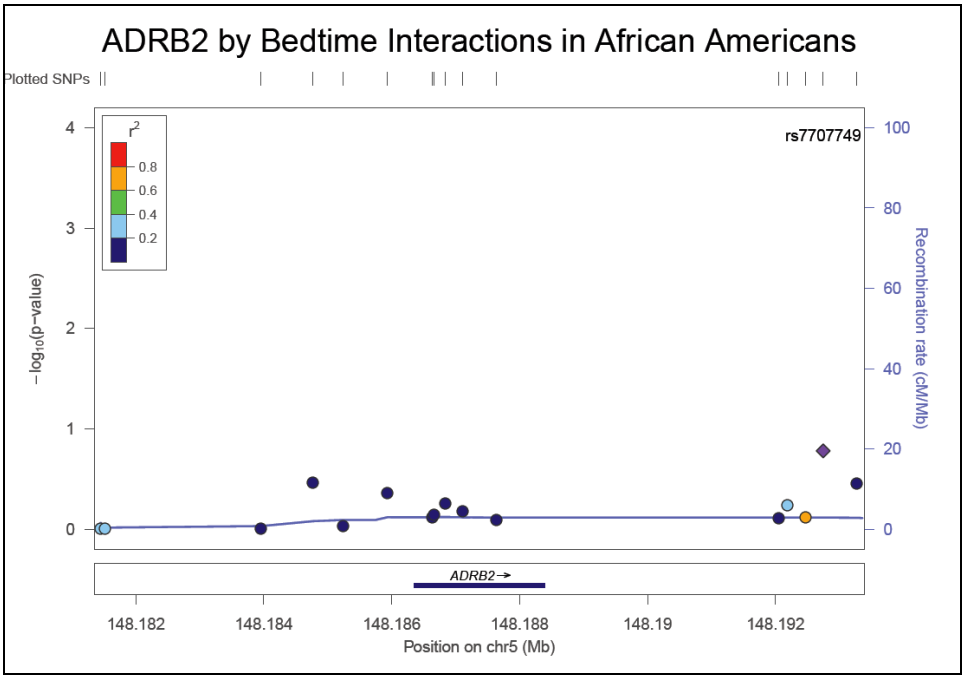


Figure A15: LocusZoom plot of the interaction between loci of the *ADRB2* gene region and Bedtime among African Americans in predicting ln(IL-6 + 1).

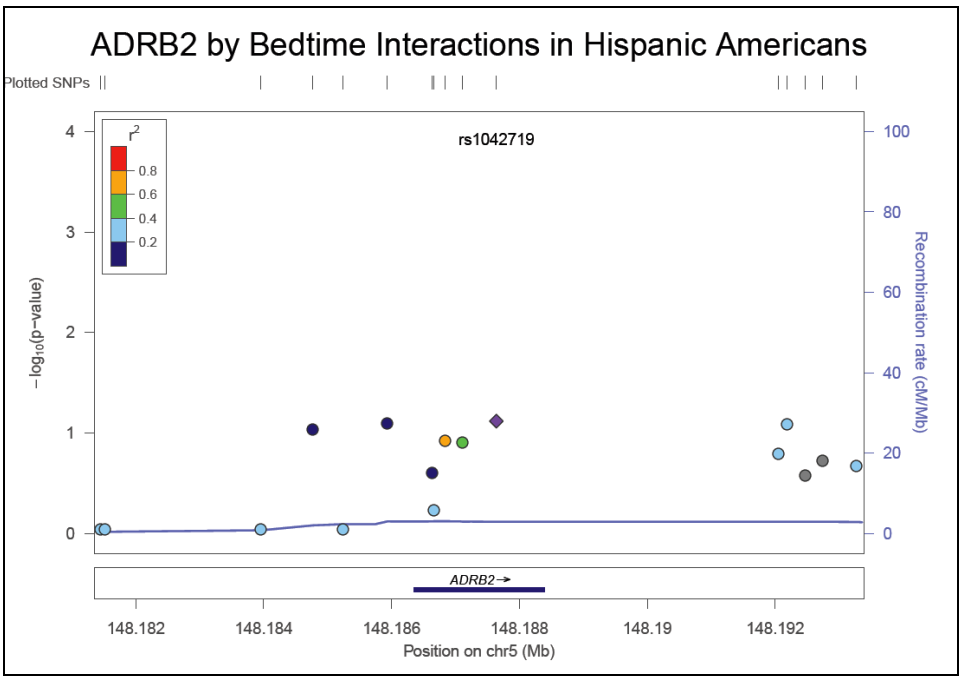


Figure A16: LocusZoom plot of the interaction between loci of the *ADRB2* gene region and Bedtime among Hispanic Americans in predicting ln(IL-6 + 1).

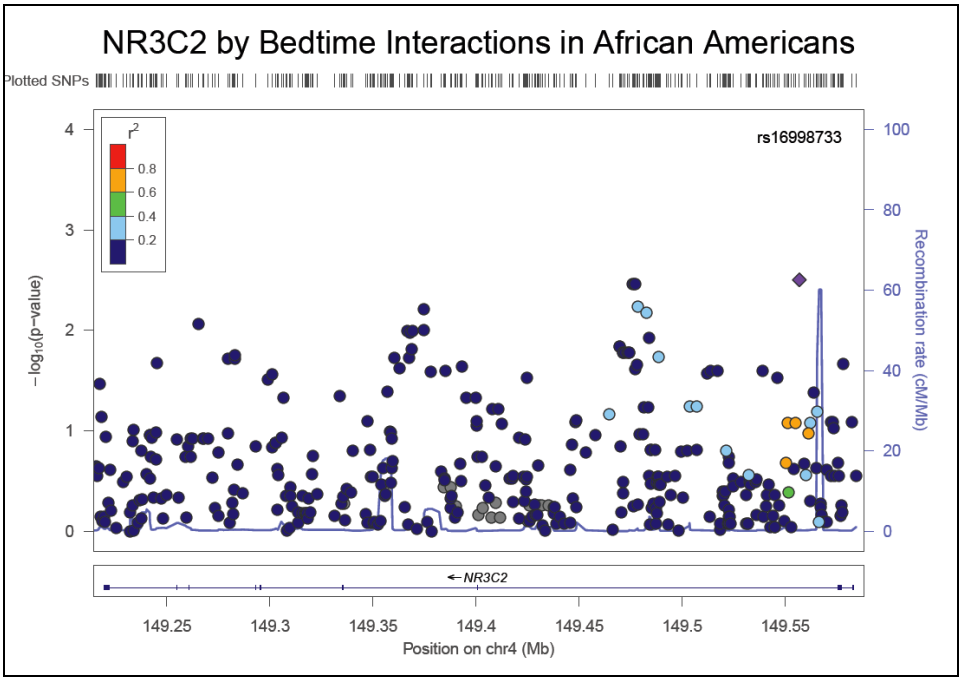


Figure A17: LocusZoom plot of the interaction between loci of the *NR3C2* gene region and Bedtime among African Americans in predicting ln(IL-6 + 1).

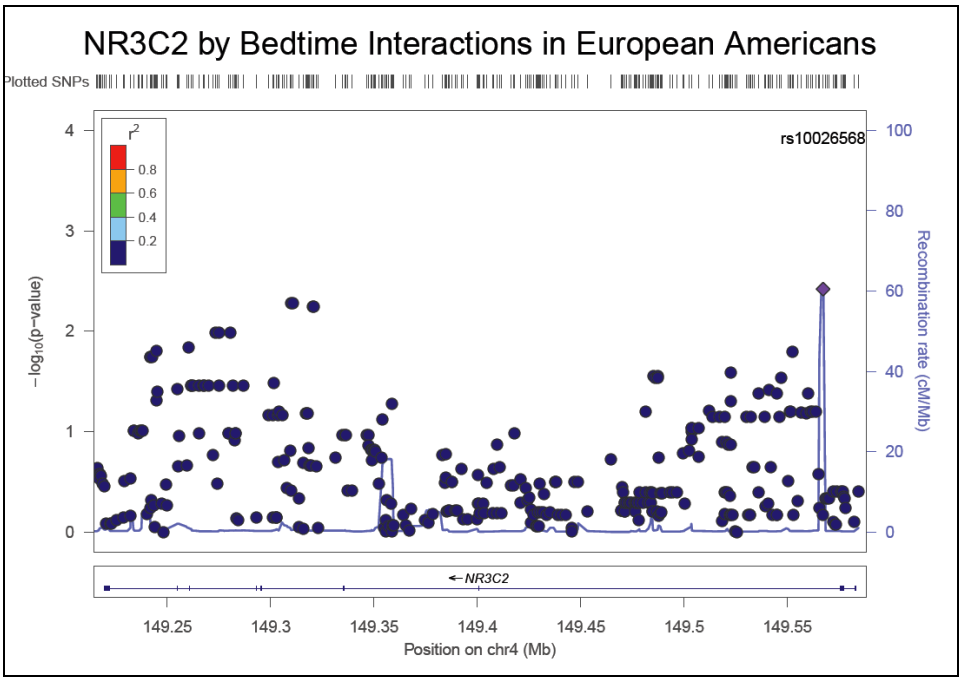


Figure A18: LocusZoom plot of the interaction between loci of the *NR3C2* gene region and Bedtime among European Americans in predicting ln(IL-6 + 1).

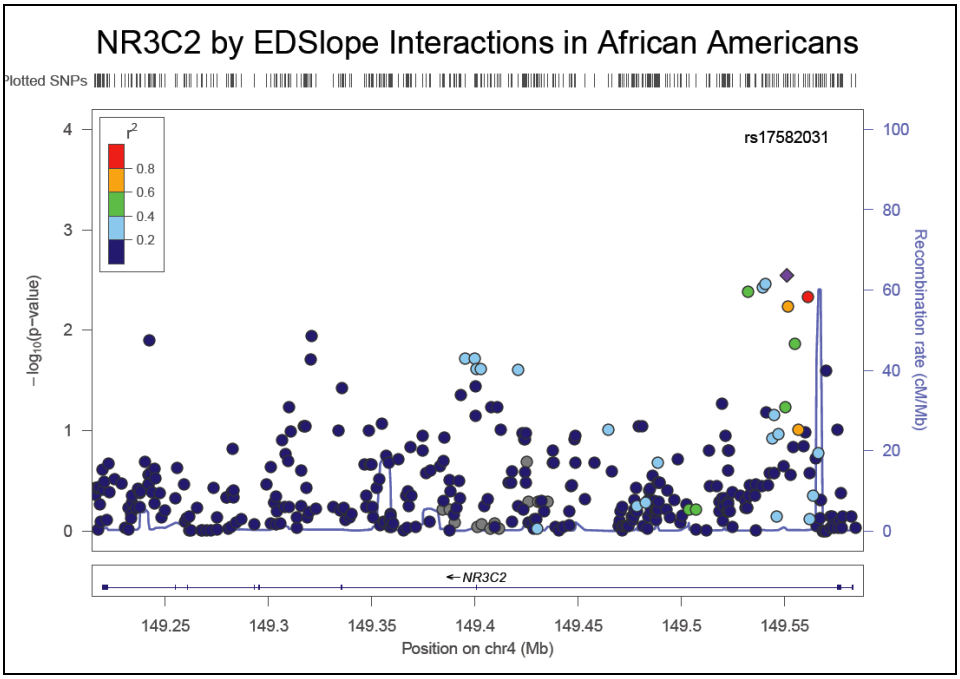


Figure A19: LocusZoom plot of the interaction between loci of the *NR3C2* gene region and EDSlope among African Americans in predicting ln(IL-6 + 1).

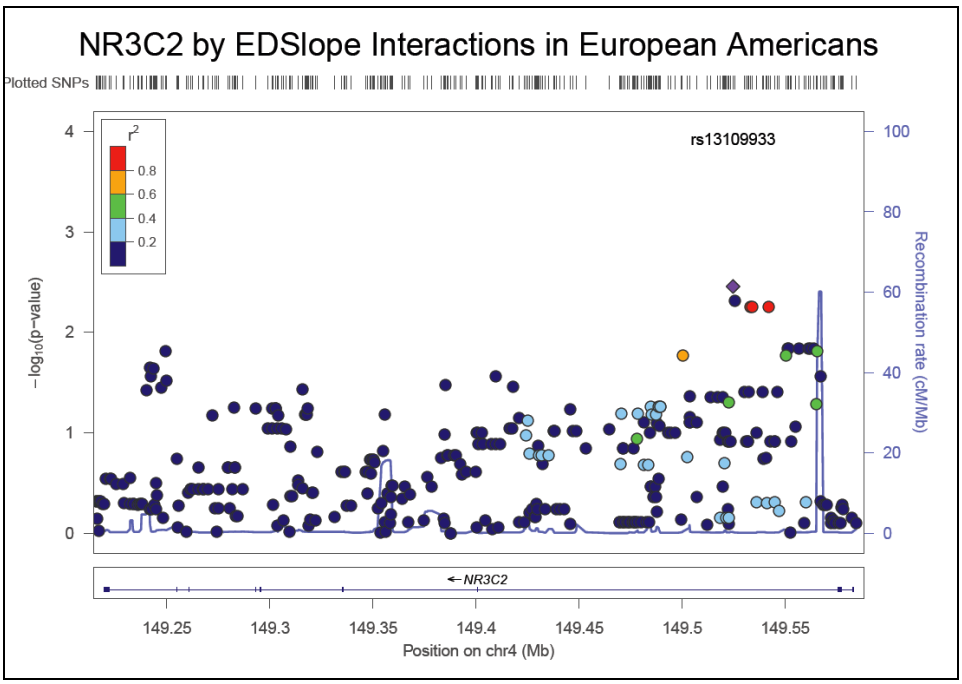


Figure A20: LocusZoom plot of the interaction between loci of the *NR3C2* gene region and EDSlope among European Americans in predicting ln(IL-6 + 1).

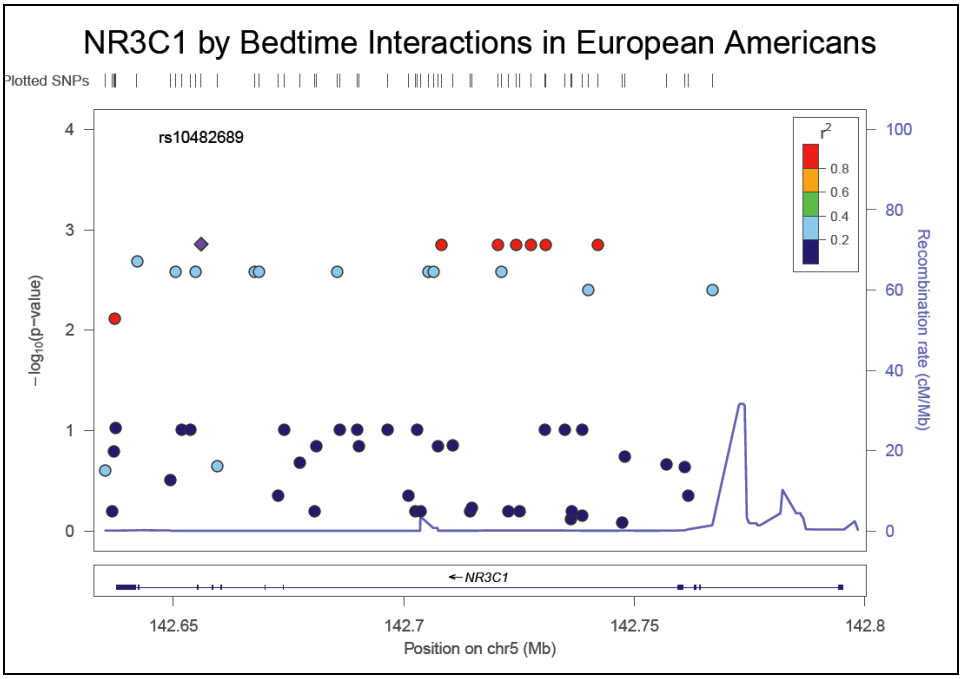


Figure A21: LocusZoom plot of the interaction between loci of the *NR3C1* gene region and Bedtime among European Americans in predicting ln(TNF-α + 1).

