2016

Examining Genetic and Environmental Influences on Alcohol use and Externalizing Behaviors in African American Adolescents

Neeru Goyal
Virginia Commonwealth University, goyaln@vcu.edu

Examining Genetic and Environmental Influences on Alcohol use and Externalizing Behaviors in African American Adolescents

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Human and Molecular Genetics at Virginia Commonwealth University School of Medicine

by

Neeru Goyal
Bachelor of Science, Virginia Commonwealth University 2009

Advisor:
Danielle M. Dick, PhD
Professor of Psychology, African American Studies, and Human & Molecular Genetics

Virginia Commonwealth University School of Medicine
Richmond, VA
March, 2016

Table of Contents

CHAPTER 1: General Introduction

Adolescence marks the developmental period in life where there exists the highest risk for initiating substance use[1], thus identifying adolescents as a high-risk group needing particular research attention to ultimately guide prevention and intervention efforts. There is a large body of literature examining the various types of substance use and its consequences; however, much of this literature focuses on European Americans[2, 3]. There is a scarcity of literature examining ethnic differences in relation to substance use, specifically, in populations of African Americans. However, research which include African American populations is of particular importance because African Americans are considered to be at greater risk for substance use related problems[1-4].

The heightened risk, in part, stems from their increased exposure to a variety of environmental stressors, (all of which have been examined extensively in relation to substance use disorders), such as higher rates of poverty, lower socioeconomic status, and decreased employment opportunities and access to resources/social services for growth and rehabilitation, in comparison to other racial/ethnic groups. A survey of children in the 100 largest metropolitan areas in the United States reported that 76% of African American children lived under worse circumstances than the worst off European Americans in those areas[5]. Higher rates of negative stereotypes and racial discrimination have been identified as one of the most important environmental stressors contributing to the health and physiological well-being among African Americans. Racial discrimination has been reported in almost every aspect of African American life, including mortgage lending, housing discrimination and residential segregation, employment opportunities, and health care access and responsiveness[6-9]. However, despite these

unique environmental stressors, studies find that African Americans are more likely to report later initiation of substance use and overall lower rates of use than their European American counterparts[9-10]. Subsequently, in order to understand this paradox, there is an emphasis on the importance of advancing research regarding substance use amongst African American populations. Moreover, it is imperative to understand the variability in developing substance use problems, which may be largely influenced by differences in use patterns and genetic diversity of the population.

Epidemiology

Alcohol. Rates, patterns, and problems associated with substance use during adolescence and young adulthood tend to vary by race. Substance use rates are much lower amongst African American youth in comparison to their European American counterparts[11]. From a nationally representative sample of 8th, 10th, and 12th graders surveyed on alcohol and other drug use, lifetime prevalence of alcohol use among African Americans was less than Caucasian Americans at all three grade levels[12]. According to data from the National Survey on Drug Use and health, African Americans, ages 12 to 20, reported past-month alcohol use at a rate of 17.3% compared to a national average of 22.8%. Furthermore, past month binge drinking among African American youth was 8.5% compared to a national average of 13.8%[11].

Surprisingly, despite having lower rates of alcohol use, African American adolescents and adults report higher levels of alcohol related problems than European Americans[9, 13], such as depression, violence, and suicide[14, 15]. Problematic alcohol use also places African Americans at a higher risk for developing drug-related medical disorders, such as alcoholic fatty liver disease, liver cirrhosis, smoking-related lung cancer, as well as, many other types of cancers (i.e. cancer of the mouth, larynx, etc.)[3]. Moreover, reports indicate that mortality rates for alcohol-related

disorders were 10% higher in African American populations than for other ethnic groups within the United States[15].

An explanation for these seemingly contradictory results is that though African American youth are more likely to initiate alcohol use at a later age than European American youth[15,17], once they have initiated use, they are less likely to desist use in comparison to European Americans[18,19]. It has also been suggested that polysubstance use may contribute to the higher risk of alcohol related problems in African Americans. For example, not only are African American youth three times as likely to initiate the use of illicit drugs before the use of alcohol and cigarettes, they are 2.3 times more likely to initiate use of both illicit drugs and alcohol or cigarettes within the same year as compared to European Americans[17]. Furthermore, African Americans may also be at a higher risk for developing alcohol related problems partly due to differences in ethanol metabolism seen in individuals of African ancestry. Increased conversion of ethanol to acetaldehyde and increased liver cirrhosis risk have been associated with alcohol-metabolizing functional polymorphisms found exclusively in populations with African ancestry[20-22].

Environment

Environmental factors have been well documented as having significant impacts on influencing substance use outcomes[23-26]. African Americans are potentially exposed to a unique set of environments in comparison to their European American counterparts. For example, African Americans are exposed to a variety of disadvantages, including but not limited to, poorer living conditions, fewer educational opportunities, lower income/economic wealth, and racial discrimination[27]. Furthermore, African American families are much more likely to report lower

levels of maternal or paternal education and living in a household with one parent instead of two[28].

However, it has been noted that historically, African Americans have considerably conservative attitudes towards alcohol use and abuse. Specifically, studies have reported that African American parents are more likely to disapprove of their children's use[9]. African Americans are also more restrictive about their alcohol use due to awareness of the particularly negative consequences (i.e. increased financial hardship, increased health problems, increased problems with the law) associated with its use[29,30]. European Americans, in comparison, are more vulnerable to both personal and interpersonal factors than African Americans[3].

Economic deprivation (i.e. being financially poor, living in poor-income and highly risky neighborhoods) is a contextual factor to which, generally, African Americans have more exposure than their European American counterparts[31]. Contrary to findings based on European American populations, it has been hypothesized that disadvantaged neighborhoods in which there are little economic advantages and higher rates of substance availability may potentially have a preventative influence among African Americans, instead of being a risk factor. Specifically, in such neighborhoods, African American adolescents are first-hand witnesses to the consequences of substance abuse and dependence and ultimately may be deterred from substance use[32,33].

There are also racial differences in alcohol expectancies (i.e. beliefs about the expected positive or negative effects of alcohol). Positive alcohol expectancies have been shown to be associated with earlier age at onset of drinking among European Americans and Hispanics[34]. However, African American female adolescents were less likely to endorse positive alcohol expectancies in comparison to European American female adolescents, thus appearing to have

more negative expectancies towards alcohol and other substance use than their European American counterparts[35]. As a result, it has been suggested that lowered positive expectancies and higher negative expectancies towards alcohol aid in hindering adolescents from alcohol use. This is reflected in the older age at onset of drinking seen amongst African Americans. In this manner, literature has documented many potentially protective environmental factors, also including cultural influences[17, 28] and racial identification[36], that explain the ethnic differences seen among African American youth. Accordingly, because there are inconsistencies in the literature regarding positive and negative risk factors that influence alcohol use among African Americans, it is also important to examine to what extent environmental influences should be given importance as underlying factors in the development of alcohol use disorders.

Twin Studies

In human studies, it can often be difficult to tease apart the effects of "nature" versus "nurture". Fortunately, there are two types of study designs that have the ability to distinguish genetic effects from environmental effects: adoption studies and twin studies. Adoption studies are a useful tool in examining the contributions of genetic and environmental influences on alcohol use disorders; however, studies containing a large sample size of African Americans are rare, and thus, African Americans have not been examined in adoption studies. Thus, we must turn to twin studies to examine the relative importance of genetic and environmental factors in the origin of alcohol use disorders.

Twin studies allow us to separate environmental influences into two categories: shared environmental factors (non-genetic, environmental factors that make a twin pair more similar on a specific phenotype, and unique environmental factors (those environmental factors that make members of a twin pair different). However, beyond assessing the contribution of environmental

influences to the origin of disease, twin studies also can examine the relative contributions of genetic influences on the variation of a given phenotype. In this manner, twin studies quantitatively examine the liability of developing substance use disorders as a result of genetic and/or environmental influences.

Statistically, twin studies are based on whether the resemblance between twins is a result of additive genetics (A), or environmental factors, which as described above, separate into common environmental (C) factors and unique environmental (E) factors. By comparing the degree of correlation between monozygotic (MZ) and dizygotic (DZ) twins, it is possible to estimate the influences that genetic and environmental factors will have, and statistically, whether an ACE or ADE (non-additive genetic factors which include dominance, denoted by D) model fits the data better[37, 38]. For this discussion of twin studies, the DZ twin correlations were more than half of the MZ twin correlation in all cases, and thus reflect results from fitting an ACE model.

Lack of Twin Studies in African Americans. To date, the majority of all twin studies have been carried out amongst European American populations. Of those that have been conducted in populations of African Americans, most are focused on female populations. Thus, it is unclear to what extent the findings from European American samples are generalizable to African Americans. Of those twin studies that do include minority twin pairs, often there are not enough African American twin pairs to accurately estimate variance components. For example, Prescott & colleagues[39] aimed to examine genetic variation in the etiology of alcoholism in European American and African American twin pairs from the Washington University Twin Study. Despite this being the first twin study to include a substantial number of African American twin participants (approximately 32% of the sample were African American), there

were still too few African American twin pairs to calculate valid genetic and environmental variance component estimates; and therefore, model fitting was only conducted in European American twins[39]. Therefore, in assessing existing findings from twin studies with African American samples, it is important to interpret the results as suggestive rather than definitive until research has expanded in this area.

Twin studies & alcohol use outcomes. With the aforementioned caveats in mind, some conclusions about alcohol use may still be drawn from the few twin studies that have been published. Although prevalence of substance use is lower and patterns of specific-substance use are different in African Americans compared to other ethnic groups, it can be confirmed that familial risk for problem drinking is a global risk factor for early initiation of substances. Early use has been confirmed as being moderately heritable across adolescents of both European American and African American populations[40].