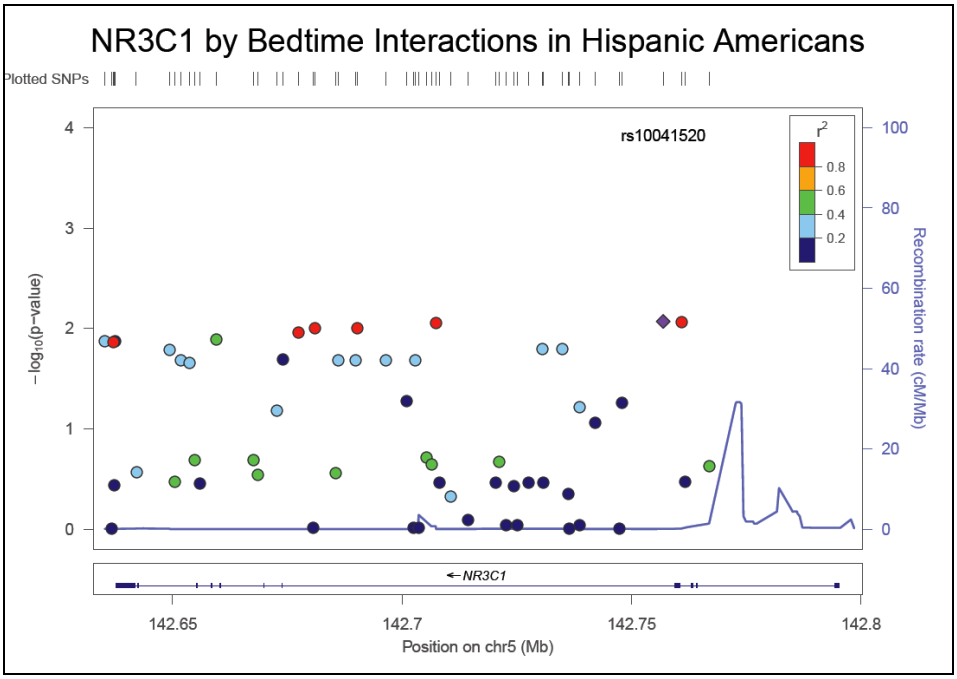


Figure A22: LocusZoom plot of the interaction between loci of the *NR3C1* gene region and Bedtime among Hispanic Americans in predicting ln(TNF-α + 1).

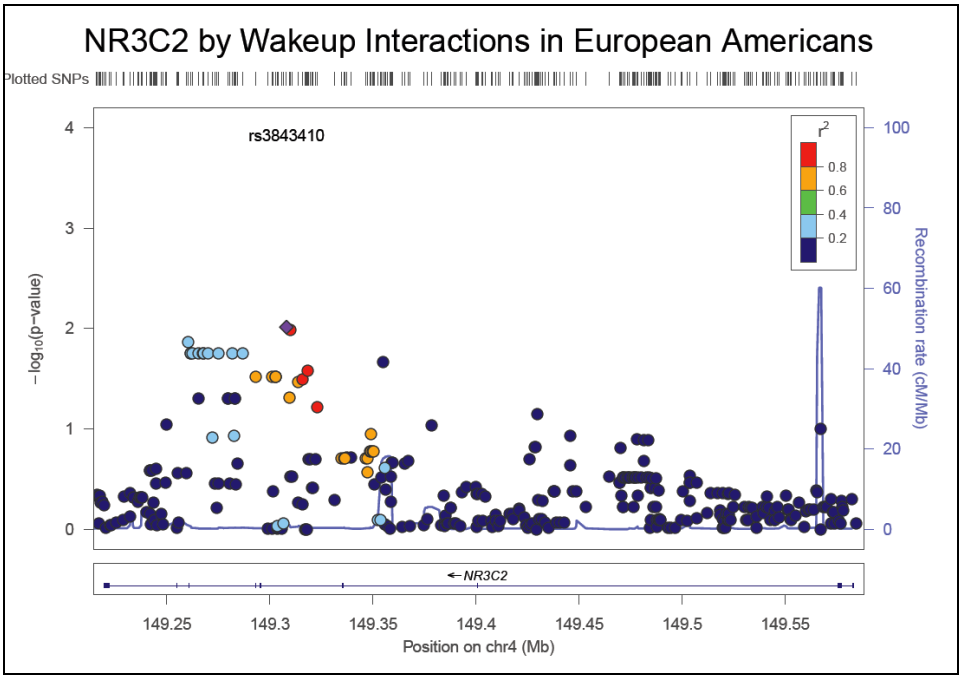


Figure A23: LocusZoom plot of the interaction between loci of the *NR3C2* gene region and Wakeup among European Americans in predicting ln(TNF-α + 1).

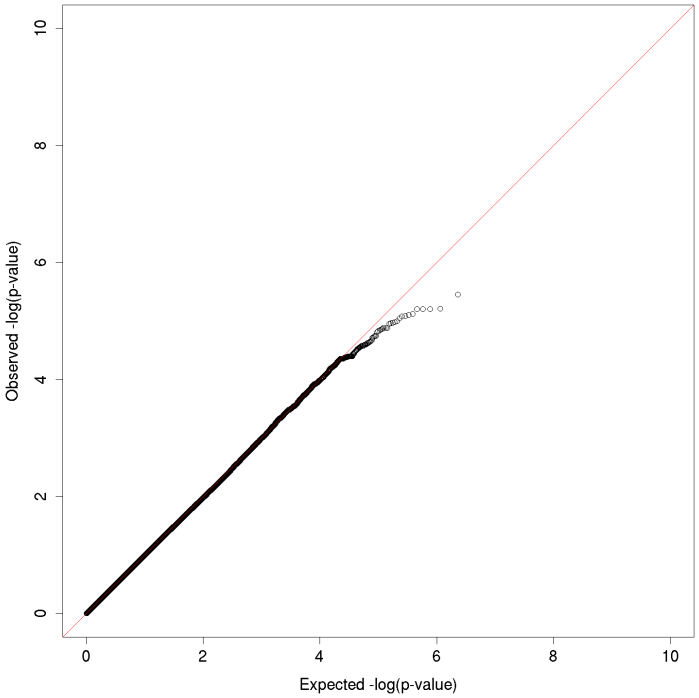


Figure A24: Quantile-Quantile plot of the observed and expected p-values for associations with AUC among European Americans.

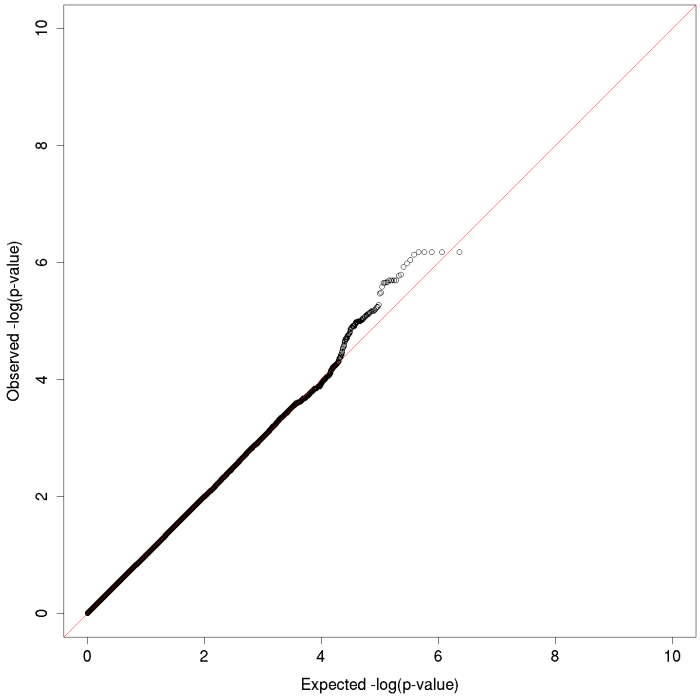


Figure A25: Quantile-Quantile plot of the observed and expected p-values for associations with Bedtime among European Americans.

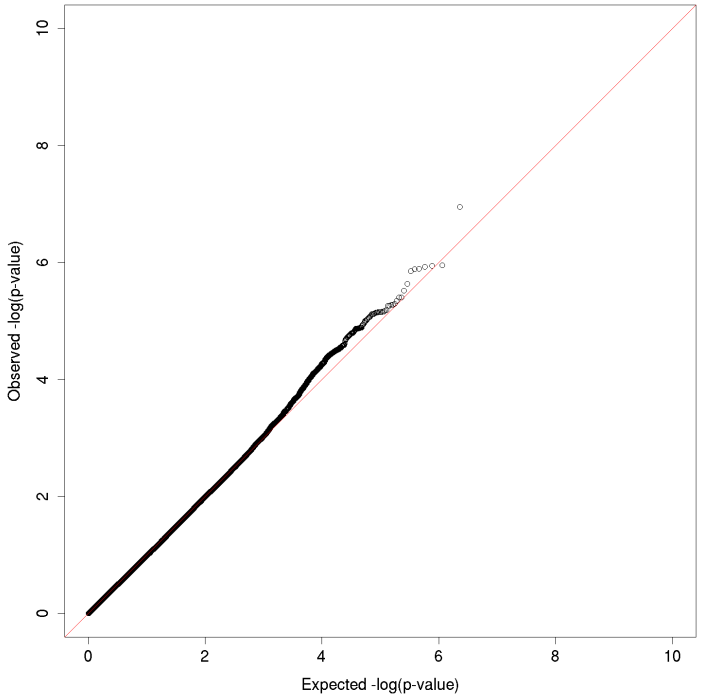


Figure A26: Quantile-Quantile plot of the observed and expected p-values for associations with CAR among European Americans.

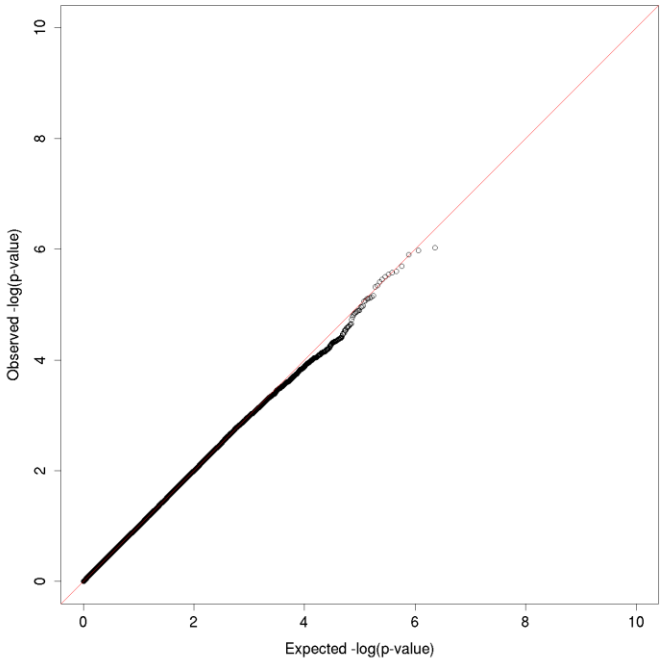


Figure A27: Quantile-Quantile plot of the observed and expected p-values for associations with EDSlope among European Americans.

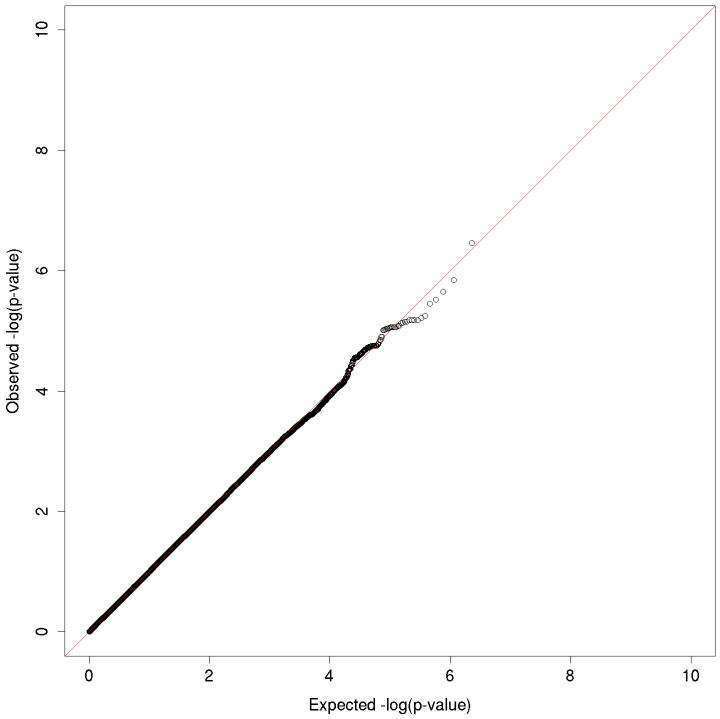


Figure A28: Quantile-Quantile plot of the observed and expected p-values for associations with LDSlope among European Americans.

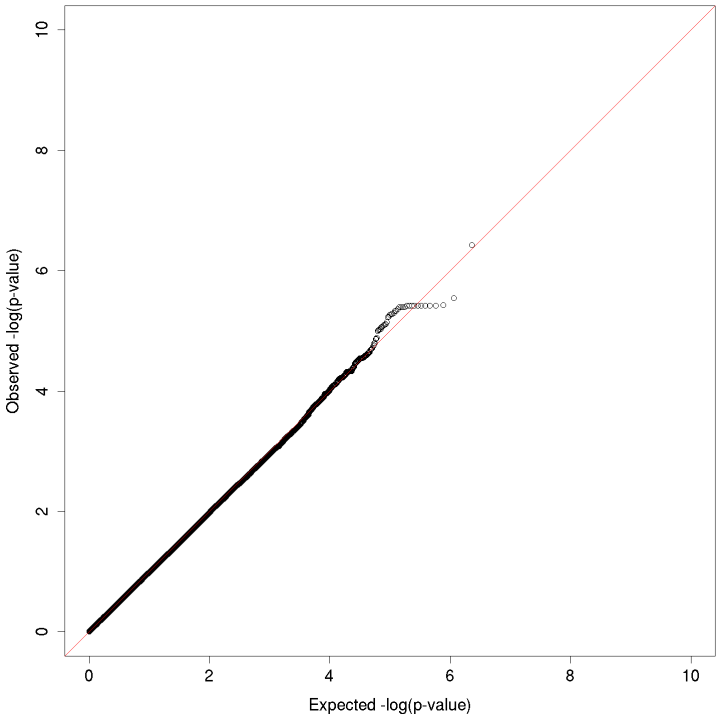


Figure A29: Quantile-Quantile plot of the observed and expected p-values for associations with ODSlope among European Americans.

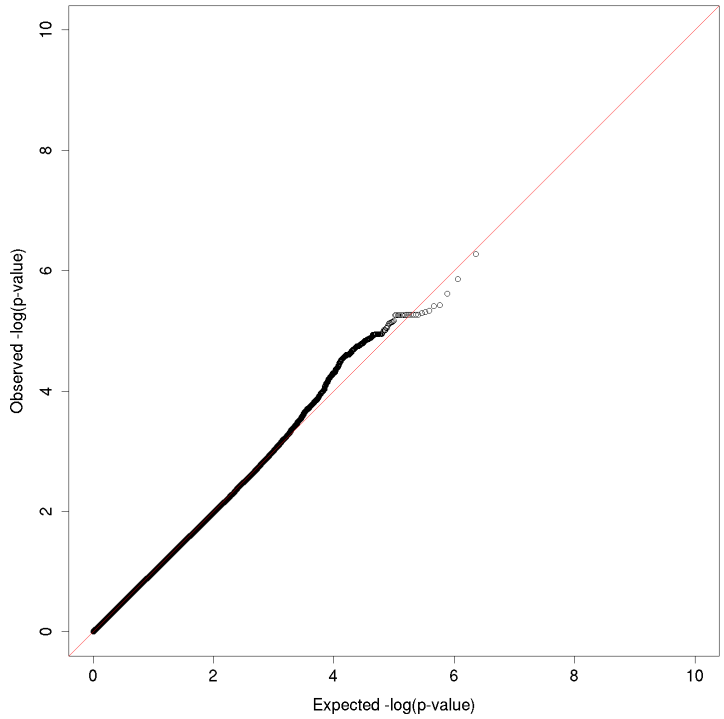


Figure A30: Quantile-Quantile plot of the observed and expected p-values for associations with Wakeup among European Americans.

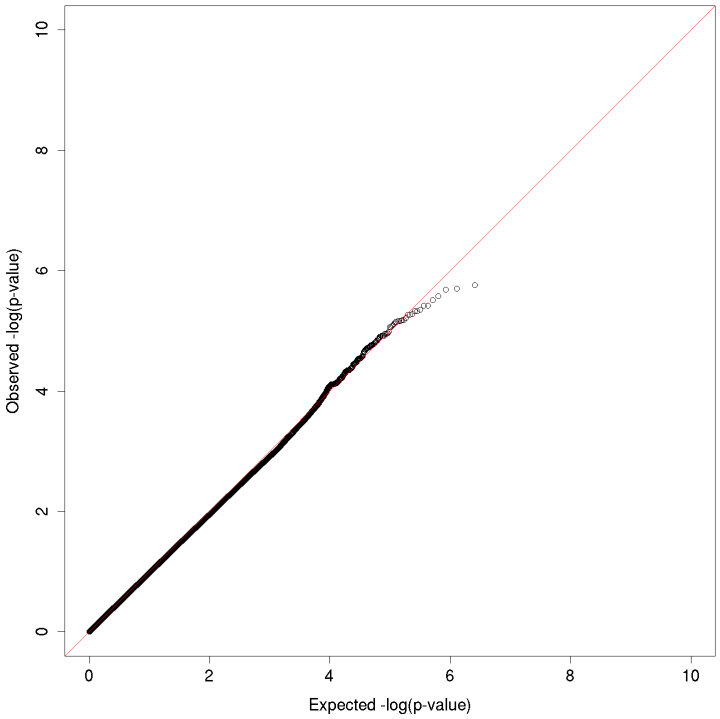


Figure A31: Quantile-Quantile plot of the observed and expected p-values for associations with AUC among African Americans.

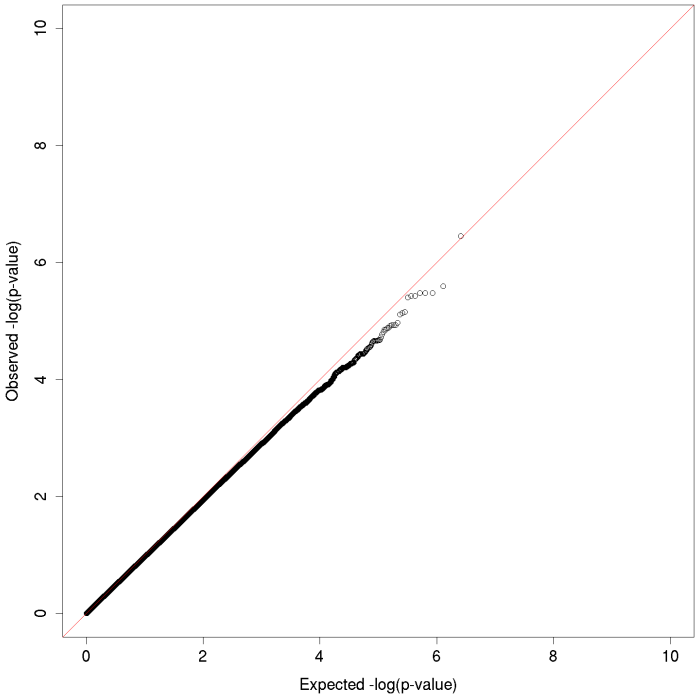


Figure A32: Quantile-Quantile plot of the observed and expected p-values for associations with Bedtime among African Americans.

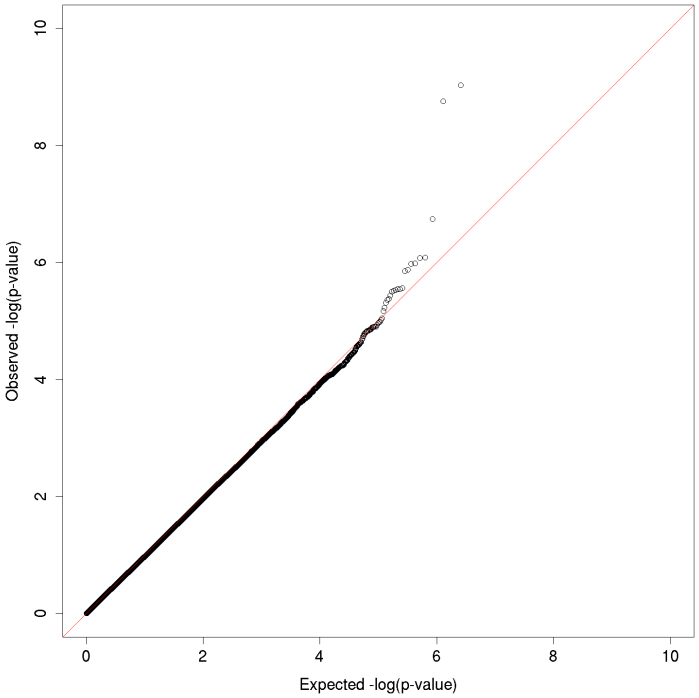


Figure A33: Quantile-Quantile plot of the observed and expected p-values for associations with CAR among African Americans. Note the two loci that exceeded the genome-wide significance threshold of $5x10^{-8}$.

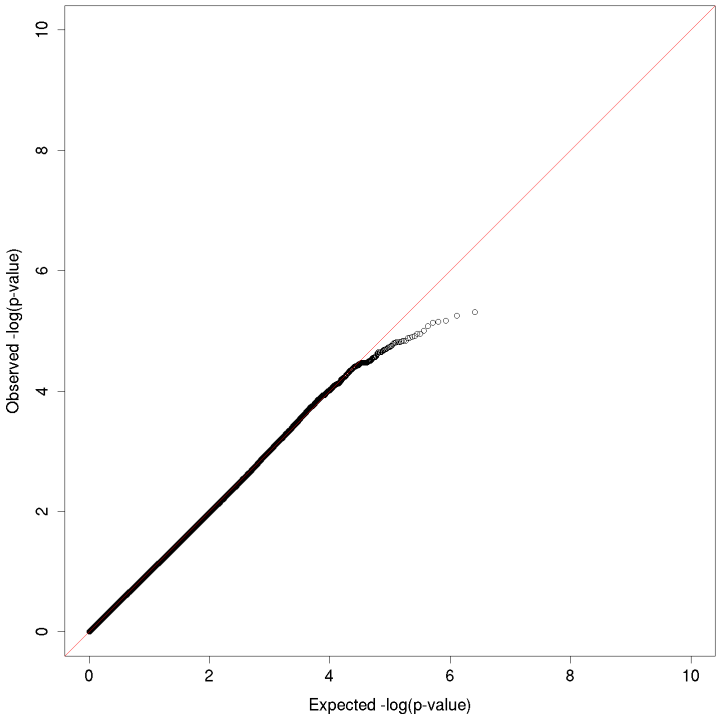


Figure A34: Quantile-Quantile plot of the observed and expected p-values for associations with EDSlope among African Americans.

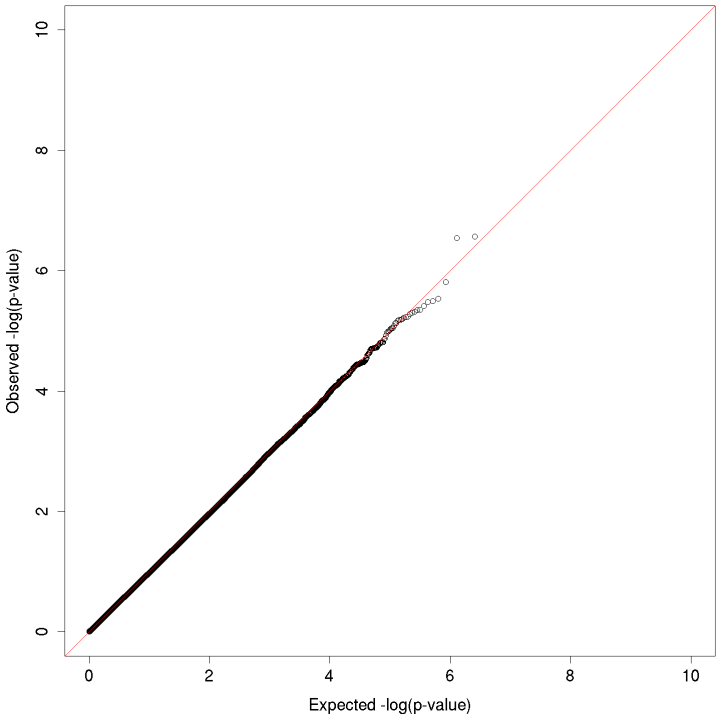


Figure A35: Quantile-Quantile plot of the observed and expected p-values for associations with LDSlope among African Americans.

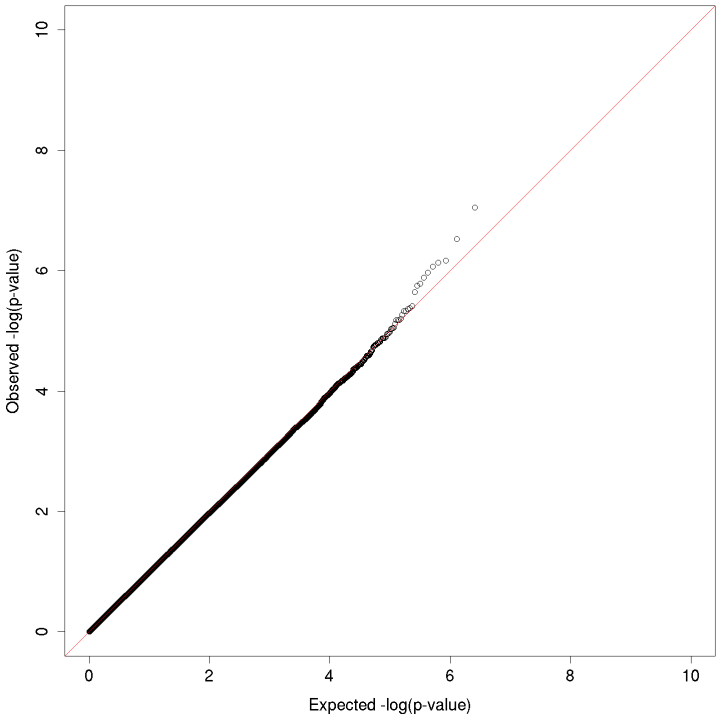


Figure A36: Quantile-Quantile plot of the observed and expected p-values for associations with ODSlope among African Americans.

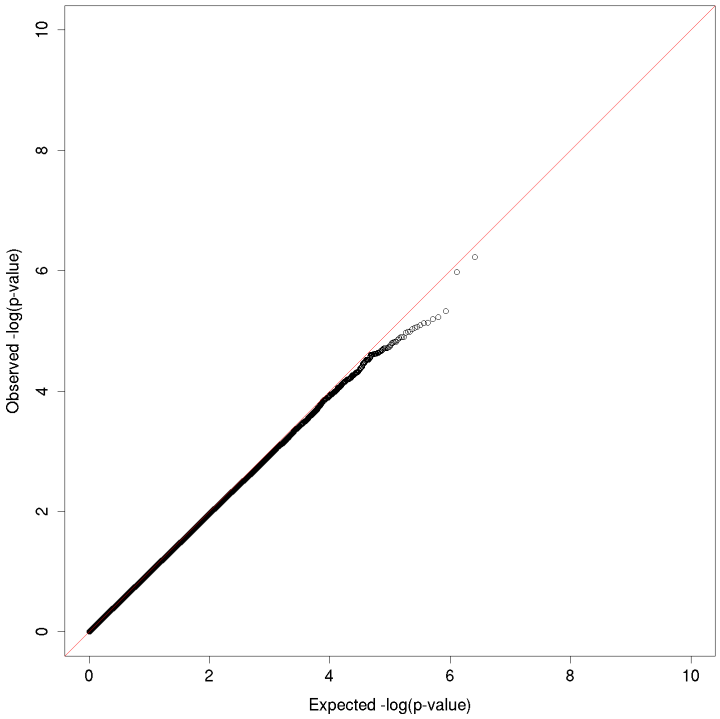


Figure A37: Quantile-Quantile plot of the observed and expected p-values for associations with Wakeup among African Americans.

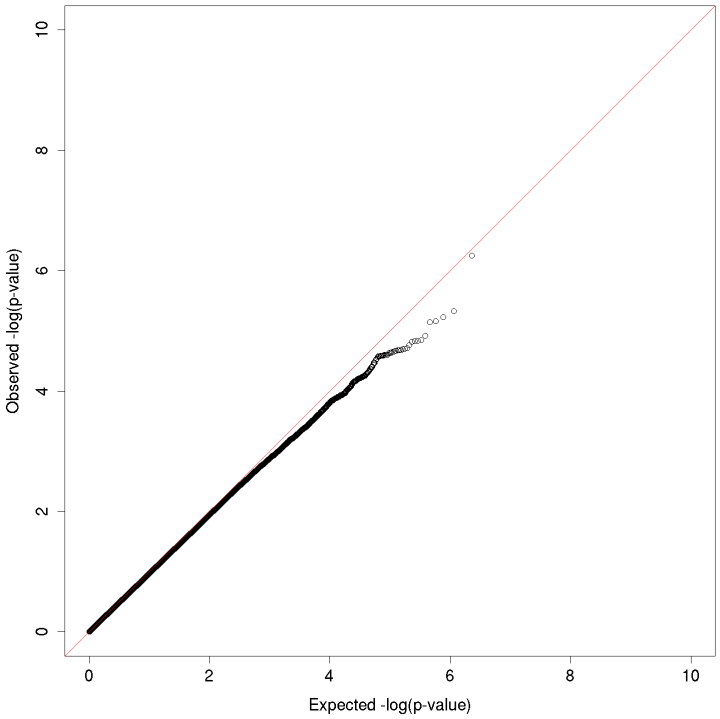


Figure A38: Quantile-Quantile plot of the observed and expected p-values for associations with AUC among Hispanic Americans.

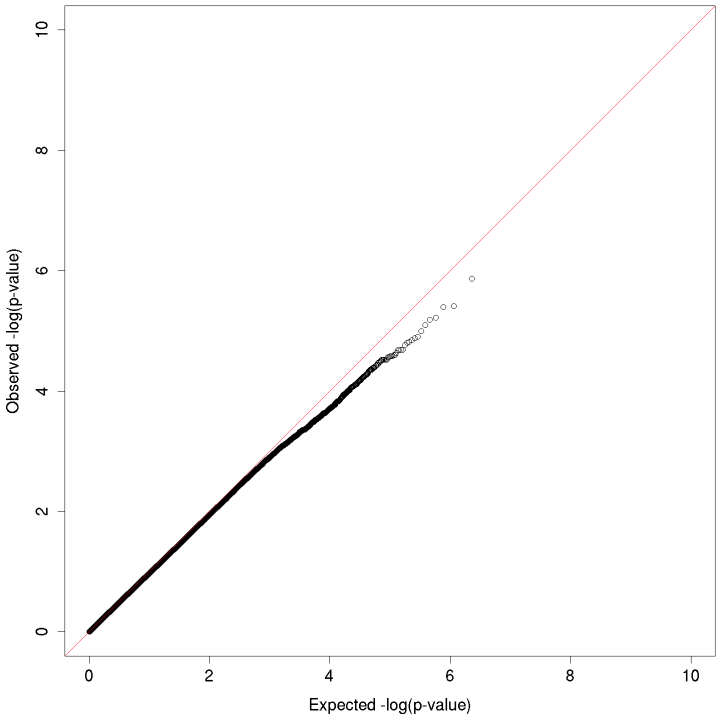


Figure A39: Quantile-Quantile plot of the observed and expected p-values for associations with Bedtime among Hispanic Americans.