However, EA and AA youth differ in the importance of environmental influences on initiation. Furthermore, from analyses conducted separately for alcohol, cigarettes, and cannabis use, a very modest or no role for shared environment on age at first use in African American females was found[40]. Shared environmental factors accounted for approximately 10% of the variance in age at first use of alcohol, but had no contribution towards cigarettes or cannabis. This is in contrast to findings from European Americans that have demonstrated that shared environmental influences account for a proportion of variance in initiation of alcohol, cigarettes, and cannabis[41, 42]. Sartor and colleagues[40] also found genetic influences to account for age at first use with 44% of variance in onset of alcohol use. In examining age at first use, the genetic contributions vary across the three substances to the same extent as the variation seen amongst the three substances in European American samples.

However, there also exist many variations in findings from twin studies which further justify the need for twin studies that are specific to African Americans. For example, in testing for sex-related transmission of alcoholism, inconsistent findings were yielded in different samples of European descent. Heath and colleagues[43] reported non-significant negative evidence to sex-specific transmission in an Australian twin sample; whereas McGue and colleagues[44] reported positive non-significant results from the Minnesota twin sample. Moreover, Prescott & Kendler[45] reported positive evidence for sex-specific transmission in the Virginia twin sample, but only when a broad definition of alcoholism was considered. These reported variations across studies of European American populations raises concern based on current knowledge about the differences in the various phenotypes (i.e. lower prevalence amongst African Americans of alcohol use), and to what extent these findings from twin studies based on European Americans are applicable to African Americans. The vast documentation of differences in substance use among African Americans hints at potentially different pathways of risk[17]. Therefore, it is absolutely necessary to conduct twin studies in African American samples. Varying environments may have different impacts on this ethnic/racial group which could result in different environmental variance contributions to the risk of alcoholism and other substance use disorders. Moreover, if there are disparate and inconsistent findings from a less genetically diverse population such as European Americans), it has been suggested that the current estimates may not portend well for African American populations, who are more genetically diverse. Thus, findings from African American samples can potentially have additional implications on the risks to substance use and the development of substance use disorders.

Genetic Influences

Identifying the genes predisposing African American adolescents to substance use disorders has also become vital in understanding the etiology of these disorders, and in implementing prevention and intervention. Over a decade ago, the Human Genome Project identified approximately 20,500 genes and three billion base pairs in the genome. Among the methodologies to analyze the plethora of information that resulted from the Human Genome Project are genome-wide association studies (GWAS). GWAS scan the entire genome to identify common genetic variations in different individuals in order to identify the variant(s) associated with a trait or disease. The basis behind GWAS is that without *a priori* evidence of a specific gene, one can identify loci in the genome that underlie susceptibility to disease, by testing whether an allele is observed more frequently in persons with disease than in persons without disease.

To date, the majority of GWAS success has come from studies in populations of European ancestry[46,47]. GWAS has identified a vast number of single nucleotide polymorphisms (SNP) associated with various disease phenotypes, including type 2 diabetes, obesity, and various types of cancer. GWAS has also transformed the search for genetic risk variants underlying complex traits, identifying associations between over 3,000 SNPs and over 700 complex traits[48].

An important advantage also provided by the GWAS study design is that it provides a means by which researchers can control for population stratification. Population stratification occurs when there are differences in allele frequencies and trait prevalences that arise as a function of population membership or, in other words, the underlying structure of the population, rather than association of the marker with the disease. The impact of confounding by population stratification has often raised concern due to its potential to cause spurious associations or false

positive associations[49,50]. Therefore, the influence of population stratification remains a concern in genetic association studies. However, the advantage of GWAS is that there are such a large number of markers which can be tested across the genome to genetically determine ethnic background which can then consequently be used as a covariate to control for potential stratification issues.

It has been suggested that no single population is sufficient to identify every risk variant, underlying disease, applicable to all populations[47]. Further, because complex traits are the result of several genetic factors (i.e. polygenic) and the interplay between genetic and environmental factors, it has also been speculated that the genetic basis behind complex traits should map onto the evolution of the human genome and of human populations[47,48]. Because the genetic architecture of African Americans is different than that of European Americans, (reflecting greater exposure to evolutionary forces such as genetic drift, mutations, etc.), it remains pertinent to determine whether findings from GWAS based primarily on individuals of European descent are transferrable to African American populations.

The genetic architecture of African Americans and of other non-European populations differs in multiple ways which includes, but are not limited to, a greater genetic diversity reflected by population differences in allele frequencies, differences in linkage disequilibrium patterns, and haplotype structures. Additionally, effect sizes of risk variants are also different across ethnic groups[47,48,51]. However, it is important to keep in mind that GWAS of African Americans have found most success in examining smoking behaviors; thus the following discussion of reasons contributing to the limited availability of GWAS in African American populations will provide evidence analyzing smoking outcomes. However, there is no reason to

believe that these reasons for differences are not applicable in the context of other adverse substance use outcomes.

Difference in Allele Frequencies. Just as prevalence of alcohol use and its related disorders vary considerably, allele frequencies vary across populations. Variations in allele frequency at any given location reflect a long history of recombination, mutations, and divergence of genealogical lineages, but ultimately affect the detectability of true risk variants underlying disease. Variations in frequency limit the contribution of a specific variant to disease susceptibility across different ethnic populations[47]. However, GWAS has gained its success with populations of European ancestry largely due to the fact that there is less variation in allele frequency in comparison to the variation in individuals of African (and other) ancestry. However, when studies aim to replicate significant associations between genetic variants and disease in a non-European population, this may not be possible because the risk allele, associated with susceptibility to disease, may be very rare if not absent in another population such as African Americans[52]. On the other hand, it is very likely that genetic variants that influence outcomes in African Americans are being missed in existing GWAS done in European populations. An example of this is the variant in the *KCNQ1* gene found for type 2 diabetes in East Asian Americans that had been missed in previous GWAS done in European American populations[48, 51]. In this case, the risk allele frequency was far less in European populations, and therefore, studies examining the variant in European populations did not have enough power to detect the association. The frequency of the risk allele was greater in East Asian populations, and therefore was found to be an important risk variant[52]. Further evidence of variation in allele frequency by population was observed by Adeyemo & Rotimi[52], who found that allele frequencies correlated better among groups that shared ancestry. Allele frequencies were similar

at most loci between specific groups of African American ancestry, between the European groups, and among the respective Southeast Asian groups. Accordingly, this provides evidence of the importance of differing relative frequencies of potential variants underlying disease susceptibility, and the necessity to have appropriate sample sizes and populations to reflect these variations in frequency.

Differences in Linkage Disequilibrium. Linkage Disequilibrium (LD) is a measure of the nonrandom association between alleles or haplotypes at a given location. An LD map is formed by taking a subset of identified and genotyped markers and "mapping" the LD patterns of the SNPs (i.e. whether two SNPs are in high LD with each other), which, as an approximation, reflects their likelihood to be inherited together[47, 51]. Because individuals of African ancestry have greater genetic variation than those of European ancestry, the likelihood is lower that two SNPs, within a given stretch of the chromosome, will be correlated or in high LD. That is, much lower levels of LD are found in populations of African Americans than in individuals of European or Asian ancestry. Therefore, the formation of these LD blocks varies greatly as a function of ethnicity. For example, in a pooled sample including both European American and African American subjects, 8 blocks were identified for the *GABBR2* gene. However, when the two ethnic groups were assessed separately, differences were seen in LD block structure, such that four separate blocks were identified among African Americans, whereas three blocks were identified at the same region among European Americans[53]. Similarly, within *GABBR1*, another GABA receptor gene, the block structure of the *GABBR1* SNPs differed in the same manner between the two ethnic groups and the pooled sample. In the pooled sample, two blocks were identified; but when assessed separately, two different blocks spanning the whole gene were identified among the African Americans, and only one block was identified in the European

American sample located at the 3' end of the gene[53]. LD structure in European American populations may not replicate in African American populations because the association between a SNP and the true causal variant may not be as high in African Americans as this relationship may be in a population of European Americans. Due to such differences in LD findings, by relying on findings from European American populations, we limit the opportunity to find genes which influence alcohol use and the development of alcohol use related outcomes in an ethnically mismatched population (i.e. African American populations).

Differences in Haplotype Structure. LD maps also indicate whether a subset of SNPs at a given locus of a chromosome are in high LD with each other. In this manner, analyzing one or a few of such SNPs is often informative of multiple SNPs at that given location. Haplotype-based association analyses provide evidence of the differences seen in haplotype structure as a function of ethnicity. For example, in a study to understand the etiology of nicotine dependence (ND) among European American and African American populations, Li and colleagues[53] found that when the European American and African Americans subsamples were assessed separately, the haplotypes with which there was a significant association to ND, as measured by the Fagerström Test for Nicotine Dependence (FTND) and its related phenotypes (i.e. heaviness of smoking as defined by two separate measures) differed. Firstly, there were differences in the haplotype location in the gene. In the European American sample, haplotypes (significantly associated to ND) existed in the introns among the transmembrane and exons encoding the cytoplasmic domain of *GABBR2*, whereas in the African American sample, haplotypes existed in the introns among exons encoding the binding domain of *GABBR2*[53].

Additionally, distinctions among the ethnic groups were seen in directions of the haplotype effects on FTND and the measure assessing heaviness of smoking. The A-A-T

haplotype of a 3-SNP combination, that was found to be significant to one of the three measures of ND and heaviness of smoking, had a protective effect against ND among African Americans. However, in the European American sample, the G-C-C haplotype from this combination served as a "risk haplotype" for developing ND[53]. Such findings reiterate that genetic vulnerability to substance use and its related outcomes differ in subjects of European and African ancestry. This further emphasizes the importance of conducting gene finding studies in samples of African Americans to better understand the etiology of developing alcohol use disorders in this genetically different population, which may not be achieved when genetic studies are limited to samples of European American individuals.