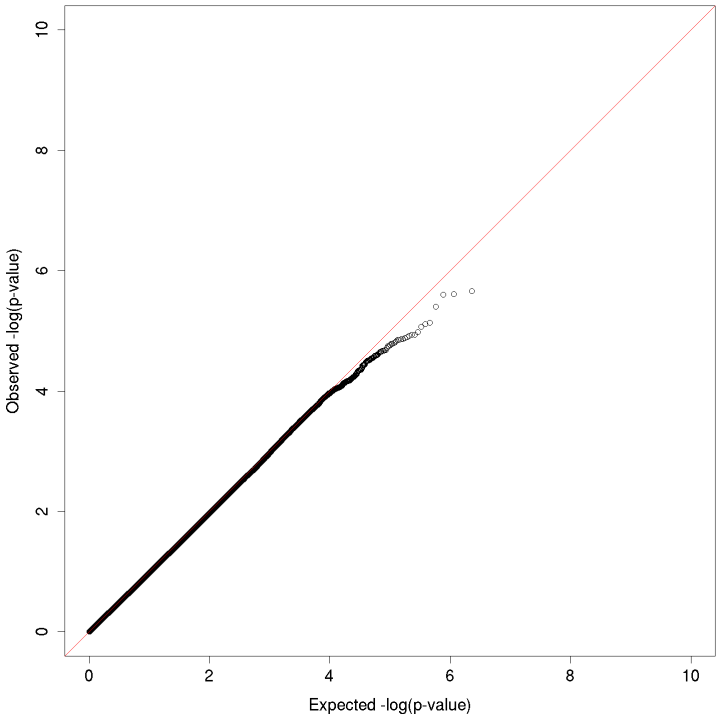


Figure A40: Quantile-Quantile plot of the observed and expected p-values for associations with CAR among Hispanic Americans.

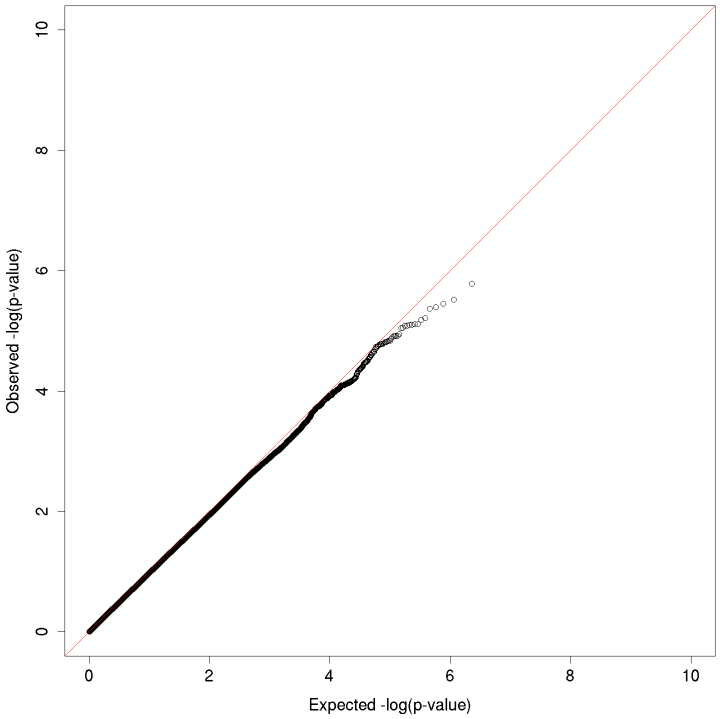


Figure A41: Quantile-Quantile plot of the observed and expected p-values for associations with EDSlope among Hispanic Americans.

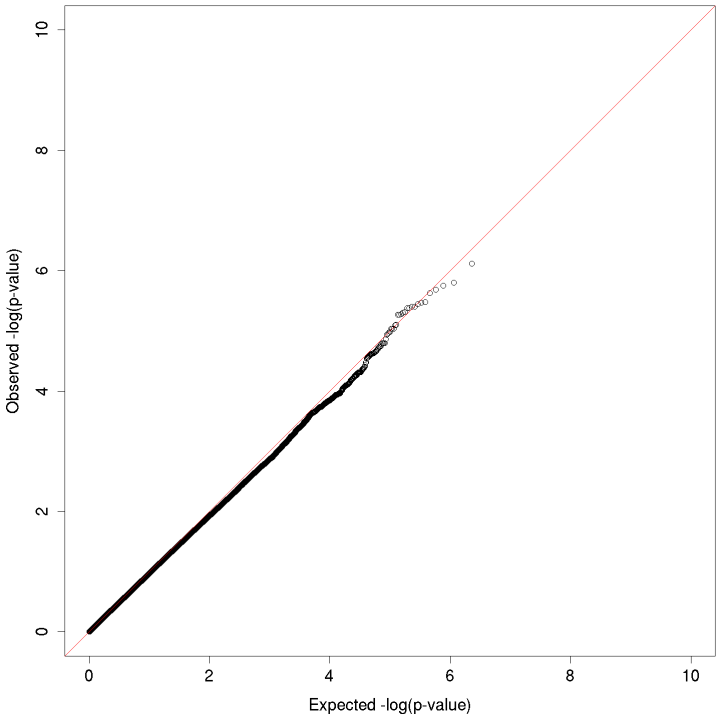


Figure A42: Quantile-Quantile plot of the observed and expected p-values for associations with LDSlope among Hispanic Americans.

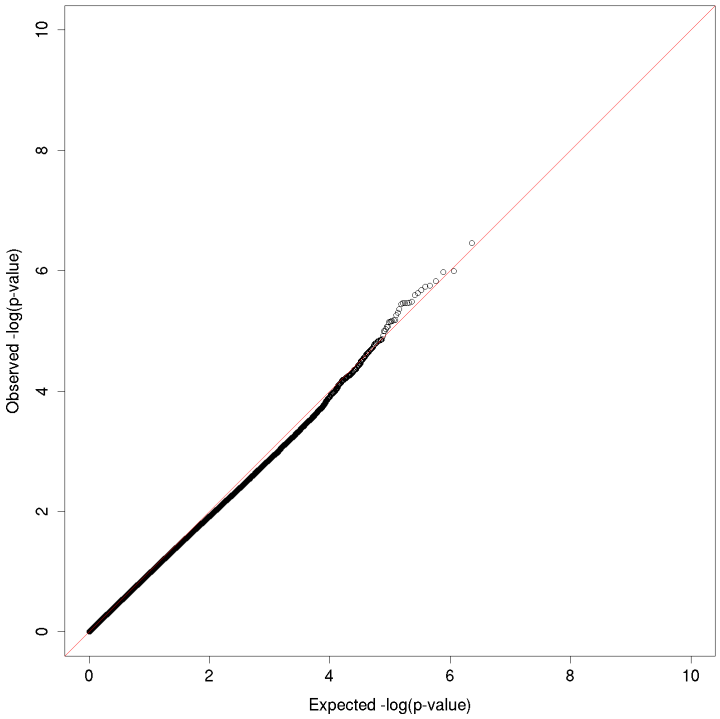


Figure A43: Quantile-Quantile plot of the observed and expected p-values for associations with ODSlope among Hispanic Americans.

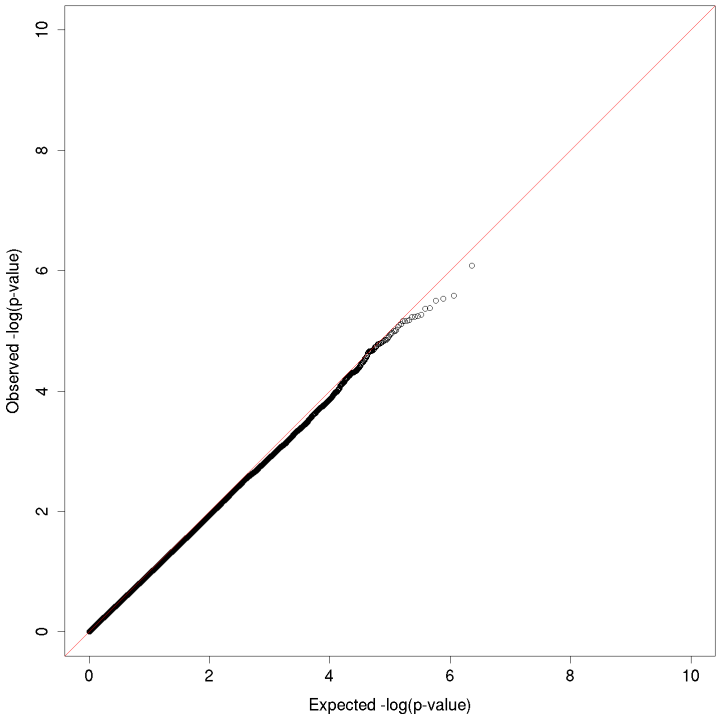
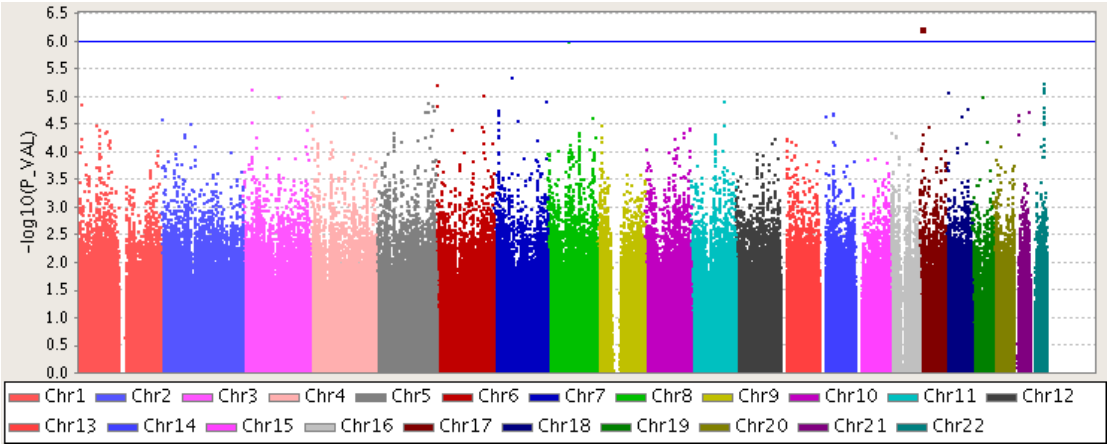
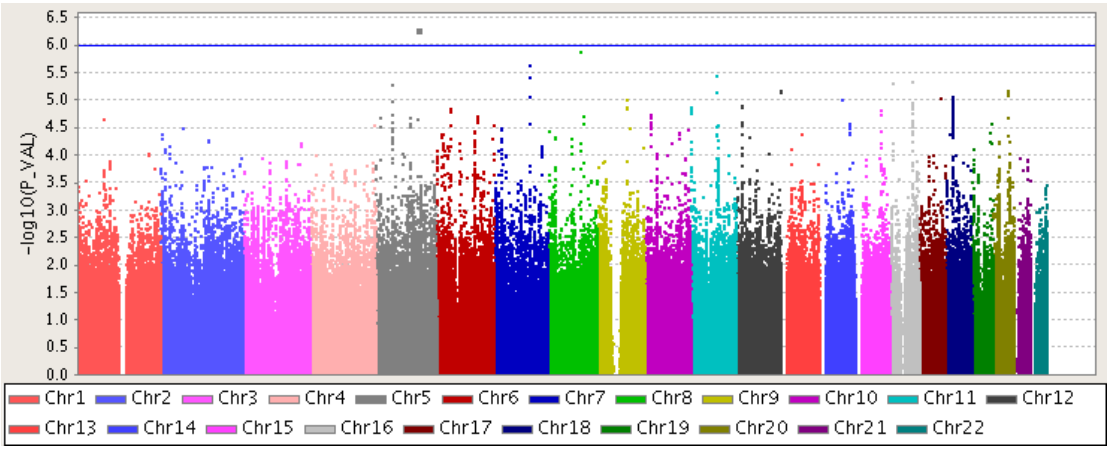


Figure A44: Quantile-Quantile plot of the observed and expected p-values for associations with Wakeup among Hispanic Americans.

AFA



EA



HIS

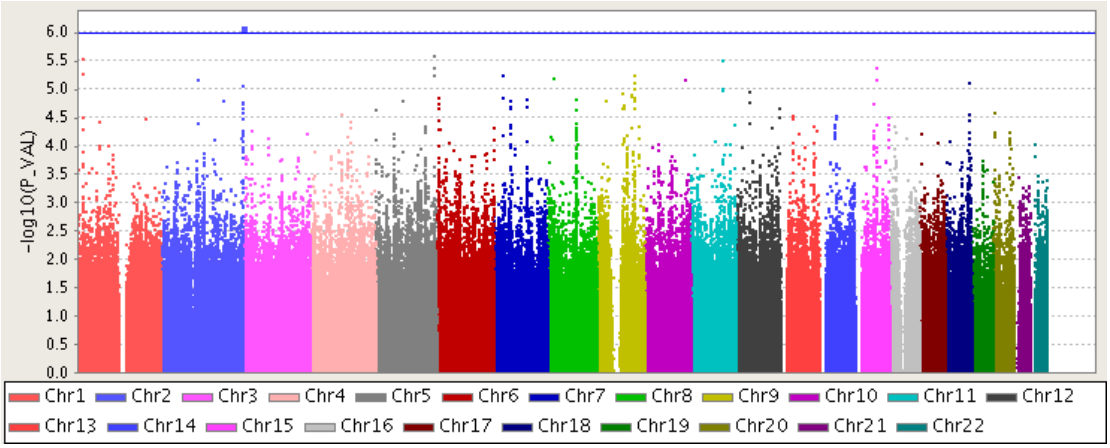
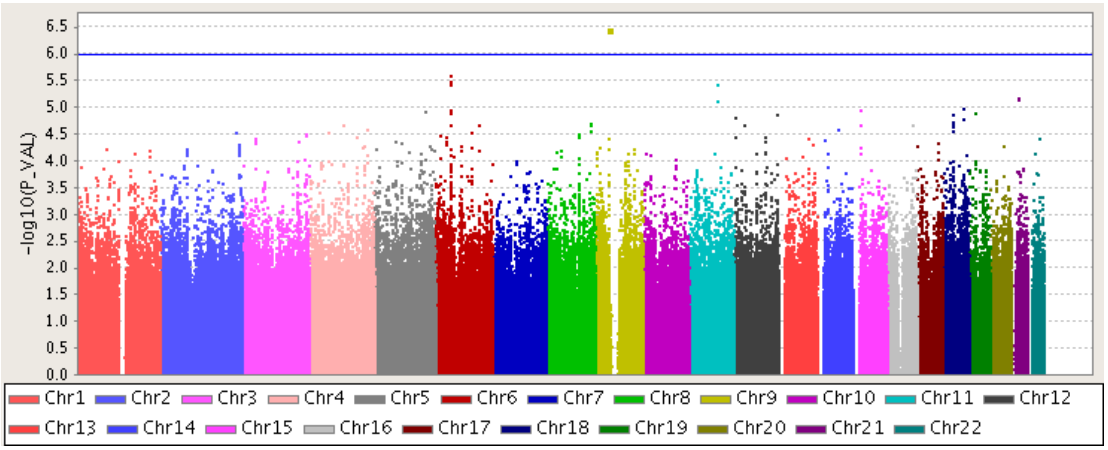
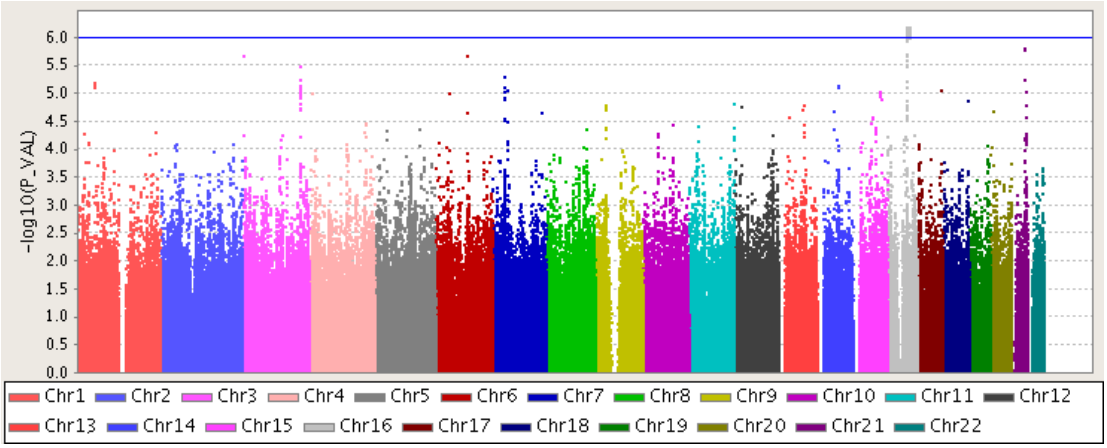


Figure A45: GWAS results for Wakeup. Analyses were run separately for African Americans (AFA), European Americans (EA), and Hispanic Americans (HIS). Individual SNP p-values located above the blue line are suggestive of genome-wide significance $-\log_{10}(p<1x10^{-6})$.

AFA



EA



HIS

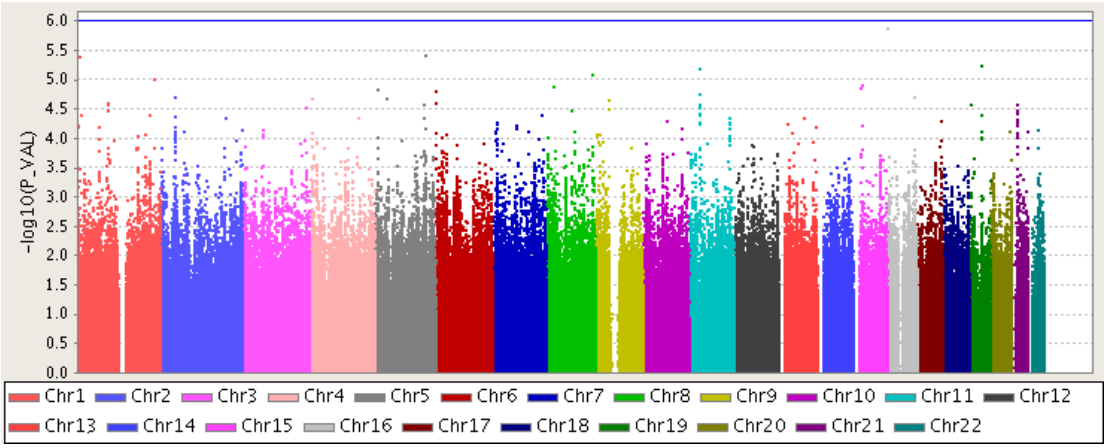
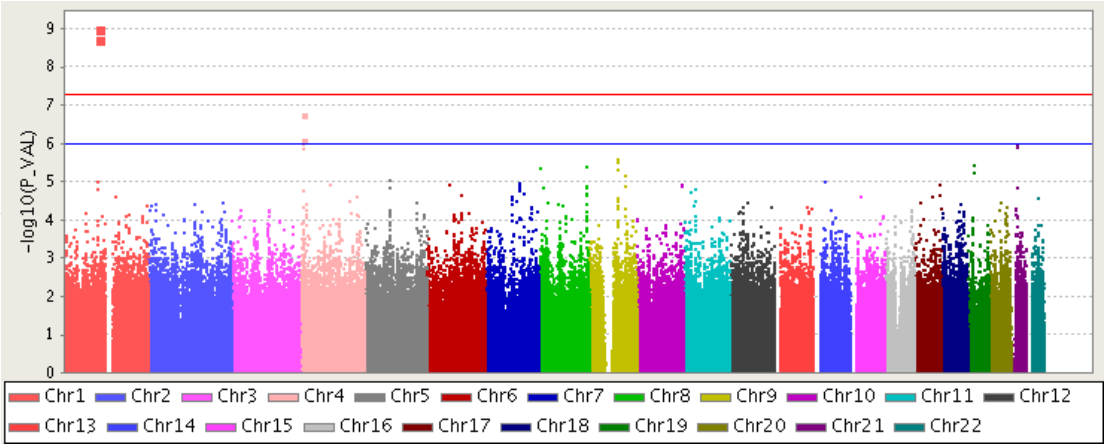
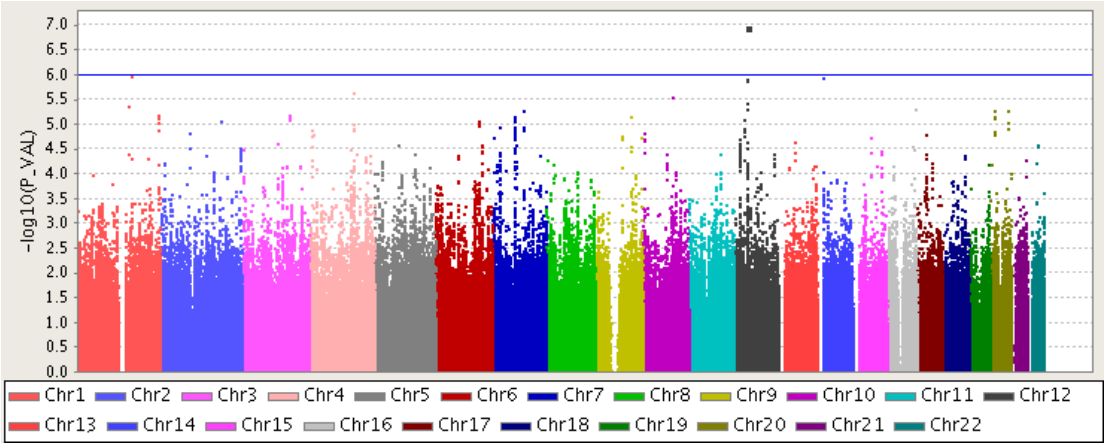


Figure A46: GWAS results for Bedtime. Analyses were run separately for African Americans (AFA), European Americans (EA), and Hispanic Americans (HIS). Individual SNP p-values located above the blue line are suggestive of genome-wide significance $-\log_{10}(p<1x10^{-6})$.

AFA



EA



HIS

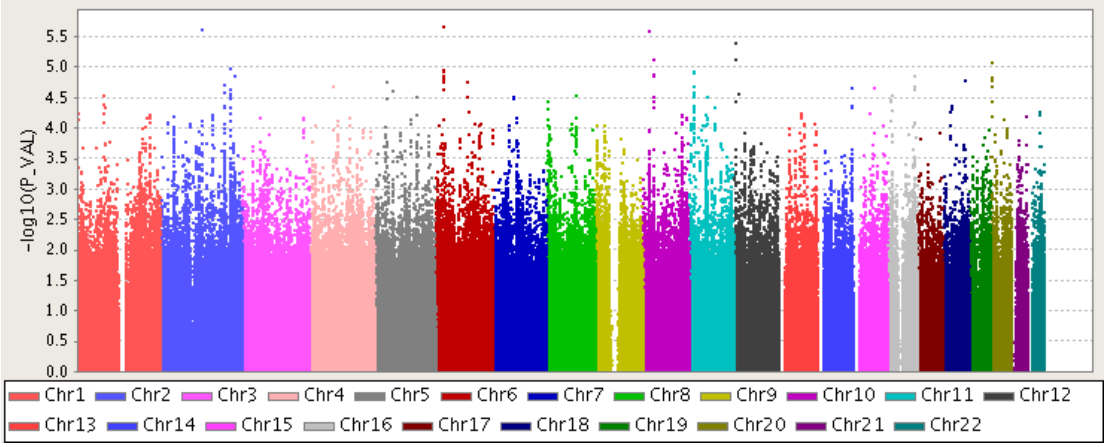
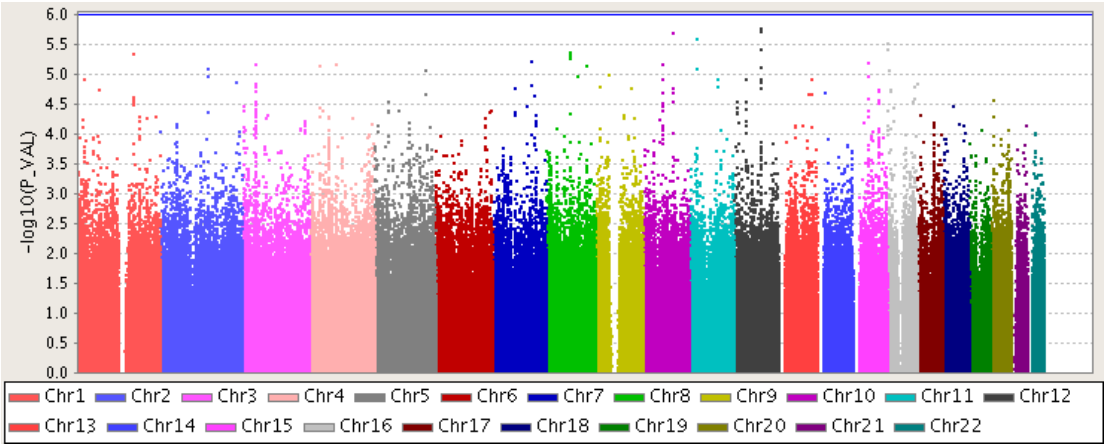
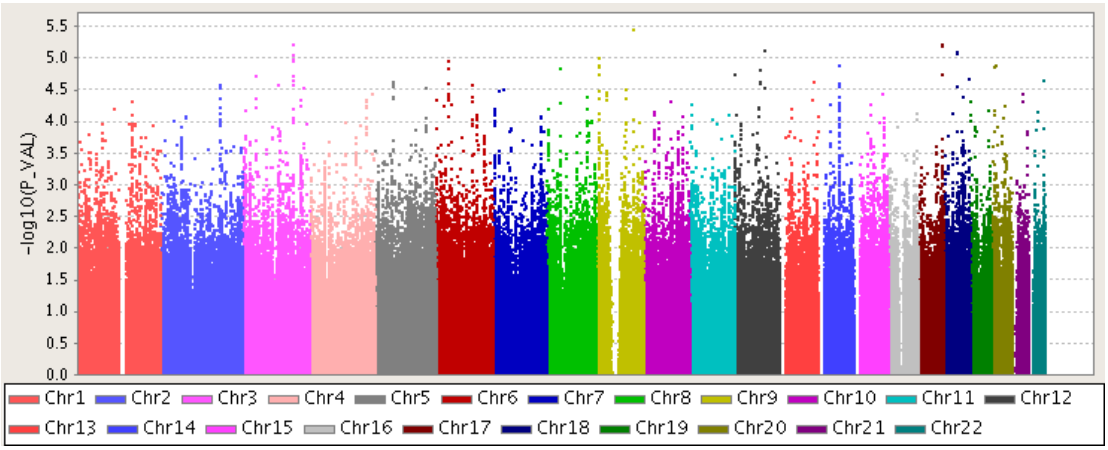


Figure A47: GWAS results for cortisol awakening response (CAR). Analyses were run separately for African Americans (AFA), European Americans (EA), and Hispanic Americans (HIS). Individual SNP p-values located above the blue line are suggestive of genome-wide significance $-\log_{10}(p<1x10^{-6})$. Individual SNP p-values located above the red line have reached genome-wide significance $-\log_{10}(p<5x10^{-8})$.

AFA



EA



HIS

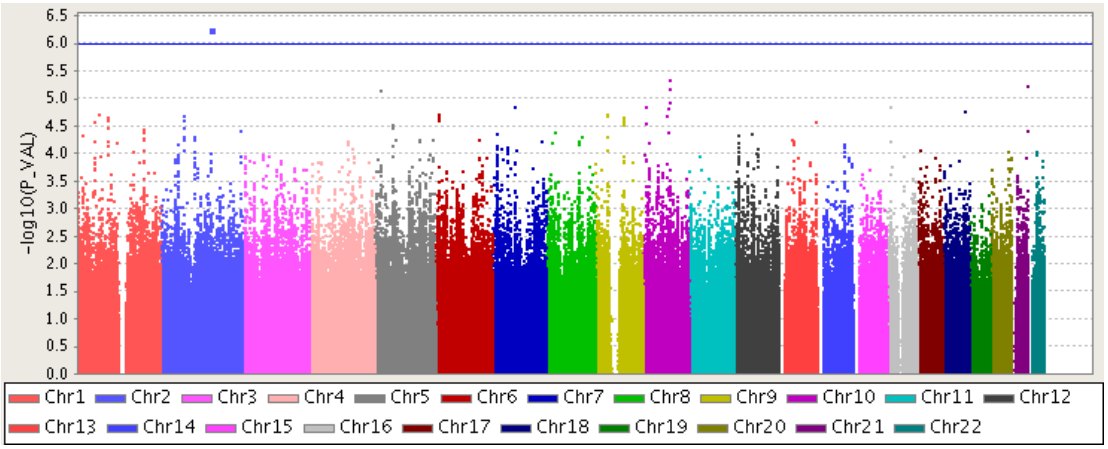
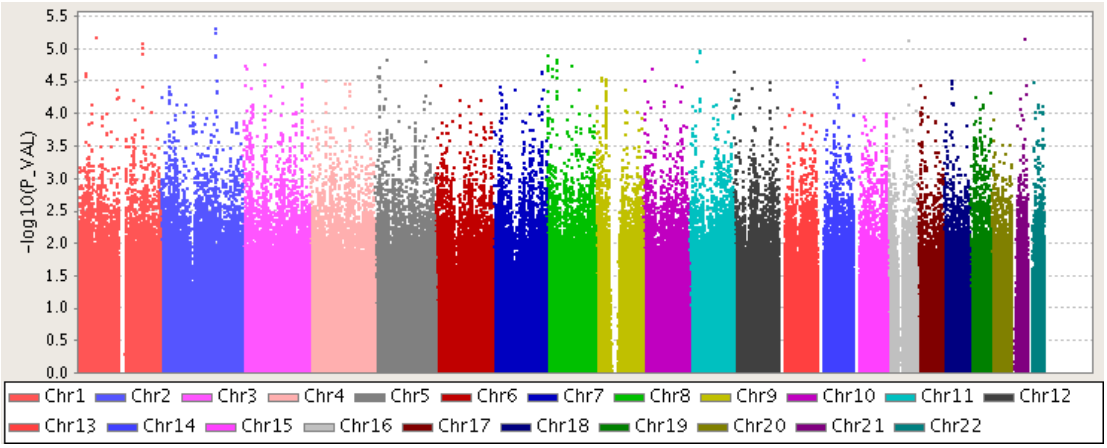
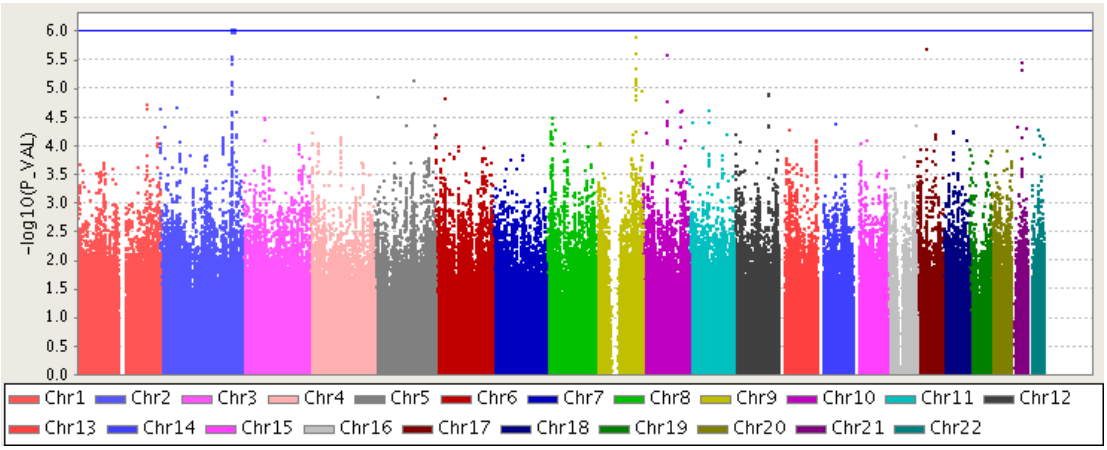


Figure A48: GWAS results for Area under the Curve (AUC). Analyses were run separately for African Americans (AFA), European Americans (EA), and Hispanic Americans (HIS). Individual SNP p-values located above the blue line are suggestive of genome-wide significance $-\log_{10}(p<1x10^{-6})$.