Differences in Effect Size. Lastly, effect sizes for variants reported to be associated with smoking behaviors in larger GWAS of European ancestry often tend to be notably smaller in African American populations[47]. These differing effect sizes suggest that there are most likely additional loci and potentially hundreds of more risk variants of small effect sizes that underlie risk for complex traits. However, these remain undiscovered because in order to detect such variants, much larger power and inclusion of individuals of other ethnicities is required[47].

Genetic similarities between European Americans & African Americans, & its implications. The aforementioned discussion identifies the many documented differences between the genetic architecture of European Americans and African Americans to clearly warrant expanding GWAS to further include African Americans. However, similarities do exist among the findings that give rise to the potential of portability of findings between the two racial groups. In the very first meta-analysis including over 30,000 subjects of African ancestry (66.1% female), multiple SNPs across a variety of genes, reported in risk for substance use behaviors (specifically nicotine use) from GWAS of European ancestry, were found to be significantly

correlated to functional variants in populations of African ancestry[19]. One such SNP, rs2036527, located near the 5' end of gene *CHRNA5* on chromosome 15q25.1 (a region previously identified to be strongly associated to smoking related phenotypes, including cigarettes smoked per day[54]), showed a genome-wide significant association with number of cigarettes smoked per day. This SNP was also correlated with two index signals reported in previous GWAS of European ancestry[19]. Thus, studies including populations of African ancestry not only have the potential to discover new variants, but to also confirm existing risk variants and locations from GWAS of European ancestry. That is, findings from this meta-analysis of African populations and previous findings from GWAS of European ancestry both identified regions of chromosome 15q25 as significantly associated with smoking behaviors.

These findings also imply that such association signals may originate from the same causal variant[48]. In other words, disease etiology, from a genetic standpoint, may be common across different ethnic populations, but perhaps specific risk variants are population-specific[48]. With that said, current GWAS efforts are, in part, focused on multi-ethnic replication efforts for already identified risk variants with limited efforts on novel gene-findings.

However, it is important to note that replication attempts warrant concern. Replicability rates are much lower in African American populations. Whereas the replicability rates are high among individuals of European descent (85.6%), it has been calculated that the replicability rates in populations of African descent are as low as 9.6%, for reasons that are not attributable to statistical power[46]. Furthermore, for many reasons that have been previously discussed, it is difficult to ensure that the marker SNP that is tagging the causal variant in one population is the same in a different ethnic population.

In summary, GWAS holds various strengths and weaknesses, which aid and inhibit the ability to fully characterize complex traits. First, GWAS findings based on individuals of European ancestry have limited ability to derive conclusions that also apply to African populations. Without taking into account the influence of factors such as platform used in the study and/or whether imputation (the ability to predict the genotypes at SNP positions that have not actually been sequenced) occurred, it cannot be concluded that the number of existing risk SNPs shared between populations is necessarily representative of the actual overlap in genetic background between populations. Second, findings from GWAS cannot be solely relied on to explain the intricacy of genetically heterogeneous traits; it should be noted that results of GWAS findings only reflect a small fraction of the total genetic risk. The majority of SNPs account for small changes in the phenotype (i.e. smoking quantity)[19].

However, GWAS findings do guide future directions in this line of research. Genomic efforts including larger association studies will aid in clarifying allele and haplotype frequencies, and the frequency and degree that existing and undiscovered risk variants are shared across ethnic groups. Such efforts thus require GWAS of larger sample sizes of African ancestry. Other efforts have also attempted to combine smaller cohorts into a large consortium to provide increased power in discovering new genetic associations for complex traits such as substance use disorders. Additionally, whereas GWAS holds power in identifying common variants for common traits, advancements in high throughput sequencing has allowed for the discovery of rare variants (with potentially larger effect sizes). Consequently, custom arrays are being developed to include rare variants so that we may estimate the extent to which rare variants influence risk for disease.

Gene x Environment. The very essence of substance use and its related traits is that they are complex traits, defined by the combined influences of genetic and environmental influences. Gene x environment (GxE) interactions refer to the interaction between genetic risk factors and environmental risk factors to affect the risk for development of substance use disorders[55]. Studies focusing solely on either genetic influences or environmental influences on alcohol use disorders are highly confounded by ignoring the influence of the other. For example, genetic studies that ignore environmental influences may be missing out on identifying specific genetic effects that underlie disorders, especially if such genetic effects are only present when an individual is concurrently exposed to the environmental influence[55, 56]. For example, in a population-based sample of Finnish twins, it was found that the importance of genetic factors on alcohol problems was enhanced among individuals with a higher level of education[57]. Additionally, using a sample of European and African Americans adolescent drinkers, a significant GxE interaction was found such that in individuals who reported none or limited peer drinking, a well-established genetic variant associated with alcohol use disorders (*ADH1B*) had a strong protective effect for adolescent alcohol drinking behaviors. However, in the presence of high levels of peer drinking, the protective effects were greatly attenuated[58]. Salvatore and colleagues[59] similarly found that not only did higher polygenic scores predict a greater number of alcohol problems in a sample of European descent, but also these genetic influences were significantly more pronounced under conditions of low parental knowledge or high peer deviance. These findings demonstrate examples of the vast literature documenting the role of specific environmental factors moderating the importance of genetic effects, and/or vice versa.

Conclusion

This chapter provides evidence that the epidemiology (i.e. difference in prevalence rates and substance use patterns), and environmental (i.e. varying effects by social, cultural, and interpersonal factors) and genetic (i.e. greater genetic diversity and varying effect sizes of risk variants) determinants underlying susceptibility to developing substance abuse or dependence may vary between subjects of European and African ancestry. Furthermore, advanced methodological designs (such as twin studies, GWAS, sequencing of rare variants, and GxE interactions) suggest the importance of extending this line of research to include members of under-represented ethnic groups. Specifically, future investigations, which include African American adolescent populations, will further refine the genetic and environmental underpinnings of complex traits such as alcohol and drug dependence.

Overview of current research

The aim of this thesis is to expand our understanding of the etiology of alcohol misuse and related disorders in African Americans using genetically informative study designs. Specifically, we take advantage of the candidate-gene approach and polygenic score analysis to extend the literature specific to African American populations. Chapter 2 explores gene x environment (GxE) interactions through the candidate gene approach to explore the relationship between two genes chosen on their potential relevance to stress response and adolescent alcohol use and misuse, among African American youth living in highly impoverished neighborhoods, as moderated by stressful life events. Chapter 3 implements polygenic score analyses to examine the effect of an aggregate of markers. We explore whether polygenic risk for alcohol dependence – derived from GWAS estimates in one discovery sample – predict alcohol use and broader externalizing behaviors and interact with stressful life events to predict alcohol use/misuse among high-risk African American youth.

CHAPTER 2: Genes involved in stress response and alcohol use among high-risk African American youth

Abstract

Background: Genetic and environmental factors influence substance use behaviors in youth. One of the known environmental risk factors is exposure to life stressors. The aim of this project is to study the interaction between *NR3C1* and *CRHBP*, genes thought to be involved in stress pathways, exposure to stressful life events, and adolescent alcohol use/misuse.
Methods: The sample included 541 African American individuals (ages 13-18) from the Genes, Environment, and Neighborhood Initiative, a subset of the Mobile Youth Survey sample from whom DNA and more extensive phenotypic data were collected. Participants were selected from high poverty neighborhoods in Mobile, Alabama with potential exposure to a variety of extreme life stressors.
Results: A measure of stressful life events was significantly predictive of alcohol use/misuse. In addition, this association was significantly dependent upon the number of putative risk variants at rs1715749, a SNP in *CRHBP* ($p \leq 0.006$). There was no significant interaction between *NR3C1* and stressful life events with respect to alcohol use/misuse, after taking into account multiple testing.
Conclusions: These findings suggest that *CRHBP* variants are potentially relevant for adolescent alcohol use/misuse among African American youth populations being reared within the context of stressful life events, and warrants replication.

Introduction

Adolescent substance use is a major public health problem, with alcohol being the most widely used drug by youth [60]. Alcohol use among African Americans is an area of major concern because alcohol use is related to three of the four leading causes of death among African Americans between the ages of 12 and 20, including homicide, unintentional injuries, and suicide[61]. Further, despite lower rates of substance use among African American adolescents, they display more problematic trajectories of drinking as they age into adulthood and consequently report higher levels of substance related problems than white Americans[62]. For example, though African American youth are more likely to initiate smoking at a much later age than white youth, once they have initiated use, they are less likely to desist use [63,64]. Consequently, African American youth are categorized as a population at greater risk for alcohol and substance use and misuse[65]. For example, rates of heavy drinking and alcohol-related problems remain high in African American individuals aged 18 to 29 as compared to European Americans[66]. African Americans are also more likely to face disadvantaged environmental conditions, such as poverty[65]. Yet, there is a scarcity of research examining adolescent substance use among African American youth living in high poverty neighborhoods[8].

Stressful Life Events

According to the National Comorbidity Replication Survey (NCS-R), 53% of adults have experienced some kind of major life stressor before the age of 18[67]. Of these stressors, the most common consist of parental divorce, family violence, economic adversity, parental death, and mental illness. While the biological stress response system is essential to human survival, it has been found that chronic or over-activation of the stress response system results in an increased

vulnerability for not only physiological problems, but also an increased risk for psychopathologies such as anxiety, depression, and alcohol and other drug dependence[68].