AFA



EA



HIS

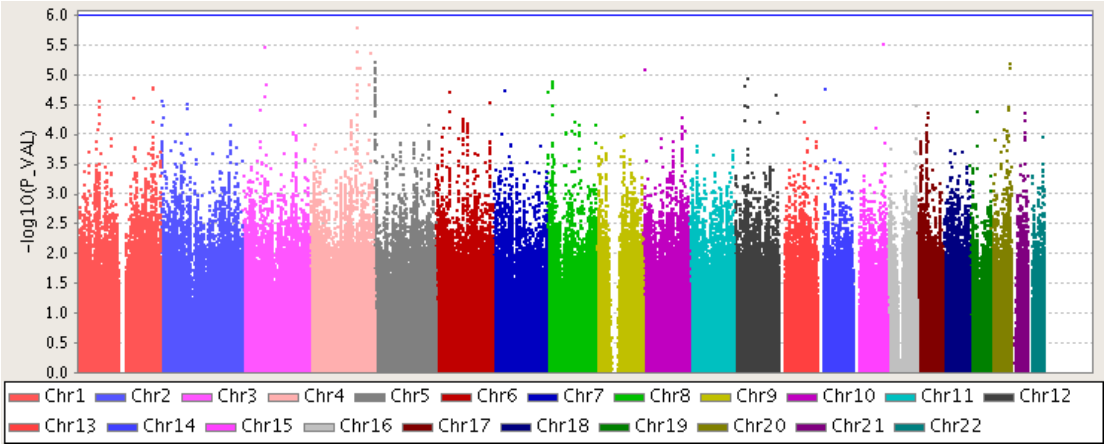
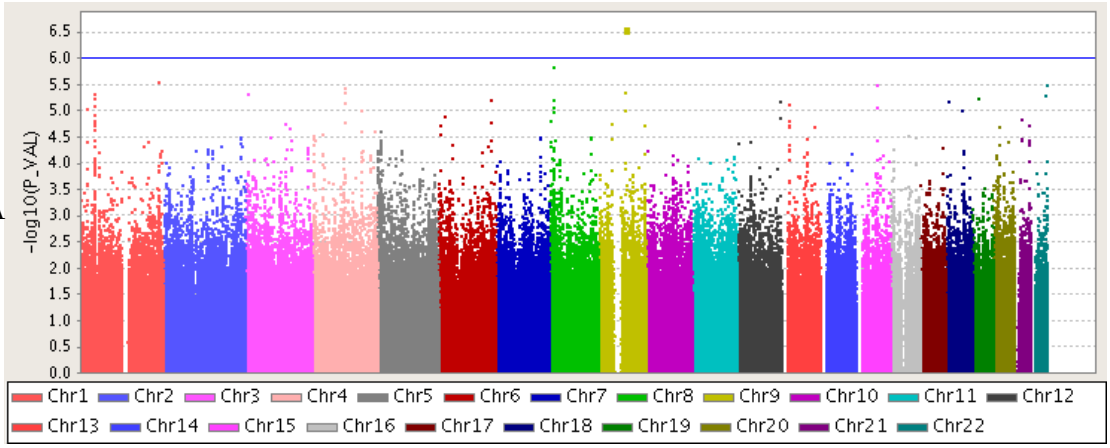
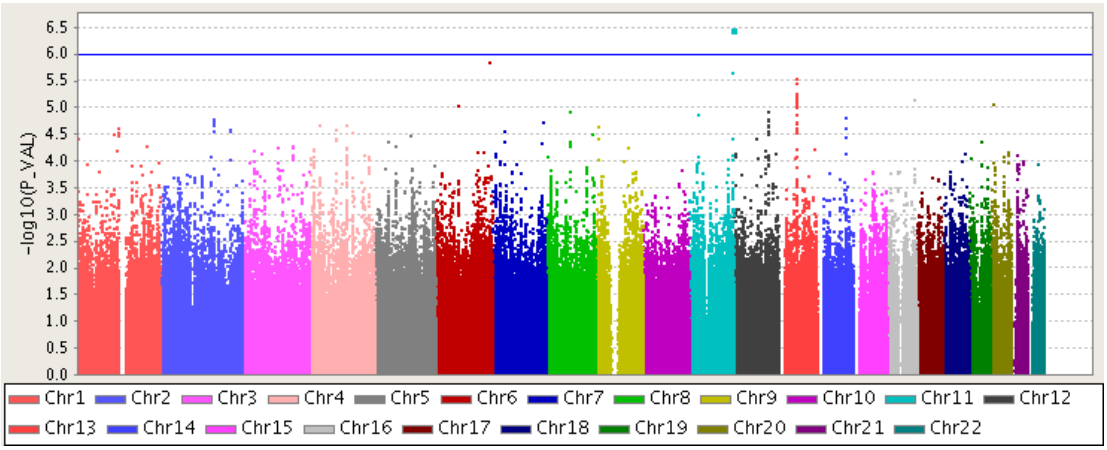


Figure A49: GWAS results for Early Decline Slope (EDSlope). Analyses were run separately for African Americans (AFA), European Americans (EA), and Hispanic Americans (HIS). Individual SNP p-values located above the blue line are suggestive of genome-wide significance $-\log_{10}(p<1x10^{-6})$.

AFA



EA



HIS

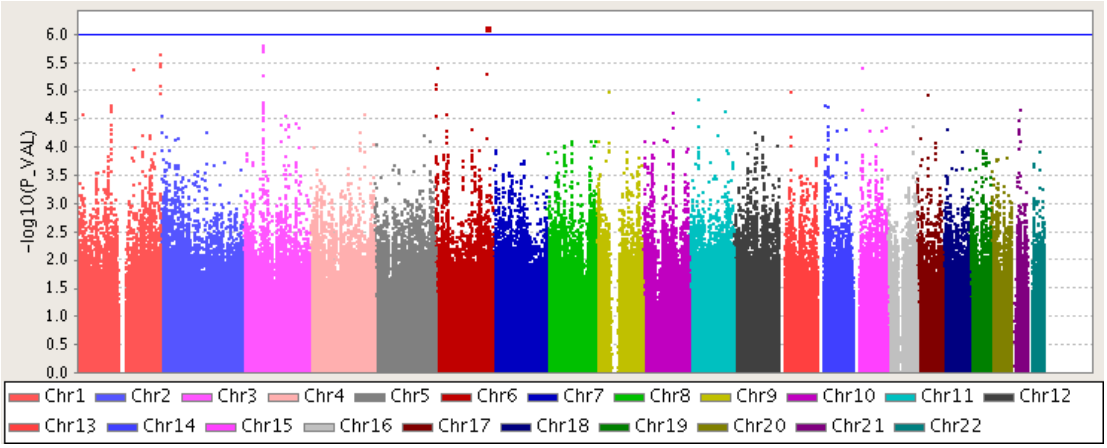
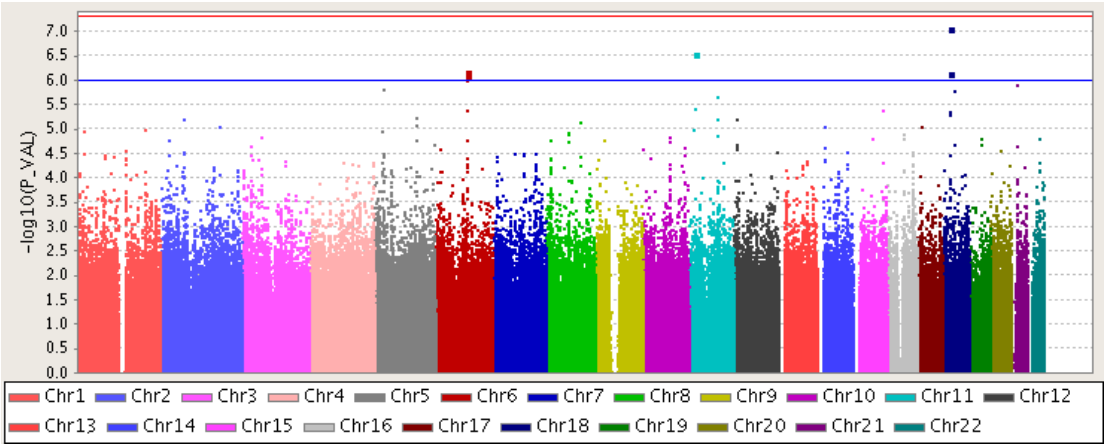
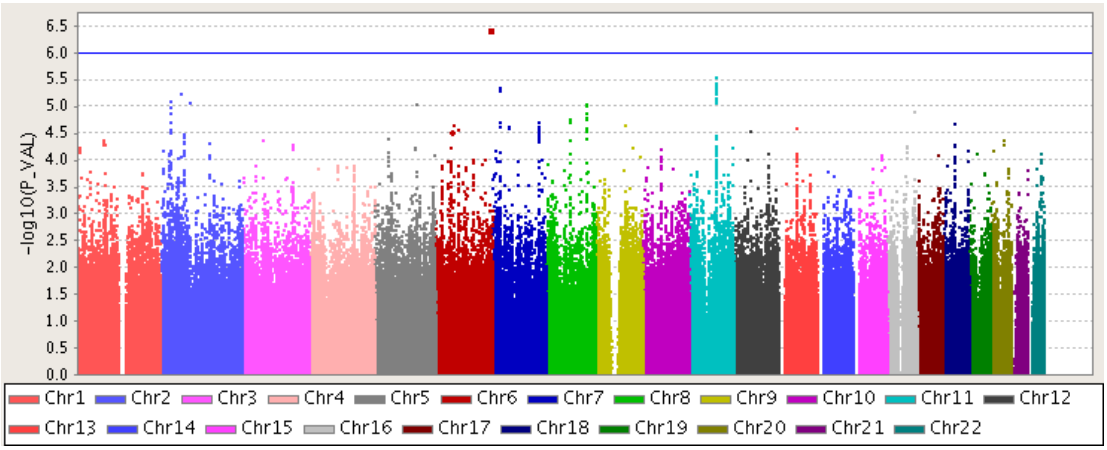


Figure A50: GWAS results for Late Decline Slope (LDSlope). Analyses were run separately for African Americans (AFA), European Americans (EA), and Hispanic Americans (HIS). Individual SNP p-values located above the blue line are suggestive of genome-wide significance $-\log_{10}(p<1x10^{-6})$.

AFA



EA



HIS

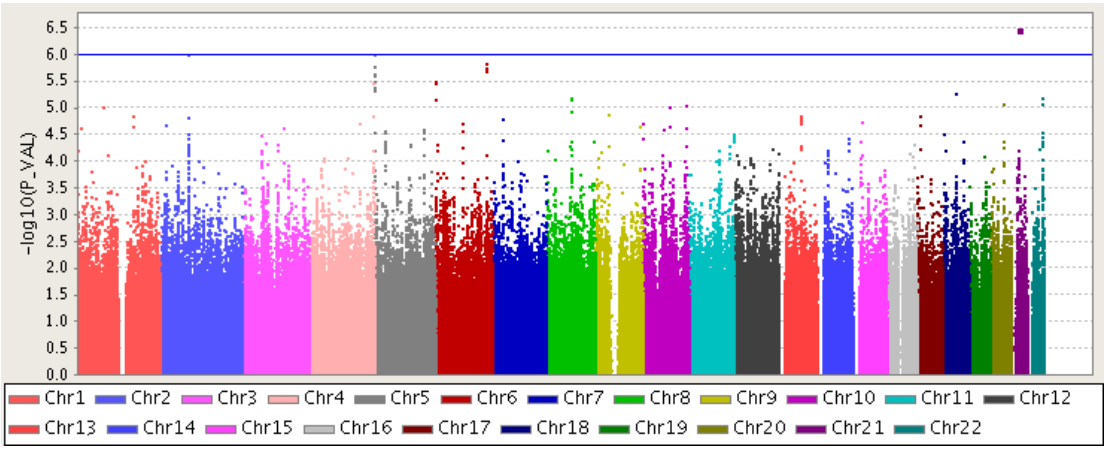


Figure A51: GWAS results for Overall Decline Slope (ODSlope). Analyses were run separately for African Americans (AFA), European Americans (EA), and Hispanic Americans (HIS). Individual SNP p-values located above the blue line are suggestive of genome-wide significance $-\log_{10}(p<1x10^{-6})$. Individual SNP p-values located above the red line have reached genome-wide significance $-\log_{10}(p<5x10^{-8})$.

REFERENCES

(1) Adam EK, Gunnar MR. Relationship functioning and home and work demands predict individual differences in diurnal cortisol patterns in women. *Psychoneuroendocrinology* 2001 Feb;26(2):189-208.

(2) Kirschbaum C, Hellhammer DH. Salivary cortisol in psychobiological research: an overview. *Neuropsychobiology* 1989;22(3):150-169.

(3) Pruessner JC, Wolf OT, Hellhammer DH, Buske-Kirschbaum A, von Auer K, Jobst S, et al. Free cortisol levels after awakening: a reliable biological marker for the assessment of adrenocortical activity. *Life Sci.* 1997;61(26):2539-2549.

(4) Knorr U, Vinberg M, Kessing LV, Wetterslev J. Salivary cortisol in depressed patients versus control persons: a systematic review and meta-analysis. *Psychoneuroendocrinology* 2010 Oct;35(9):1275-1286.

(5) Cohen S, Schwartz J, Epel E, Kirschbaum C, Sidney S, Seeman T. Socioeconomic status, race, and diurnal cortisol decline in the Coronary Artery Risk Development in Young Adults (CARDIA) Study. Psychosomatic medicine 2006;68(1):41-50.

(6) DeSantis A, Adam E, Doane L, Mineka S, Zinbarg R, Craske M. Racial/ethnic differences in cortisol diurnal rhythms in a community sample of adolescents. Journal of adolescent health 2007;41(1):3-13.

(7) Steptoe A, Kunz-Ebrecht S, Owen N, Feldman P, Willemsen G, Kirschbaum C, et al. Socioeconomic status and stress-related biological responses over the working day. Psychosomatic medicine 2003;65(3):461-470.

(8) Bjorntorp P, Holm G, Rosmond R. Hypothalamic arousal, insulin resistance and Type 2 diabetes mellitus. Diabetic medicine 1999;16(5):373-383.

(9) Rosmond R, Chagnon YC, Holm G, Chagnon M, Prusse L, Lindell K, et al. A glucocorticoid receptor gene marker is associated with abdominal obesity, leptin, and dysregulation of the hypothalamic-pituitary-adrenal axis. *Obes.Res.* 2000;8(3):211-218.

(10) Matthews K, Schwartz J, Cohen S, Seeman T. Diurnal cortisol decline is related to coronary calcification: CARDIA study. Psychosomatic medicine 2006;68(5):657-661.