Adverse childhood events have been strongly related to alcohol use in early and mid-adolescence[69] and to the subsequent development of alcohol dependence[70]. Stressful life events (SLE) have also been linked to increased drug use over time among adolescents[71,72], and as a prominent predictor of early alcohol and drug use[73]. Consequently, in recent years, the role of SLE has been an area of increasing interest because of its noted influence on substance use outcomes. However, the associations between SLE and substance use in at-risk African American youth have received relatively little empirical attention. This represents an important gap in the literature because, in comparison to their white peers, African American youth experience higher rates of violence and poverty[74]. In the few studies that have examined SLE in African Americans, violent victimization was suggestive of playing an important role in prolonging substance use in a longitudinal study following African Americans ages 6 to 42[75]. Furthermore, Doherty and colleagues[75] also found that life-traumas involving coercion and force can also be highly predictive of drug dependence among both Caucasian and African American populations.

Influence of Genetic Factors

Substance use is not only influenced by environmental factors, but is also a function of genetic factors[76-80]. Further, studies have suggested that specific environmental factors can moderate the importance of genetic effects. Genes implicated in stress response are especially strong candidates for observing gene-environment interaction. For example, Covault and colleagues[81] found that among college students being homozygous for the 5-HTTLPR short-allele was associated with an increased risk for drinking outcomes (including drinking frequency

and drinking intentions) if they had experienced multiple negative events in the past year relative to their counterparts who had low (or no) exposure to negative life events. The drinking of students homozygous for the long allele did not differ as a function of negative life events. These results parallel other findings of the interaction between the 5-HTTLPR short allele and childhood maltreatment exposure on use of alcohol in children[81]. Similarly, gene-environment interactions have also been observed with a variant of the gene for the dopamine type 2 receptor (*DRD2* Taq1 polymorphism). Madrid and colleagues[83] found that variability in stress exposure interacted with the *DRD2* Taq1 polymorphism in predicting risk for alcoholism, such that carriers of the A1 allele were at an increased risk for alcoholism when exposed to higher levels of stressors in comparison to lower levels of stressors. These results parallel findings by Bau and colleagues[84], such that *DRD2* Taq1 A1 allele interacted with measures of stress to predict severity of alcoholism.

Another effort to extend the genotype-environment interaction literature included examining the role of *CRHR1*, which codes for the corticotropin releasing hormone receptor in the pituitary gland. Interest in *CRHR1* as a candidate gene for the interaction between environmental stress and alcohol use resulted from animal studies[85]. The gene x environment interaction was also tested in a sample of 15 year olds of predominantly European descent, selected from the Mannheim Study of Children at Risk. These results indicated that variation in the *CRHR1* gene and the greater number of negative life events during the previous 3 years was significantly associated to increasing rates of lifetime heavy alcohol use and levels of excessive use per occasion[86]. Another candidate gene includes *PER2*, which codes for the period circadian protein homolog 2 protein in humans. *PER2* is a circadian clock gene, which influences the adaptation of an organism to its internal and external environment through governing circadian

rhythms, which in turn has been found to be influenced by heavy alcohol use. Recent findings from the Mannheim study of Children at Risk indicated a protective effect of the minor allele of *PER2* on the susceptibility to alcohol use in young adults exposed to a higher number of stressful life events during the previous three years[87].

In this study, we examined the genes *NR3C1* and *CRHBP* based on their potential relevance for stress response. *NR3C1* codes for the glucocorticoid receptor that, when bound to glucocorticoids, acts as a transcription factor mediating the adaptation to environmental challenges and stress[88]. A number of functional polymorphisms have been identified that impact sensitivity or resistance to glucocorticoids[89,90], which is released following stress- (including alcohol-) induced activation of the hypothalamic-pituitary-adrenocortical (HPA) axis. The known functional role of *NR3C1* variants in regulating the body's response to environmental challenges and psychosocial stress point to *NR3C1* as a high priority candidate gene for this study. Laboratory studies have suggested a role for *NR3C1* in self-administration of drugs of abuse in animal models[88].

CRHBP codes for the corticotropin-releasing hormone binding protein (CRH-BP), which regulates the availability of CRH to act at its receptors and inhibits CRH activation of the HPA axis. Activity of the CRH-HPA system has long been known to shape effects of environmental impacts during development (including SLE and substance use) on responses to later life stressors and impact risk for psychiatric disorders[91]. In a study by Ray[92], it was found that a genetic variant in *CRHBP* was associated with variations in alcohol craving. Specifically, the T-allele homozygotes at rs10055255 (located within intron 6 of the *CRHBP* gene) reported greater alcohol craving during stress-induced conditions but not in the neutral conditions, greater negative moods following stress imagery but not after the neutral imagery, as well as greater

stress-induced tension, compared to A-allele carriers. In addition, *CRHBP* variants have been associated with alcohol dependence occurring concurrently with two highly comorbid conditions that themselves are known to be linked with high stress exposure—anxiety[34] and depressive symptoms[92].

Objective of the Present Study

With the increasing documentation of the influence of specific environmental factors in moderating the importance of genetic effects, it is important to identify specific genes and environments that act together. The present study examined the associations among *NR3C1* and *CRHBP* genotypes and adolescent alcohol use/misuse in African American youth living in high poverty neighborhoods as moderated by SLE. Consistent with the theoretical mechanism outlined by Shanahan and Hofer[95], we expected genetic variance associated with alcohol use/misuse would increase under conditions of higher levels of SLE[95]. Specifically, we predicted genetic effects would be most pronounced under conditions of high SLE and attenuated under conditions of low SLE.

Methods

Sample

The sample included individuals from the Genes, Environment, and Neighborhood Initiative (GENI). This group of individuals (ages 13-18) includes a subset of participants from the Mobile Youth Survey (MYS) sample from whom DNA was collected, as well as more extensive phenotypic data[96,97]. The MYS is a community-based, multiple cohort longitudinal study of adolescents who live in impoverished neighborhoods in Mobile, Alabama. The study began in 1998 with the goal of studying the etiology of risk behaviors among adolescents living in extreme poverty and how factors (such as family, school, and neighborhood) affect risk

behaviors. The GENI study was developed with the primary aim of understanding gene-environment interplay for these risk behaviors.

Participation in GENI involved an extensive interview, using an audio computerized self-administered interview (ACASI) approach, for all eligible adolescents from GENI families and their caregiver. The interviews involved questions related to the primary outcomes of interest including sexual risk taking, substance use and externalizing problems as well as on exposure to stressors, neighborhood conditions, and other potential risk or protective factors. Additionally, select candidate genes were genotyped based on the existing literature, which connects these genes to the risk behaviors of interest to GENI. Consequently, GENI aimed to investigate the gene-environment interplay in relation to a variety of risk behaviors among African American youth in urban, high poverty neighborhoods. The Institutional Review Boards at Northwestern University, Virginia Commonwealth University, the University of Illinois at Chicago, and the University of Alabama approved procedures for this study.

The total GENI sample included 592 participants; however, there was a small group of individuals from whom responses were missing on key analytic variables. Therefore, this analytic sample represents 541 individuals (mean age [SD] = 15.89 [1.43]; 51.6% female).

Measures

Alcohol Use. Data on alcohol use/misuse were collected from the AIDS Risk Behavior Assessment (ARBA) scale[98-103]. A principal component analysis was conducted using three items based on frequency, quantity, and binge drinking: "In the last 12 months, how many days did you drink alcohol?", "Think of all the times you have had a drink during the last 12 months, how many drinks did you usually have each time?", and "Over the last 12 months, on how many days did you drink 5+ drinks in a row, within a couple of hours?". Standard definitions of "drinks"

were provided. Factor scores were calculated on all participants in the sample, including those who did not endorse ever drinking. In this manner, initiation and use dimensions are collapsed together, which is more appropriate among a sample of this age group than is the case with a sample of adults.

Stressful Life Events. The Exposure to Stressors scale measures total amount of exposure to life stressors, as well as frequency of exposure to these stressors[104,105]. This is a 16-item scale assessing items within three major categories: life transitions, circumscribed events, and exposure to violence during the last 12 months. Response options included "yes" (coded as 1) or "no" (coded as 0). Frequency of exposure to these stressors was assessed based on the event occurring "once", "twice", or "three or more times". For the purposes of the present study, we used a sum score of SLE. Previous studies have found that the joint effect of exposure to multiple adverse events is stronger than the effect of a single adverse event. For example, Pilowsky and colleagues[70] found that individuals who experienced two or more adverse childhood events are at increased risk for lifetime alcohol dependence.

Genotyping

DNA was obtained via saliva samples using Oragene collection kits under the supervision of a specially trained interviewer. Saliva samples were labeled anonymously and sent to the Virginia Institute for Psychiatric and Behavioral Genetics (Richmond, Virginia), where DNA extraction and genotyping occurred. In total, DNA samples have been obtained from 579 individuals, representing 98.3% of the total GENI sample.

A total of 18 SNPs were genotyped across the two genes, all of which were based on HapMap data from the Nigerian Yoruba population, in order to capture the genetic variability in individuals of African descent. All SNPs were in Hardy-Weinberg equilibrium. Genotyping

was conducted using fluorescence polarization detection of template-directed dye-terminator incorporation (FP-TDI) with appropriate AcycloPrime SNP detection kit for specific polymorphisms (PerkinElmer, Boston) and an automated allele-scoring platform[47].

The genotyping success rate for *NR3C1* and *CRHBP* within this sample was > 98% for all variants. Haploview[107] was used to estimate linkage disequilibrium (r^2) across the full set of genotyped SNPs (summarized in Figure 1). A multiple testing correction across the SNPs was performed using the web-based software SNPSpD[108], which takes into account the number of SNPs genotyped and the linkage disequilibrium (LD) structure between them. Based on this test, we used adjusted significance values of $p = 0.006$ for *NR3C1* and $p = 0.007$ for *CRHBP* as evidence for association and interaction.