(11) Champaneri S, Xu X, Carnethon MR, Bertoni AG, Seeman T, Roux AD, et al. Diurnal salivary cortisol and urinary catecholamines are associated with diabetes mellitus: the Multi-Ethnic Study of Atherosclerosis. *Metabolism* 2011 Dec 28.

(12) Desantis AS, Diezroux AV, Hajat A, Aiello AE, Golden SH, Jenny NS, et al. Associations of salivary cortisol levels with inflammatory markers: The Multi-Ethnic Study of Atherosclerosis. *Psychoneuroendocrinology* 2011 Dec 16.

(13) Ice GH. Factors influencing cortisol level and slope among community dwelling older adults in Minnesota. *J.Cross Cult.Gerontol.* 2005 Jun;20(2):91-108.

(14) Kirschbaum C, Kudielka BM, Gaab J, Schommer NC, Hellhammer DH. Impact of gender, menstrual cycle phase, and oral contraceptives on the activity of the hypothalamus-pituitary-adrenal axis. *Psychosom.Med.* 1999 Mar-Apr;61(2):154-162.

(15) Franz C, York T, Eaves L, Mendoza S, Hauger R, Hellhammer D, et al. Genetic and environmental influences on cortisol regulation across days and contexts in middle-aged men. *Behav.Genet.* 2010;40(4):467-479.

(16) Wust S, Federenko I, Hellhammer DH, Kirschbaum C. Genetic factors, perceived chronic stress, and the free cortisol response to awakening. *Psychoneuroendocrinology* 2000 Oct;25(7):707-720.

(17) Young EA, Aggen SH, Prescott CA, Kendler KS. Similarity in saliva cortisol measures in monozygotic twins and the influence of past major depression. *Biol.Psychiatry* 2000 Jul 1;48(1):70-74.

(18) Kirschbaum C, Wust S, Faig HG, Hellhammer DH. Heritability of cortisol responses to human corticotropin-releasing hormone, ergometry, and psychological stress in humans. *J.Clin.Endocrinol.Metab.* 1992 Dec;75(6):1526-1530.

(19) Froehlich JC, Zink RW, Li TK, Christian JC. Analysis of heritability of hormonal responses to alcohol in twins: beta-endorphin as a potential biomarker of genetic risk for alcoholism. *Alcohol.Clin.Exp.Res.* 2000 Mar;24(3):265-277.

(20) Bartels M, Van den Berg M, Sluyter F, Boomsma DI, de Geus EJC. Heritability of cortisol levels: review and simultaneous analysis of twin studies. *Psychoneuroendocrinology* 2003 2;28(2):121-137.

(21) Joels M, Karst H, DeRijk R, de Kloet ER. The coming out of the brain mineralocorticoid receptor. *Trends Neurosci.* 2008;31(1):1-7.

(22) DeRijk R, de Kloet ER. Corticosteroid receptor genetic polymorphisms and stress responsivity. *Endocrine* 2005;28(3):263-270.

(23) Derijk R, de Kloet ER. Corticosteroid receptor polymorphisms: determinants of vulnerability and resilience. *Eur.J.Pharmacol.* 2008;583(2-3):303-311.

(24) Huizenga NA, Koper JW, De Lange P, Pols HA, Stolk RP, Burger H, et al. A polymorphism in the glucocorticoid receptor gene may be associated with and increased sensitivity to glucocorticoids in vivo. *J.Clin.Endocrinol.Metab.* 1998;83(1):144-151.

(25) van Rossum EFC, Koper J, Huizenga NATM, Uitterlinden A, Janssen,Joop A M J L., Brinkmann A, et al. A polymorphism in the glucocorticoid receptor gene, which decreases sensitivity to glucocorticoids in vivo, is associated with low insulin and cholesterol levels. *Diabetes* 2002;51(10):3128-3134.

(26) van Rossum EFC, Voorhoeve P, te Velde SJ, Koper J, Delemarre-van de Waal H, Kemper HCG, et al. The ER22/23EK polymorphism in the glucocorticoid receptor gene is associated with a beneficial body composition and muscle strength in young adults. *J.Clin.Endocrinol.Metab.* 2004;89(8):4004-4009.

(27) LINDON RL. Stress diseases. *J.R.Inst.Public.Health.* 1952 Sep;15(9):231-243.

(28) RUSK HA. Stress in the world: the individual and the doctor. *Med.Ann.Dist.Columbia* 1958 May;27(5):259-262.

(29) Rosmond R, Dallman MF, Bjrntorp P. Stress-related cortisol secretion in men: relationships with abdominal obesity and endocrine, metabolic and hemodynamic abnormalities. The Journal of clinical endocrinology and metabolism 1998;83(6):1853-1859.

(30) McEwen BS. Protective and damaging effects of stress mediators: central role of the brain. *Dialogues Clin.Neurosci.* 2006;8(4):367-381.

(31) Hyman SE. How adversity gets under the skin. *Nat.Neurosci.* 2009 Mar;12(3):241-243.

(32) Young E, Abelson J, Lightman S. Cortisol pulsatility and its role in stress regulation and health. *Front.Neuroendocrinol.* 2004;25(2):69-76.

(33) Lightman SL, Windle RJ, Julian MD, Harbuz MS, Shanks N, Wood SA, et al. Significance of pulsatility in the HPA axis. *Novartis Found.Symp.* 2000;227:244-57.

(34) Conway-Campbell B, McKenna M, Wiles C, Atkinson H, de Kloet ER, Lightman S. Proteasome-dependent down-regulation of activated nuclear hippocampal glucocorticoid receptors determines dynamic responses to corticosterone. *Endocrinology* 2007;148(11):5470-5477.

(35) Reul JM, de Kloet ER. Two receptor systems for corticosterone in rat brain: microdistribution and differential occupation. *Endocrinology* 1985;117(6):2505-2511.

(36) de Kloet ER, Jols M, Holsboer F. Stress and the brain: from adaptation to disease. *Nature Reviews.Neuroscience* 2005;6(6):463-475.

(37) DeRijk R, Wust S, Meijer O, Zennaro M, Federenko I, Hellhammer D, et al. A common polymorphism in the mineralocorticoid receptor modulates stress responsiveness. *Journal of Clinical Endocrinology and Metabolism* 2006;91(12):5083-5089.

(38) Wust S, Van Rossum EF, Federenko I, Koper J, Kumsta R, Hellhammer D. Common polymorphisms in the glucocorticoid receptor gene are associated with adrenocortical responses to psychosocial stress. *J.Clin.Endocrinol.Metab.* 2004;89(2):565-573.

(39) Kumsta R, Entringer S, Koper J, van Rossum EFC, Hellhammer D, Wst S. Sex specific associations between common glucocorticoid receptor gene variants and hypothalamus-pituitary-adrenal axis responses to psychosocial stress. *Biol.Psychiatry* 2007;62(8):863-869.

(40) Kirschbaum C, Hellhammer DH. Salivary cortisol in psychoneuroendocrine research: recent developments and applications. *Psychoneuroendocrinology* 1994;19(4):313-333.

(41) Harris B, Watkins S, Cook N, Walker RF, Read GF, Riad-Fahmy D. Comparisons of plasma and salivary cortisol determinations for the diagnostic efficacy of the dexamethasone suppression test. *Biol.Psychiatry* 1990;27(8):897-904.

(42) Reid JD, Intrieri RC, Susman EJ, Beard JL. The relationship of serum and salivary cortisol in a sample of healthy elderly. *J.Gerontol.* 1992;47(3):P176-P179.

(43) Tunn S, Mllmann H, Barth J, Derendorf H, Krieg M. Simultaneous measurement of cortisol in serum and saliva after different forms of cortisol administration. *Clin.Chem.* 1992;38(8):1491-1494.

(44) Dallman MF. Stress update Adaptation of the hypothalamic-pituitary-adrenal axis to chronic stress. *Trends Endocrinol.Metab.* 1993 Mar;4(2):62-69.

(45) Chrousos GP. The concepts of stress and stress system disorders. Overview of physical and behavioral homeostasis. *JAMA : the journal of the American Medical Association* 1992 -03-04;267(9):1244-52.

(46) Ranjit N, Diez-Roux AV, Sanchez B, Seeman T, Shea S, Shrager S, et al. Association of salivary cortisol circadian pattern with cynical hostility: multi-ethnic study of atherosclerosis. *Psychosom.Med.* 2009 Sep;71(7):748-755.

(47) Do DP, Diez Roux AV, Hajat A, Auchincloss AH, Merkin SS, Ranjit N, et al. Circadian rhythm of cortisol and neighborhood characteristics in a population-based sample: the Multi-Ethnic Study of Atherosclerosis. *Health Place* 2011 Mar;17(2):625-632.

(48) Koper JW, Stolk RP, de Lange P, Huizenga NA, Molijn GJ, Pols HA, et al. Lack of association between five polymorphisms in the human glucocorticoid receptor gene and glucocorticoid resistance. *Hum.Genet.* 1997;99(5):663-668.

(49) Rao F, Zhang L, Wessel J, Zhang K, Wen G, Kennedy B, et al. Adrenergic polymorphism and the human stress response. *Ann.N.Y.Acad.Sci.* 2008;1148:282-296.

(50) Armando I, Jezova M, Bregonzio C, Baiardi G, Saavedra JM. Angiotensin II AT1 and AT2 receptor types regulate basal and stress-induced adrenomedullary catecholamine production through transcriptional regulation of tyrosine hydroxylase. *Ann.N.Y.Acad.Sci.* 2004;1018:302-309.

(51) Rao F, Zhang L, Wessel J, Zhang K, Wen G, Kennedy B, et al. Tyrosine hydroxylase, the rate-limiting enzyme in catecholamine biosynthesis: discovery of common human genetic variants governing transcription, autonomic activity, and blood pressure in vivo. *Circulation* 2007;116(9):993-1006.

(52) Barbeau P, Litaker M, Jackson R, Treiber F. A tyrosine hydroxylase microsatellite and hemodynamic response to stress in a multi-ethnic sample of youth. *Ethnicity disease* 2003;13(2):186-192.

(53) MacMillan LB, Hein L, Smith MS, Piascik MT, Limbird LE. Central hypotensive effects of the alpha2a-adrenergic receptor subtype. *Science* 1996;273(5276):801-803.

(54) Finley JC, O'Leary M, Wester D, MacKenzie S, Shepard N, Farrow S, et al. A genetic polymorphism of the alpha2-adrenergic receptor increases autonomic responses to stress. *J.Appl.Physiol.* 2004;96(6):2231-2239.

(55) Schwartz A, Gerin W, Davidson K, Pickering T, Brosschot J, Thayer J, et al. Toward a causal model of cardiovascular responses to stress and the development of cardiovascular disease. *Psychosom.Med.* 2003;65(1):22-35.

(56) Li GH, Faulhaber HD, Rosenthal M, Schuster H, Jordan J, Timmermann B, et al. Beta-2 adrenergic receptor gene variations and blood pressure under stress in normal twins. *Psychophysiology* 2001;38(3):485-489.

(57) Caspi A, Sugden K, Moffitt T, Taylor A, Craig I, Harrington H, et al. Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* 2003;301(5631):386-389.

(58) Koenen KC, Aiello AE, Bakshis E, Amstadter AB, Ruggiero KJ, Acierno R, et al. Modification of the association between serotonin transporter genotype and risk of posttraumatic stress disorder in adults by county-level social environment. *Am.J.Epidemiol.* 2009 Mar 15;169(6):704-711.

(59) Stein M, Schork N, Gelernter J. Gene-by-environment (serotonin transporter and childhood maltreatment) interaction for anxiety sensitivity, an intermediate phenotype for anxiety disorders. *Neuropsychopharmacology* 2008;33(2):312-319.

(60) Anguelova M, Benkelfat C, Turecki G. A systematic review of association studies investigating genes coding for serotonin receptors and the serotonin transporter: II. Suicidal behavior. *Mol.Psychiatry* 2003;8(7):646-653.

(61) Risch N, Herrell R, Lehner T, Liang KY, Eaves L, Hoh J, et al. Interaction between the serotonin transporter gene (5-HTTLPR), stressful life events, and risk of depression: a meta-analysis. *JAMA* 2009 Jun 17;301(23):2462-2471.

(62) Wust S, Kumsta R, Treutlein J, Frank J, Entringer S, Schulze T, et al. Sex-specific association between the 5-HTT gene-linked polymorphic region and basal cortisol secretion. *Psychoneuroendocrinology* 2009;34(7):972-982.

(63) Hariri A, Mattay V, Tessitore A, Kolachana B, Fera F, Goldman D, et al. Serotonin transporter genetic variation and the response of the human amygdala. *Science* 2002;297(5580):400-403.

(64) Williams RB, Marchuk DA, Gadde KM, Barefoot JC, Grichnik K, Helms MJ, et al. Central nervous system serotonin function and cardiovascular responses to stress. *Psychosom.Med.* 2001;63(2):300-305.

(65) Arnaldi G, Angeli A, Atkinson AB, Bertagna X, Cavagnini F, Chrousos GP, et al. Diagnosis and complications of Cushing's syndrome: a consensus statement. The Journal of clinical endocrinology and metabolism 2003;88(12):5593-5602.

(66) Pasquali R, Cantobelli S, Casimirri F, Capelli M, Bortoluzzi L, Flamia R, et al. The hypothalamic-pituitary-adrenal axis in obese women with different patterns of body fat distribution. The Journal of clinical endocrinology and metabolism 1993;77(2):341-346.

(67) Weaver JU, Kopelman PG, McLoughlin L, Forsling ML, Grossman A. Hyperactivity of the hypothalamo-pituitary-adrenal axis in obesity: a study of ACTH, AVP, beta-lipotrophin and cortisol responses to insulin-induced hypoglycaemia. Clinical endocrinology 1993;39(3):345-350.

(68) Chiodini I, Torlontano M, Scillitani A, Arosio M, Bacci S, Di Lembo S, et al. Association of subclinical hypercortisolism with type 2 diabetes mellitus: a case-control study in hospitalized patients. European journal of endocrinology 2005;153(6):837-844.

(69) Chiodini I, Adda G, Scillitani A, Coletti F, Morelli V, Di Lembo S, et al. Cortisol secretion in patients with type 2 diabetes: relationship with chronic complications. Diabetes Care 2007;30(1):83-88.

(70) Rosmond R, Wallerius S, Wanger P, Martin L, Holm G, Bjrntorp P. A 5-year follow-up study of disease incidence in men with an abnormal hormone pattern. Journal of internal medicine 2003;254(4):386-390.

(71) Hautanen A, Rikknen K, Adlercreutz H. Associations between pituitary-adrenocortical function and abdominal obesity, hyperinsulinaemia and dyslipidaemia in normotensive males. Journal of internal medicine 1997;241(6):451-461.

(72) Duclos M, Marquez Pereira P, Barat P, Gatta B, Roger P. Increased cortisol bioavailability, abdominal obesity, and the metabolic syndrome in obese women. *Obes.Res.* 2005 Jul;13(7):1157-1166.

(73) Ranjit N, Diez-Roux A, Shea S, Cushman M, Seeman T, Jackson S, et al. Psychosocial factors and inflammation in the multi-ethnic study of atherosclerosis. Archives of internal medicine 2007;167(2):174-181.

(74) Jain S, Mills P, von Knel R, Hong S, Dimsdale J. Effects of perceived stress and uplifts on inflammation and coagulability. Psychophysiology 2007;44(1):154-160.

(75) Weik U, Herforth A, Kolb-Bachofen V, Deinzer R. Acute stress induces proinflammatory signaling at chronic inflammation sites. Psychosomatic medicine 2008;70(8):906-912.

(76) Miller G, Chen E, Sze J, Marin T, Arevalo JMG, Doll R, et al. A functional genomic fingerprint of chronic stress in humans: blunted glucocorticoid and increased NF-kappaB signaling. Biological psychiatry 2008;64(4):266-272.

(77) Cole S, Hawkley L, Arevalo J, Sung C, Rose R, Cacioppo J. Social regulation of gene expression in human leukocytes. GenomeBiology.com 2007;8(9):R189-R189.

(78) Bild D, Bluemke D, Burke G, Detrano R, Diez Roux AV, Folsom A, et al. Multi-ethnic study of atherosclerosis: objectives and design. *Am.J.Epidemiol.* 2002;156(9):871-881.

(79) Kudielka B, Broderick J, Kirschbaum C. Compliance with saliva sampling protocols: electronic monitoring reveals invalid cortisol daytime profiles in noncompliant subjects. *Psychosom.Med.* 2003;65(2):313-319.

(80) Hajat A, Diez-Roux A, Franklin TG, Seeman T, Shrager S, Ranjit N, et al. Socioeconomic and race/ethnic differences in daily salivary cortisol profiles: The Multi-Ethnic Study of Atherosclerosis. *Psychoneuroendocrinology* 2010 7;35(6):932-943.

(81) Rosner B. Percentage Points for a Generalized ESD Many-Outlier Procedure. *Technometrics* 1983;25(2):165-172.

(82) IMPUTE2. Available at: http://mathgen.stats.ox.ac.uk/impute/impute_v2.html#home.

(83) DeSantis AS, DiezRoux AV, Hajat A, Golden SH, Jenny NS, Sanchez BN, et al. Associations of salivary cortisol levels with metabolic syndrome and its components: the multi-ethnic study of atherosclerosis. *J.Clin.Endocrinol.Metab.* 2011 Nov;96(11):3483-3492.

(84) Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X. Rare-variant association testing for sequencing data with the sequence kernel association test. *Am.J.Hum.Genet.* 2011 Jul 15;89(1):82-93.

(85) Champaneri S, Xu X, Carnethon MR, Bertoni AG, Seeman T, Desantis AS, et al. Diurnal salivary cortisol is associated with body mass index and waist circumference: The multi-ethnic study of atherosclerosis. *Obesity (Silver Spring)* 2012 Oct 3.

(86) Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. *Genome Res.* 2002 Jun;12(6):996-1006.

(87) R Development Core Team. R: A language and environment for statistical computing. 2011;2.14.0.

(88) Pruim RJ. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* 2010 -09-15;26(18):2336-7.

(89) SNPTest version 2. Available at: https://mathgen.stats.ox.ac.uk/genetics_software/snptest/old/snptest.html.

(90) Golden Helix I. SNP & Variation Suite. ;7.

(91) Morgenthaler S, Thilly WG. A strategy to discover genes that carry multi-allelic or mono-allelic risk for common diseases: a cohort allelic sums test (CAST). *Mutat.Res.* 2007 Feb 3;615(1-2):28-56.

(92) Madsen BE, Browning SR. A groupwise association test for rare mutations using a weighted sum statistic. *PLoS Genet.* 2009 Feb;5(2):e1000384.

(93) Neale BM, Rivas MA, Voight BF, Altshuler D, Devlin B, Orho-Melander M, et al. Testing for an unusual distribution of rare variants. *PLoS Genet.* 2011 Mar;7(3):e1001322.

(94) Broderick JE, Arnold D, Kudielka BM, Kirschbaum C. Salivary cortisol sampling compliance: comparison of patients and healthy volunteers. *Psychoneuroendocrinology* 2004 Jun;29(5):636-650.

(95) Lander ES, Schork NJ. Genetic dissection of complex traits. *Science* 1994 Sep 30;265(5181):2037-2048.

(96) Goddard KA, Hopkins PJ, Hall JM, Witte JS. Linkage disequilibrium and allele-frequency distributions for 114 single-nucleotide polymorphisms in five populations. *Am.J.Hum.Genet.* 2000 Jan;66(1):216-234.

(97) Bryc K, Velez C, Karafet T, Moreno-Estrada A, Reynolds A, Auton A, et al. Colloquium paper: genome-wide patterns of population structure and admixture among Hispanic/Latino populations. *Proc.Natl.Acad.Sci.U.S.A.* 2010 May 11;107 Suppl 2:8954-8961.

(98) Lin X, Lee S, Christiani DC, Lin X. Test for interactions between a genetic marker set and environment in generalized linear models. *Biostatistics* 2013 March 05.

(99) Veen G, Giltay EJ, van Vliet IM, Derijk RH, Klaassens ER, van Pelt J, et al. C-reactive protein polymorphisms are associated with the cortisol awakening response in basal conditions in human subjects. *Stress* 2011 Mar;14(2):128-135.

(100) Mattei J, Demissie S, Falcon LM, Ordovas JM, Tucker K. Allostatic load is associated with chronic conditions in the Boston Puerto Rican Health Study. *Soc.Sci.Med.* 2010 Jun;70(12):1988-1996.

(101) Lundberg U. Coping with Stress: Neuroendocrine Reactions and Implications for Health. *Noise Health.* 1999;1(4):67-74.

(102) Van Hulle CA, Shirtcliff EA, Lemery-Chalfant K, Goldsmith HH. Genetic and environmental influences on individual differences in cortisol level and circadian rhythm in middle childhood. *Horm.Behav.* 2012 Jun;62(1):36-42.

(103) Adam EK, Kumari M. Assessing salivary cortisol in large-scale, epidemiological research. *Psychoneuroendocrinology* 2009 Nov;34(10):1423-1436.

(104) Velders FP, Kuningas M, Kumari M, Dekker MJ, Uitterlinden AG, Kirschbaum C, et al. Genetics of cortisol secretion and depressive symptoms: a candidate gene and genome wide association approach. *Psychoneuroendocrinology* 2011 Aug;36(7):1053-1061.

(105) Kurina LM, Weiss LA, Graves SW, Parry R, Williams GH, Abney M, et al. Sex differences in the genetic basis of morning serum cortisol levels: genome-wide screen identifies two novel loci specific to women. *J.Clin.Endocrinol.Metab.* 2005 Aug;90(8):4747-4752.

(106) Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010 Sep 1;26(17):2190-2191.

(107) Huedo-Medina TB, Sanchez-Meca J, Marin-Martinez F, Botella J. Assessing heterogeneity in meta-analysis: Q statistic or I2 index? *Psychol.Methods* 2006 Jun;11(2):193-206.

(108) Urbanek P, Fetka I, Meisler MH, Busslinger M. Cooperation of Pax2 and Pax5 in midbrain and cerebellum development. *Proc.Natl.Acad.Sci.U.S.A.* 1997 May 27;94(11):5703-5708.

(109) O'Brien P, Morin P,Jr, Ouellette RJ, Robichaud GA. The Pax-5 gene: a pluripotent regulator of B-cell differentiation and cancer disease. *Cancer Res.* 2011 Dec 15;71(24):7345-7350.

(110) Luykx JJ, Vinkers CH, Bakker SC, Visser WF, van Boxmeer L, Strengman E, et al. A common variant in ERBB4 regulates GABA concentrations in human cerebrospinal fluid. *Neuropsychopharmacology* 2012 Aug;37(9):2088-2092.

(111) Iwakura Y, Nawa H. ErbB1-4-dependent EGF/neuregulin signals and their cross talk in the central nervous system: pathological implications in schizophrenia and Parkinson's disease. *Front.Cell.Neurosci.* 2013;7:4.

(112) Hollmen M, Liu P, Kurppa K, Wildiers H, Reinvall I, Vandorpe T, et al. Proteolytic processing of ErbB4 in breast cancer. *PLoS One* 2012;7(6):e39413.

(113) Paatero I, Lassus H, Junttila TT, Kaskinen M, Butzow R, Elenius K. CYT-1 isoform of ErbB4 is an independent prognostic factor in serous ovarian cancer and selectively promotes ovarian cancer cell growth in vitro. *Gynecol.Oncol.* 2013 Apr;129(1):179-187.

(114) Siegel DA, Huang MK, Becker SF. Ectopic dendrite initiation: CNS pathogenesis as a model of CNS development. *Int.J.Dev.Neurosci.* 2002 Jun-Aug;20(3-5):373-389.

(115) Crown Human Genome Center. GeneCards. ;3.0.

(116) Montpetit A, Boily G, Sinnett D. A detailed transcriptional map of the chromosome 12p12 tumour suppressor locus. *Eur.J.Hum.Genet.* 2002 Jan;10(1):62-71.

(117) Montpetit A, Larose J, Boily G, Langlois S, Trudel N, Sinnett D. Mutational and expression analysis of the chromosome 12p candidate tumor suppressor genes in pre-B acute lymphoblastic leukemia. *Leukemia* 2004 Sep;18(9):1499-1504.

(118) Stephens M, Balding DJ. Bayesian statistical methods for genetic association studies. *Nat.Rev.Genet.* 2009 Oct;10(10):681-690.

(119) Teo YY, Ong RT, Sim X, Tai ES, Chia KS. Identifying candidate causal variants via trans-population fine-mapping. *Genet.Epidemiol.* 2010 Nov;34(7):653-664.

(120) Kalman B, Lublin FD. The genetics of multiple sclerosis. A review. *Biomed.Pharmacother.* 1999 Sep;53(8):358-370.

(121) Fox ER, Musani SK, Barbalic M, Lin H, Yu B, Ogunyankin KO, et al. Genome-Wide Association Study of Cardiac Structure and Systolic Function in African Americans: The Candidate Gene Association Resource (CARe) Study. *Circ.Cardiovasc.Genet.* 2013 Feb 1;6(1):37-46.

(122) Peters U, Jiao S, Schumacher FR, Hutter CM, Aragaki AK, Baron JA, et al. Identification of Genetic Susceptibility Loci for Colorectal Tumors in a Genome-wide Meta-analysis. *Gastroenterology* 2012 Dec 22.

(123) Hancock DB, Artigas MS, Gharib SA, Henry A, Manichaikul A, Ramasamy A, et al. Genome-wide joint meta-analysis of SNP and SNP-by-smoking interaction identifies novel loci for pulmonary function. *PLoS Genet.* 2012 Dec;8(12):e1003098.

(124) International Consortium for Blood Pressure Genome-Wide Association Studies, Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* 2011 Sep 11;478(7367):103-109.

(125) Majewski J, Schwartzentruber J, Lalonde E, Montpetit A, Jabado N. What can exome sequencing do for you? *J.Med.Genet.* 2011 Sep;48(9):580-589.

(126) Lander ES. Initial impact of the sequencing of the human genome. *Nature* 2011 Feb 10;470(7333):187-197.

(127) Bao L, Cui Y. Prediction of the phenotypic effects of non-synonymous single nucleotide polymorphisms using structural and evolutionary information. *Bioinformatics* 2005 May 15;21(10):2185-2190.

(128) Burke DF, Worth CL, Priego EM, Cheng T, Smink LJ, Todd JA, et al. Genome bioinformatic analysis of nonsynonymous SNPs. *BMC Bioinformatics* 2007 Aug 20;8:301.

(129) Tian J, Wu N, Guo X, Guo J, Zhang J, Fan Y. Predicting the phenotypic effects of non-synonymous single nucleotide polymorphisms based on support vector machines. *BMC Bioinformatics* 2007 Nov 16;8:450.

(130) Jorgensen TJ, Ruczinski I, Kessing B, Smith MW, Shugart YY, Alberg AJ. Hypothesis-driven candidate gene association studies: practical design and analytical considerations. *Am.J.Epidemiol.* 2009 Oct 15;170(8):986-993.

(131) Wilkening SS. Is there still a need for candidate gene approaches in the era of genome-wide association studies? *Genomics (San Diego, Calif.)* 2009 -05;93(5):415-419.

(132) Baxevanis AD. The importance of biological databases in biological discovery. *Curr.Protoc.Bioinformatics* 2011 Jun;Chapter 1:Unit 1.1.

(133) Adriaens ME, Jaillard M, Waagmeester A, Coort SL, Pico AR, Evelo CT. The public road to high-quality curated biological pathways. *Drug Discov.Today* 2008 Oct;13(19-20):856-862.

(134) Teufel A, Krupp M, Weinmann A, Galle PR. Current bioinformatics tools in genomic biomedical research (Review). *Int.J.Mol.Med.* 2006 Jun;17(6):967-973.

(135) Freudenberg J, Propping P. A similarity-based method for genome-wide prediction of disease-relevant human genes. *Bioinformatics* 2002;18 Suppl 2:S110-5.

(136) Perez-Iratxeta C, Bork P, Andrade MA. Association of genes to genetically inherited diseases using data mining. *Nat.Genet.* 2002 Jul;31(3):316-319.

(137) Turner FS, Clutterbuck DR, Semple CA. POCUS: mining genomic sequence annotation to predict disease genes. *Genome Biol.* 2003;4(11):R75.

(138) Adie EA, Adams RR, Evans KL, Porteous DJ, Pickard BS. Speeding disease gene discovery by sequence based candidate prioritization. *BMC Bioinformatics* 2005 Mar 14;6:55.

(139) Aerts S, Lambrechts D, Maity S, Van Loo P, Coessens B, De Smet F, et al. Gene prioritization through genomic data fusion. *Nat.Biotechnol.* 2006 May;24(5):537-544.

(140) McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat.Rev.Genet.* 2008 May;9(5):356-369.

(141) Altshuler D, Daly M. Guilt beyond a reasonable doubt. *Nat.Genet.* 2007 Jul;39(7):813-815.

(142) Seldin MF, Pasaniuc B, Price AL. New approaches to disease mapping in admixed populations. *Nat.Rev.Genet.* 2011 Jun 28;12(8):523-528.

(143) Smith MW, O'Brien SJ. Mapping by admixture linkage disequilibrium: advances, limitations and guidelines. *Nat.Rev.Genet.* 2005 Aug;6(8):623-632.

(144) Florez JC, Price AL, Campbell D, Riba L, Parra MV, Yu F, et al. Strong association of socioeconomic status with genetic ancestry in Latinos: implications for admixture studies of type 2 diabetes. *Diabetologia* 2009 Aug;52(8):1528-1536.

(145) Sankararaman S, Sridhar S, Kimmel G, Halperin E. Estimating local ancestry in admixed populations. *Am.J.Hum.Genet.* 2008 Feb;82(2):290-303.

(146) Pasaniuc B, Sankararaman S, Kimmel G, Halperin E. Inference of locus-specific ancestry in closely related populations. *Bioinformatics* 2009 Jun 15;25(12):i213-21.

(147) Pugach I, Matveyev R, Wollstein A, Kayser M, Stoneking M. Dating the age of admixture via wavelet transform analysis of genome-wide data. *Genome Biol.* 2011;12(2):R19-2011-12-2-r19. Epub 2011 Feb 25.

(148) Tang H, Coram M, Wang P, Zhu X, Risch N. Reconstructing genetic ancestry blocks in admixed individuals. *Am.J.Hum.Genet.* 2006 Jul;79(1):1-12.

(149) Sundquist A, Fratkin E, Do CB, Batzoglou S. Effect of genetic divergence in identifying ancestral origin using HAPAA. *Genome Res.* 2008 Apr;18(4):676-682.

(150) Price A, Patterson N, Plenge R, Weinblatt M, Shadick N, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat.Genet.* 2006;38(8):904-909.

(151) Price AL, Tandon A, Patterson N, Barnes KC, Rafaels N, Ruczinski I, et al. Sensitive detection of chromosomal segments of distinct ancestry in admixed populations. *PLoS Genet.* 2009 Jun;5(6):e1000519.

(152) Paşaniuc B, Kennedy J, Măndoiu I. Imputation-based local ancestry inference in admixed populations. *Bioinformatics Research and Applications* 2009:221-233.

(153) Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat.Genet.* 2003 Feb;33(2):177-182.

(154) NCI-NHGRI Working Group on Replication in Association Studies, Chanock SJ, Manolio T, Boehnke M, Boerwinkle E, Hunter DJ, et al. Replicating genotype-phenotype associations. *Nature* 2007 Jun 7;447(7145):655-660.