Statistical Analyses

For these analyses, each of the SNPs was coded 0, 1, or 2, reflecting an additive genetic model. This coding is in reference to the number of copies of the minor allele (Table 1). We used bivariate correlations to assess the association between SLE and alcohol use/misuse. We then used linear regression models in SPSS Statistics (Version 20.0) to assess the additive and interactive effects for genotype and SLE in predicting alcohol use/misuse. The interaction was modeled by creating cross product terms between each SNP and total SLE, (centered on its mean to aid in interpretation). The covariates of child age and sex were accounted for in calculating the main effect. Covariates accounted for in calculating the interaction effects were sex and child age, the gene-specific SNP, and total SLE. Regression models were conducted separately for each SNP.

In addition, simulations from recent research[109] demonstrate that using a cross product interaction term with a 3-level genotype can lead to spurious results under some conditions and

may not accurately capture the nature of the interaction. It has been suggested that a reparameterization of the regression equation with additional degrees of freedom is a better way to represent the nature of the interaction effects, rather than the single cross-product term that is more commonly used with a three-category coding of the genotype. Therefore, we also fit an extended parameterization of the interaction model involving greater degrees of freedom, for SNPs that yielded interactions with the single cross-product term, to determine whether the predicted interaction lines accurately represented the shape of the interaction in the data.

Results

Descriptive Statistics

The number of participants who reported ever use of alcohol was 168 (31.1%). Of those who reported ever using alcohol, 59 participants (35.1%) reported drinking alcohol 1-2 days over the past 12 months, and 51 participants (30.3%) reported drinking alcohol three or more days in the past 12 months. An average of 2.40 drinks (SD=1.39) was consumed per drinking occasion. For those subjects who reported drinking in the past year, the number of participants who reported engaging in binge drinking (having 5 or more drinks in a row within a couple of hours) on one to two days over the last 12 months was 25 (22.7%), whereas 19 participants (17.1%) engaged in binge drinking on three or more days over the past 12 months. Overall, participants reported exposure to an average of 3.85 SLE (SD=3.08). There was a modest correlation between SLE and alcohol use/misuse ($r = 0.19$, $p \leq 0.01$). An alcohol factor score was created such that higher scores indicated increased drinking frequency, quantity, and heaviness (range = -.41-6.34, mean [SD] = 0 [1]). The factor score explained 76% of the variance in alcohol use/misuse outcome. Factor loadings based on principal component analysis for the 3 items were 0.88, 0.90, and 0.83, respectively.

Regression Analyses

Results of moderated multiple regressions of all SNPs and total SLE predicting alcohol use/misuse, are shown in Table 1. The main effect SLE on predicting alcohol use and misuse was significant (R^2=0.08, *B*=0.06, $p \leq 0.01$). There were no significant main effects of any of the SNPs on the outcome, based on the Nyholt correction.

In *CRHBP*, the interaction between SLE and rs1715749 was significant in predicting alcohol use/misuse after applying the Nyholt correction (R^2=0.09, *B*=-0.06, $p \leq 0.006$) (Table 1). Post-hoc analyses including a test of simple slopes for the significant interaction between SLE and rs1715749 indicated that the association between SLE and alcohol use/ misuse was significant for the C/C genotypic group (*B*=0.11, $p \leq 0.001$) and C/T genotypic group (*B*=0.06, $p \leq 0.002$) but not for the T/T genotypic group (*B*=0.001, p =0.97). A regression plot to illustrate the interaction effect is shown in Figure 2A. After correcting for multiple testing, there were no significant interactions between SLE and *NR3C1* (Table 1).

Figure 2B is a plot of the regression lines for each genotype from the reparameterization of the regression equation for rs1715749 in *CRHBP* that yielded a significant interaction effect with the cross-product term. In a comparison of the regression plots representing the predicted values from the standard regression equation using the single cross-product term (Figure 2A) to the regression plots representing the extended parameterization (Figure 2B), we can see that using the single cross product term relatively accurately captures the nature of the interaction in relation to the ordering of the genotypic categories and slope of the predicted regression lines for each genotype.

In a set of supplementary analyses, linear regression models were run among only those subjects who reported any drinking over the past year (n=168). The main effect of SLE on

alcohol use and misuse remained significant (R^2=0.08, *B*=0.06, $p \leq 0.05$) (Table 2). There were no significant main effects of any of the SNPs on the outcomes, based on the Nyholt correction. Though not significant after applying the Nyholt adjusted p-value, the interaction between SLE and rs1715749 remained nominally significant in predicting alcohol use/misuse (R^2=0.12, *B*=-0.09, $p \leq 0.05$) (Table 2). Thus, even though the moderation effect is attenuated (as is expected given the reduced sample size), the effects observed in the full sample remain consistent when non-drinkers are excluded. This suggests that our results are not entirely driven by the distribution of the alcohol factor.

Discussion

The present study examined the associations between genotypes in *CRHBP* and *NR3C1* and adolescent alcohol use/misuse in a sample of African Americans adolescents living in high poverty neighborhoods. We found a significant main effect of SLE on alcohol use/misuse such that higher levels of SLE were associated with higher levels of alcohol use/misuse. There was no evidence for main effects for either of the genes studied here after correcting for multiple testing. We found a significant interaction between rs1715749 in *CRHBP* and SLE in predicting alcohol use/misuse. Individuals having two copies of the minor allele may be more resilient to environments of high SLE than those individuals with zero or one copies of minor allele in the same environment of high SLE. Specifically, when we tested regions of significance for each genotype group contrast (i.e. 0 versus 1 copy of the minor allele), we found that individuals with zero copies of the minor allele had significantly fewer alcohol use problems as a function of low SLE than individuals with two copies of the minor allele in the same conditions of SLE. These findings support the concept of differential susceptibility[110], which suggests that individuals who are highly susceptible to their environments fare more poorly in negative environments (i.e. high

SLE) but also fare much better in positive environments (i.e. low SLE) compared to individuals who are less susceptible to their environments. After taking into account multiple testing, there were no significant interactions between any of the SNPs from *NR3C1* and SLE in predicting alcohol use/misuse.

The SNP in *CRHBP* with which we find evidence of gene-environment interaction is novel in that this SNP has not been examined in relation to alcohol related outcomes in prior research. The SNP is located in the promoter region of the gene and lies within a distinct haplotype block from SNPs in the 3' UTR region for which previous studies have evidenced associations with other psychiatric disorders[92,94]. Among existing studies examining variants of *CRHBP*, Ray and colleagues[33] found that in a sample of non-treatment seeking heavy drinkers, homozygotes for the T-allele of rs10055255 reported higher stress-induced craving for alcohol. However, this SNP is not in LD with rs1715749[107]. Evidence for the role of *CRHBP* in stress and substance use also comes from animal models in which *CRHBP* in rat brain has been shown to modulate effects of corticotropin releasing hormone on stress-induced relapse to drug abuse[111]. Other studies have shown that *CRHBP* variants are potentially relevant for adolescent alcohol use/misuse[92,93]; however, we did not find any main effects in our African American population.

Limitations

The results from this study should be considered in the context of several limitations. First, the effects discussed in this study were identified within a sample of African American youth living in high poverty neighborhoods; however, it is possible that the same findings might not be found within a different ethnic or socioeconomic sample. Drinking behaviors differ significantly across varying ethnic groups[66]. Adolescents of European ancestry begin drinking at an earlier age, and drink greater quantities with more frequency than adolescents of African

ancestry[112-114]. It is also noteworthy that endorsement of alcohol use is much lower among African American adolescents in comparison to their white counterparts, which has been suggested to be, in part, a result of greater levels of disapproval of substance use among African American populations[115]. Rates of alcohol use in this sample were comparable to the prevalence found nationally[116]. Specifically, ever drinking alcohol was endorsed by 30.7% of the participants in this sample, whereas the national prevalence for alcohol use in African American youth is 33.4%[116]. Rates of alcohol use in this sample were also comparable to other samples of African American youth living in public housing such that the prevalence of lifetime alcohol use among youth, ages 11 to 21, was reported as 35.3%[74]. Accordingly, genetic effects associated with alcohol use may not be as easily detected in African American samples because of low endorsement. For example, in the present study, supplementary analyses were conducted among the subsample who endorsed any drinking over the past year. However, as expected, due to the reduced sample size owing to the low endorsement of alcohol use, the interaction between SLE and rs1715749 was nominally significant ($p < 0.05$) in predicting alcohol use/misuse, but did not remain significant after applying the Nyholt correction. Nonetheless, it is important to note that the effect size of the interaction between SLE and rs1715749 in predicting alcohol use/misuse is larger than the resulting effect size seen when analyses included the whole sample and the direction of the effect of the interaction between SLE and rs1715749 in predicting alcohol use/misuse remains the same as compared to the resulting interaction effect seen when analyses included the whole sample. Thus, future analyses run on samples of only those who endorse alcohol use in the past year would require larger sample sizes to address the limitation of low endorsement of alcohol use that is characteristic of African American adolescents.

Second, other issues affecting lack of consistency of effects seen across populations are differences in minor allele frequency[117]. Differing allele frequencies between populations can affect the ability to detect effects from one population to the next; therefore, studies of individuals of African descent may not detect the same signals (as seen in European Americans). Further, this may suggest that causal variants differ between populations of different ethnicities, thus highlighting the need to extend genotyping efforts to further elucidate potentially important effects that vary between populations. This underscores the importance of conducting genetically informative studies in African Americans and other minority populations.

Lastly, the present study examines only two genes involved in stress response. However, future investigations should include additional candidate genes that go beyond the commonly selected genes (e.g., 5-*HTTLPR* and *DRD2*). Other stress genes, such as *CRHR1* and *PER2*, have also been shown to interact with SLE and alcohol related outcomes, such as adolescent heavy alcohol use[86] and more drinks per occasion[87] in individuals of predominantly European descent. However, though not genotyped within this sample, the role of such additional stress-related genes should be included and extended to samples of African American youth in future analyses.

Conclusion

In conclusion, these results extend the growing literature on the role for *CRHBP* in substance use and other stress-related disorders. Notably, these findings suggest that *CRHBP* rs1715749 may contribute toward the risk of alcohol use/misuse in African Americans, such that the effect of this gene on alcohol use/misuse can vary as a function of exposure to SLE. Identifying how specific environmental variables interact with genetic variants to influence substance use behaviors is necessary to develop a better understanding of the etiology of complex behaviors. Identifying environmental risk factors to developing problematic substance

use outcomes also promotes a better understanding of those social contexts under which genetic predispositions are expressed. In addition, such interactions also aid in developing more targeted prevention and treatment programs for adolescent individuals at risk for developing substance use problems and/or for adolescents exposed to a wide variety of stressors. Initial analyses of genotype-intervention interaction studies suggest that children who are most at risk for substance use and externalizing outcomes may also be those who are most likely to benefit from intervention[118-120].

CHAPTER 3: Polygenic risk scores predict externalizing behaviors and interact with stressful life events in a sample of high-risk African American youth

Abstract

Background: Literature has provided increasing evidence making apparent that alcohol use and externalizing disorders are influenced by overlapping genetic factors. Environmental exposures, such as exposure to stressors, have also been shown to moderate the impact of genetic risks for alcohol use disorders and externalizing psychopathology. We aim to assess whether polygenic risk for alcohol dependence - as derived from GWAS estimates in one African American subsample - predicts alcohol use/misuse and broader externalizing behaviors in a sample of African American adolescents. Further, we test the interaction between polygene score and alcohol use/misuse as moderated by stressful life events. Lastly, we aim to conduct secondary analyses using the polygenic risk scores to examine the extent to which the polygenic scores from a racially/ethnically mismatched sample predict alcohol use/misuse and externalizing disorders in our African American adolescent subsample.
Methods: Using the single nucleotide polymorphisms weights derived from GWAS results from the largest sample for alcohol dependence to date, we create polygenic risk scores in an independent validation sample, the Genes, Environment, Neighborhood Initiative (GENI), This is a subset of 542 individuals, ages 12-18, from whom DNA and more extensive phenotypic data were collected, from the Mobile Youth Survey, a community-based longitudinal study of adolescents in impoverished neighborhoods of Mobile, Alabama.
Results: Stressful life events significantly predicted alcohol use and misuse such that higher levels of SLE were associated with higher levels of alcohol use/misuse. Polygenic scores as derived from the African American sample significantly predicted alcohol use/misuse across all SNP p-value thresholds, and predicted the externalizing behavior composite at p-value thresholds of 0.005 and 0.01, after accounting for gender and age. Further findings also included a significant interaction between polygenic risk and SLE in predicting alcohol use/misuse and externalizing behaviors. Secondary analyses of polygenic scores derived from the European American sample also predicted alcohol use/misuse at three of the selected p-value thresholds, but did not achieve significance when predicting the externalizing behaviors composite.
Conclusions: These findings extend the limited genetic research on GxE effects in African Americans. This study further illustrate the importance of the polygenic approach in understanding the etiology and genetic architecture of general externalizing behaviors including alcohol use and its related outcomes among populations of African ancestry.

Introduction

Traditionally, individuals of African descent have been underrepresented in research aimed at understanding genetic influences on complex psychiatric outcomes. This is true across the various types of study designs that are commonly used for genetic research, including twin studies[121-123], candidate gene studies[124-126], linkage studies[127-129], and most recently, genome-wide association studies (GWAS)[130-132].

There are likely many factors contributing to this under-representation of individuals of African descent in genetic research. It is more challenging to recruit African Americans to participate in genetic research: a review of multiple studies that measured consent rates for genetic research participation found that African Americans had significantly lower levels of consent[133]. This likely stems from a number of factors including cultural or racial stigma, mistrust of scientific research, concerns about discrimination and confidentiality, perceptions of being used by the researching scientist, and lack of interest or perception that there is no perceived benefit from participation in such a study[133-137]. Another factor contributing to the lack of diversity in genetic studies is that the majority of GWAS conducted in the United States have used existing cohorts as the initial sample population. These cohorts were mostly created before there was an increased effort to include minority populations in biomedical research and therefore, only included persons of a single ancestral background[138]. Finally, participation in genetic studies of complex behavioral disorders such as alcohol use and related outcomes may be further hindered because of increased stigma surrounding psychiatric disorders in individuals of African descent. For example, Schnittker and colleagues[139] reported that individuals of African descent were less likely to accept genetic or familial influences as causative factors of psychiatric

disorders. Thus, likely for this myriad of reasons, to date the majority of studies on the genetics of alcohol use and related outcomes have been studies of individuals of European descent.

This is problematic because genetic studies in other ancestral populations may not directly translate to populations of African descent, and/or may miss genetic variants that are important in populations of African descent. Studies of human evolution and mitochondrial DNA point to African populations as being one of the oldest populations of humans. Thus, the genetic architecture of African populations is characterized by greater levels of genetic diversity than European populations, resulting from greater exposure to evolutionary forces experienced by the African population, such as genetic drift, mutations, etc[140, 141]. African populations are characterized by differences in linkage disequilibrium (LD) and allele frequencies in comparison to other non-African populations[47,51,140]. These factors contribute to the challenge of translating genetic findings from populations of African descent, and underscore the need to conduct genetic studies in populations of African descent.

Recently, Gelernter and colleagues[142] conducted a GWAS with 16,087 African American and European American subjects - the largest sample for alcohol dependence GWAS to date, and the first to report genome wide significance in a GWAS of individuals of African ancestry. In the study, genes previously significantly associated with alcohol dependence such as *ALDH1B* and *ALDH2* were successfully replicated in the African American sample. However, a novel significant risk loci was also identified, mapping to the *ADH* gene cluster on chromosome 4 and extending centromerically to a region that includes non-*ADH* genes. Findings also suggested biological convergence across populations, evident by genome wide significant findings, in both the African American and European American samples, for different SNPs but at the same locus. For example, significant associations were identified on chromosome 2 at rs1437396, between

MTIF2 and *CCDC88A*, yielding support from both the African American subsample and the subsample of European descent. However, it is important to note that most associations were reported to be population-specific, such as the identification of SNP rs1789882 as the first genome wide significant finding reported in African Americans for alcohol dependence, which was not genome wide significant in the European American populations[142].

In the current study, we aim to assess whether polygenic risk scores, as derived from GWAS estimates in Gelernter's[142] African American subsample, predict alcohol use/misuse and broader externalizing behaviors in an adolescent sample of African Americans. We studied broader externalizing behavior in addition to alcohol outcomes in our adolescent sample, based previous evidence showing that broader externalizing behaviors, such as delinquent behavior and aggressive behavior, serve as precursors for the development of alcohol use and its related outcomes[143, 144]. Additionally, twin studies show that alcohol use and externalizing disorders are influenced by overlapping genetic factors and indicate that genes that influence adult alcohol dependence may manifest earlier in development as behavior problems[76,145-147].

Further, environmental exposures have also been shown to moderate the impact of genetic risks for alcohol use disorders and externalizing psychopathology. One such environment that has been commonly identified as an important risk factor for alcohol use and externalizing behaviors is exposure to stressors[72,122]. Specifically, research indicates that the experience of major stressful life events induces alcohol use and its related problems among individuals who are at an increased genetic risk[122,148]. The MYS sample was ascertained from impoverished neighborhoods in order to study gene-environment interplay under high-risk environmental conditions. African Americans have increased exposure to a variety of stressors, including high rates of unemployment, poverty, violent crime, exposure to segregation, incarceration, health-

care disparities, and homicide[149-151]. Therefore, in the current study, we also test for interaction between the polygene score and stressful life events in predicting alcohol use and externalizing behavior in our African American adolescent sample, using polygenic risk scores derived from the African American sample of the Gelernter study[142]. We conduct secondary analyses using the polygenic risk scores derived from the European American subsample in the Gelernter study to examine the extent to which the polygenic scores from a racially/ethnically mismatched sample predict alcohol use/misuse and externalizing disorders in our African American adolescent sample.

Methods

Sample

The sample included individuals from the Genes, Environment, and Neighborhood Initiative (GENI). This group of individuals (ages 13-18) includes a subset of participants from the Mobile Youth Survey (MYS) sample from whom DNA, and more extensive phenotypic data were collected[96,97]. The MYS is a community-based, multiple cohort longitudinal study of adolescents who live in impoverished neighborhoods in Mobile, Alabama. The study began in 1998 with the goal of studying the etiology of risk behaviors among adolescents living in extreme poverty and how factors (such as family, school, and neighborhood) affect risk behaviors. The GENI study was developed with the primary aim of understanding gene-environment interplay for these risk behaviors.

Participation in GENI involved an extensive interview, using an audio computerized self-administered interview (ACASI) approach, for all eligible adolescents from GENI families and their caregiver. The interviews involved questions related to the primary outcomes of interest including sexual risk taking, substance use and externalizing problems as well as exposure to

stressors, neighborhood conditions, and other potential risk or protective factors. Consequently, GENI aimed to investigate the gene-environment interplay in relation to a variety of risk behaviors among African American youth in urban, high poverty neighborhoods. The Institutional Review Boards at Northwestern University, Virginia Commonwealth University, the University of Illinois at Chicago, and the University of Alabama approved procedures for this study.

The present analytic sample represents 562 African American youth (mean age [SD] = 15.92 [1.43], age range = 13-19; 50.71% female), which includes 306 families with one individual, 99 families with 2 siblings, 18 families with 3 trios, and 1 family with 4 individuals.

Measures

Alcohol Use. Data on alcohol use/misuse were collected from the AIDS Risk Behavior Assessment (ARBA) scale[98-103]. A principal component analysis was conducted, with SPSS Statistics Version 22, using three items based on frequency, quantity, and binge drinking: "In the last 12 months, how many days did you drink alcohol?", "Think of all the times you have had a drink during the last 12 months, how many drinks did you usually have each time?", and "Over the last 12 months, on how many days did you drink 5+ drinks in a row, within a couple of hours?". Standard definitions of "drinks" were provided. In addition, heavy episodic/binge drinking puts individuals at risk for future alcohol-related problems, and is commonly considered an indicator of alcohol misuse[152]. Thus, we refer to the outcome as a measure of alcohol use/misuse. Factor scores were calculated on all participants in the sample, including those who did not endorse ever drinking. In this manner, initiation and use dimensions are collapsed together, which is more appropriate among a sample of this age group than is the case with a sample of adults. The three items loaded strongly onto a single factor, explaining 76% of the

variance in the alcohol use/misuse outcome. Factor loadings based on principal component analysis for the 3 items were 0.88, 0.90, and 0.83, respectively. An alcohol factor score was created such that higher scores indicated increased drinking frequency, quantity, and heaviness (range = -.41-6.34, mean [SD] = 0 [1]).

Externalizing Behaviors. Externalizing behaviors were assessed using the Youth Self Report, derived from the Child Behavior Checklist (CBCL)[153]. The Youth Self Report includes 112-items to assess eight-subscale items. In the current study, externalizing behaviors were assessed using the total score of the following subscales: rule-breaking/delinquent behavior and aggressive behavior. Examples of items assessing rule-breaking behavior include "I lie or cheat" and "I set fires". Examples of items assessing aggressive behavior include "I am mean to others" and "I destroy things belonging to others." Response options included "Not true" (coded as 0), "Somewhat or Sometimes true" (coded as 1), and "Very True or Often true" (coded as 2).

Stressful Life Events. The Exposure to Stressors scale measures total amount of exposure to life stressors, as well as frequency of exposure to these stressors[104, 105]. This is a 16-item scale assessing items within three major categories: life transitions, circumscribed events, and exposure to violence during the last 12 months. Response options included "yes" (coded as 1) or "no" (coded as 0). Frequency of exposure to these stressors was assessed based on the event occurring "once", "twice", or "three or more times". For the purposes of the present study, we used a sum score of SLE. Previous studies have found that the joint effect of exposure to multiple adverse events is stronger than the effect of a single adverse event. For example, Pilowsky and colleagues[70] found that individuals who experienced two or more adverse childhood events are at increased risk for lifetime alcohol dependence.

Genotyping and quality control

DNA was obtained via saliva samples using Oragene collection kits under the supervision of a specially trained interviewer. Saliva were labeled anonymously and sent to the Rutgers University Cell & DNA Repository (Piscataway, New Jersey) to be genotyped on the Affymetrix Biobank Version 2 Array, which contains both rare variation (exome and structural variation), as well as an imputation GWAS grid.

We performed all standard quality control checks[154-156]. For example, such checks included standardizing the naming convention of all samples, including control samples, and removal of all poorly performing SNPs which failed quality control. In addition, gender concordancy checks were run to assess discrepancies between phenotypically reported genders versus gender calls based on genotypic information. Inability to accurately obtain a gender call resulted in exclusion of the sample (n=27). Furthermore, any unintentional sample replicates were also removed. For each duplicate pair, the sample with the higher call rate was carried forward.

Identity by descent (IBD) estimation. We ran IBD estimates (the degree of recent shared ancestry for a pair of individuals) incorporating the ancestral proportions and a pruned subset of SNPs (~100,000 SNPs) based on their linkage disequilibrium (LD, r^2) to generate pi-hat estimates for each individual using the software REAP[157]. LD pruning retains those markers that are independent from each other. PI-hat estimates are calculated using the equation P(IBD=2)+0.5*P(IBD=1), where P = proportion, P(IBD=2) = the probability that two alleles are identical-by-descent, and P(IBD=1) = the probability that 1 allele is identical-by-descent.

The pairwise PI-hat equation is used to gauge the level of relatedness between two samples by determining the probable number of shared alleles at any given marker. In this manner, we can verify and/or identify the presence of twins or unintentional duplicates.

Furthermore, based on theoretical values as provided by the literature[157], the expectation is that PI-hat=1 for duplicates or monozygotic twins and PI-hat=0.5 for biological siblings. However, due to genotyping error, LD and population structure, there can be variation in the theoretical values[154]. We can also check whether an individual's genotype matches their reported relationship in the pedigree file using PI-hat estimation.

Imputation. The reference panel used for imputation was the March, 2012 (Phase 2) 1000 Genomes data. Imputation was carried out using the combination of SHAPEIT[158,159] and IMPUTE2 v2[160] resulting in a total of 37,466,192 SNPs. After quality control filtering, including minor allele frequency > 0.01 and Hardy-Weinberg disequilibrium $P>1x10^{-6}$, 4,099,400 SNPs remained. Additional filtering including call rate > 95% resulted in a total of 562 individuals.

Discovery Sample GWAS

The discovery sample, from which the polygenic risk scores were computed, consisted of the African American subset of the GWAS for alcohol dependence conducted by Gelernter and colleagues[142]. The African American subset consisted of 3318 AA subjects from Gelernter and colleagues[142] discovery sample, in addition to an identically ascertained replication sample consisting of 803 AA and 1311 AA subjects from the Study of Addiction: Genetics and Environment (SAGE). To conduct secondary analyses in which we aim to compare the ability of polygenic scores from a racially/ethnically mismatched sample to predict alcohol use/misuse and externalizing disorders in our AA adolescent sample, we also used the European American subset of the GWAS for alcohol dependence conducted by Gelernter and colleagues[142]. This subsample included 2379 EA subjects from the discovery sample, 1746 EA subjects from the identically ascertained replication sample, 3784 EA alcohol dependent and control subjects from a previously collected German sample, and 2750 EA subjects from the SAGE sample, which is

available to researchers through database of Genotypes and Phenotypes (dbGaP #phs000092.v1.p1) application. The discovery sample ascertained by Gelernter was genotyped on the Illumina HumanOmni1-Quad v1.0 microarray. Genotypes were called using GenomeStudio software V2011.1 and genotyping module V1.8.4 (Illumina, San Diego, CA, USA). SAGE samples were genotyped on the Illumina Human 1M array. Additional SNPs in the German sample were genotyped individually using Taqman[161]. From the African American and European American GWAS results, a total of 4,525,278SNPs were available for the computation of polygenic risk scores in the GENI sample, after filtering to include p-value<0.5 and MAF>0.02.

Statistical Analyses

Computation of polygenic scores. We used GWAS estimates for each independent SNP from all African American individuals in the discovery sample[142] to calculate polygenic scores for each subject within the GENI sample using the --score procedure in PLINK[156], with A/T and G/C SNPs excluded and pruned for LD at the level of $R^2 > 0.5$. Specifically, we compute a linear function of the number of score alleles that an individual possesses, weighted by the product of the sign of the SNP effect from logistic regression, taken from the discovery GWAS, and the negative logarithm (base 10) of the associated GWAS p-value. Because there is no set criteria in creating polygene scores that are maximally informative[162], we conducted analyses using a series of p-value thresholds ranging from 0.005 to 0.50 to evaluate which p-value threshold maximizes the total variance accounted for in alcohol use/misuse and the externalizing behavior composite score in the adolescent replication sample. All p-value thresholds were used to create polygene scores in each subject of the GENI sample. For our secondary analyses, to address whether GWAS estimates from a European American sample will also predict alcohol use/misuse and

externalizing behaviors in our African American sample, GWAS estimates for each independent SNP from all European American individuals from the discovery sample[142] were also used in the same manner, to compute polygenic scores for individuals within the GENI sample.

Regression. Regression analyses were conducted using the SURVEYREG procedure, which accounts for family relatedness, in SAS (Version 9.3) to assess, separately, whether the polygenic score at varying p-value thresholds predicted the alcohol use/misuse factor score and the externalizing behavior composite. The gender and child age variables were used as covariates in predicting alcohol use/misuse and the externalizing behaviors composite. We also ran moderated multiple regressions to test whether the interaction between stressful life events and polygenic risk predicted alcohol use/misuse. Covariates accounted for in calculating the interaction effects were gender and child age, the polygene score specific to the varying p-value thresholds, and total SLE. In creating the polygene score and total SLE interaction, the variables were centered on the mean to aid in interpretation.

Results

The number of participants who reported ever use of alcohol was 163 (30.2%). Of those who reported ever using alcohol, 58 participants reported drinking alcohol 1-2 days over the past 12 months, and 50 participants reported drinking alcohol three or more days in the past 12 months. For those subjects who reported drinking in the past year, the number of participants who reported engaging in binge drinking (having 5 or more drinks in a row within a couple of hours) on one to two days over the last 12 months was 25, whereas 19 participants engaged in binge drinking on three or more days over the past 12 months. Table 3 contains means and standard deviations for all other variables under study.

Table 4 contains the results from the regression analyses testing the main effect of the polygenic risk scores and interaction with SLEs for alcohol use/misuse. The main effect of polygenic risk predicting alcohol use/misuse was significant across all p-value thresholds ($p \leq 0.01$), after controlling for gender and age. The main effect of SLE on predicting alcohol use and misuse was significant (R^2=0.04, b= 0.06, p<0.001). Moderated multiple regression analyses indicated that SLE moderated polygenic risk to predict alcohol use/misuse after controlling for gender, age, and the main effects of the polygenic score and SLE, at p-value thresholds of 0.005, 0.01, 0.05, and 0.10. The interaction between SLE and polygenic score using the $p \leq 0.01$ threshold accounted for the greatest proportion of variance (0.47%, range = 0.14-0.47%) in predicting alcohol use/misuse. There was a general trend of an increase in percent variance accounted for with decreasing p-value thresholds, suggesting a further reduction in noise in the signal (Table 4). Regression plots to illustrate the significant interaction effects, where low and high values (±1 *SD* of the mean) for the polygenic scores (i.e. low and high genetic risk) and SLE are shown in Figure 3.

Next we examined whether polygenic risk scores also predicted other externalizing behavior, based on the literature suggesting shared genetic effects between alcohol outcomes and antisocial behavior[122,145-147]. The main effect of SLE on predicting externalizing behaviors was significant (R^2 = 0.13, b=1.01, p<.0001). The main effect of polygenic risk predicting externalizing behaviors was significant at p-value thresholds of 0.005 and 0.01, after controlling for gender and age ($p \leq 0.05$). Separate moderated multiple regression analyses indicated that SLE moderated polygenic risk to predict externalizing behaviors after controlling for gender, age, and the main effects of the polygenic score and SLE, across all p-value thresholds. The interaction between SLE and polygenic score using the p<0.05 threshold accounted for the

greatest proportion of variance (1.24%, range = 0.63-1.24%) in predicting the externalizing behavior composite (Table 5). Under conditions of high SLE, individuals with high polygenic risk have higher levels of externalizing behaviors than when exposed to conditions of low SLE. However, under conditions of low SLE, individuals with high polygenic risk fare better than individuals with low polygenic risk.

Secondary analyses including polygene scores from EA sample

In a set of secondary analyses, polygenic risk scores derived from Gelernter's[142] EA sample were calculated to predict alcohol use/misuse and externalizing behaviors. The main effect of polygenic risk predicting alcohol use/misuse was significant at p-value thresholds of 0.01, 0.05, 0.10, after controlling for gender and age. Though there was little variability in the percent variance accounted for (range = 0.1-0.8%), a p-value threshold less than 0.05 maximized R^2 for the main effect (0.8%). The main effect of polygenic risk predicting externalizing behaviors was not significant at any p-value threshold, after controlling for gender and age (Table 6).

Discussion

The present study examined the associations between polygenic risk and alcohol use/misuse, and externalizing behaviors. In addition, we also analyzed the interaction between polygenic risk and SLE in predicting alcohol use/misuse and externalizing behaviors in a sample of African American adolescents living in high poverty neighborhoods. We found a significant main effect of SLE on alcohol use and misuse such that higher levels of SLE were associated with higher levels of alcohol use/misuse. SLE explained 4% of the variance in alcohol use/misuse outcome. Polygenic scores as derived from the African American sample significantly predicted alcohol use/misuse across all SNP p-value thresholds, and predicted the

externalizing behavior composite at p-value thresholds of 0.005 and 0.01, after accounting for gender and age. The magnitude of the associations between polygenic scores and alcohol use/misuse and externalizing behaviors was fairly consistent across the range of selected p-value thresholds. In predicting alcohol use/misuse, a p-value threshold less than 0.005 maximized the variance explained. Further, in predicting the externalizing behavior composite, a p-value threshold less than 0.01 maximized the variance explained; however, there was little variability in the percent variance accounted for (0.43-0.60%).

Further findings also included a significant interaction between polygenic risk and SLE in predicting alcohol use/misuse and externalizing behaviors. In predicting alcohol use/misuse, genetically at risk individuals (i.e. high polygenic risk) were more at risk for developing alcohol use/misuse than individuals of low genetic risk (i.e. low polygenic risk) under conditions of high SLE. The nature of the interaction depicted in Figure 4 suggest individuals with higher genetic risk who also experience higher SLE show elevations in externalizing behavior compared to individuals with a lower genetic risk. However, these individuals fare better (lower externalizing behavior) in positive environments (i.e. low SLE) compared to individuals who are at a lower genetic risk. Thus, a cross-over effect is observed for externalizing disorders at the low end of SLEs in a way that was not observed for alcohol misuse. This difference may reflect a floor effect, in that only 30.2% of the sample reported ever use of alcohol. However, there is greater variability associated with externalizing behaviors, (i.e. range=0-47, mean (*SD*) = 11.42 (8.50)), which may have allowed the detection of a cross-over effect that was not present with the alcohol outcome at this age. Furthermore, these findings parallel existing studies that demonstrate that in the context of environmental adversity (i.e. increased stressful life events), genetic influences become more pronounced in the development of externalizing behaviors[95,163]. Our results also

illustrate a significant main effect of SLE predicting externalizing behaviors which also parallels previous findings showing strong associations between externalizing behaviors and environmental variables such as increased stressful life events[163]. Finally, our secondary analyses of polygenic scores derived from the European American sample also predicted alcohol use/misuse at three of the selected p-value thresholds, but did not achieve significance when predicting the externalizing behaviors composite. Thus, although there was some prediction associated with the risk scores generated from the European American sample, the prediction was not nearly as robust as when using the racial/ethnically matched discovery sample.

The current interaction results, indicating genetic risk and alcohol use/misuse is greater among individuals exposed to a higher level of stressful life events, are also consistent with one of the only other polygene score analyses that involve substance use outcomes in a sample of African Americans[164]. Meyers and colleagues[164] reported polygenic risk scores significantly predicted cigarettes smoked per day, accounting for 3% of the overall variance in cigarettes smoked per day. Further findings from this group also paralleled the present findings such that a significant interaction between polygenic risk score and traumatic events was found to be predictive of smoking, in addition to the interaction between polygenic risk score and neighborhood social cohesion. Specifically, it was found that the association between genetic risk and smoking was greater among individuals who had experienced an increased number of traumatic events experienced[164]. A cross over effect was not seen when examining polygenic risk and traumatic events, but was evident in the interaction between polygenic risk and neighborhood social cohesion. Genetic risk for smoking was greater for individuals who had experienced an increased number of traumatic events; whereas the association was diminished for individuals of low genetic risk. Genetically at risk individuals indicated higher mean

cigarettes per day in neighborhoods characterized by decreased social cohesion than individuals at low genetic risk. Such findings highlight the importance of identifying environmental risk factors, such as stressful life events, and its interactions with genetic predispositions to influence substance use behaviors, specifically in African Americans, who are exposed to increased environmental stressors, including psychosocial, sociodemographic, and economic disparities and consequently are at a greater risk of developing alcohol related disorders than individuals of European ancestry[165]. Most importantly, identifying such environmental risk factors, and how they exacerbate or diminish genetic influences on alcohol use and its related outcomes hold important implications for public health interventions, as such environmental factors are potentially modifiable.

Limitations

The results from this study should be considered in the context of its limitations. First, the present study focuses on a unique sample – African American youth living in high poverty neighborhoods. Partly due to the high level of genetic diversity seen in African American populations[140, 166], as well as the increased environmental stressors to which they are exposed, the results of this study may not be generalizable to populations with different racial/ethnic or socioeconomic backgrounds. Second, this study is limited in its potential strength to detect significant signals due to the small sample size and further, low endorsement of alcohol use, such that only 30.2% of the sample endorsed ever use of alcohol. However, the rate of endorsement of alcohol use in this sample is comparable to the national prevalence for alcohol use in African American youth (33.4%)[116]. Third, as in any polygenic risk analysis, there exists the limitation that GWAS derived polygenic risk scores only account for common genetic variation, rather than rare genetic variation (allele frequency less than 1%)[167]. In addition, the polygenic approach

focuses on an aggregate of risk variants that can have predictive power for an outcome rather than individual SNPs[168]. Fourth, the above analyses were conducted before applying a corrected p-value for multiple testing. The Bonferroni corrected p-value significance threshold of 0.05/42 was $p<0.001$. After accounting for the Bonferroni correction, the main effect of polygenic risk scores, computed with SNPs at p-value thresholds of $p\leq0.05$ and $p\leq0.01$, predicting alcohol use misuse was significant. Additionally, the interaction effect between SLE and polygenic risk score predicting externalizing behaviors surpassed Bonferroni correction for the polygenic risk scores that were computed using SNPs with $p\leq 0.05$, 0.10, and 0.50. All other significant effects were nominally significant ($p\leq 0.05$), after Bonferroni correction.

In conclusion, these findings extend the limited genetic research on GxE effects in African Americans. Notably, these findings suggest the importance of the polygenic approach in understanding the etiology and genetic architecture of general externalizing behaviors including alcohol use and its related outcomes among populations of African ancestry. This is of particular importance since successful significant findings arising from GWAS and candidate-gene studies have been far more limited in populations of African ancestry than in populations of European ancestry.

Literature Cited

Literature Cited

1. Wu, L. T., Woody, G. E., Yang, C., Pan, J. J., & Blazer, D. G. (2011). Racial/ethnic variations in substance-related disorders among adolescents in the United States. *Arch Gen Psychiatry, 68*(11), 1176-1185. doi: 10.1001/archgenpsychiatry.2011.120.

2. Skidmore, J. R., Murphy, J. G., Martens, M., & Dennhardt, A. A. (2012). Alcohol-related consequences in African American and European American college students. *J Ethn Subst Abuse, 11*(2), 174-191. doi: 10.1080/15332640.2012.675248.

3. Biafora, F. & Zimmerman, R. (1998). Developmental patterns of African American adolescent drug use. In W. A. Vega & A. G. Gil (Eds.), Drug use and ethnicity in early adolescence (pp. 149-175). New York, NY: Plenum.

4. Rothman, E. F., Wise, L. A., Bernstein, E., & Bernstein, J. (2009). The timing of alcohol use and sexual initiation among a sample of Black, Hispanic, and White adolescents. *J Ethn Subst Abuse,* 8(2), 129-145. doi: 10.1080/15332640902896984.

5. Acevedo-Garcia D, Ospuk T, McArdle N, Williams DR. (2008). Toward a policy relevant analysis of geographic and racial/ethnic health disparities. *Health Affairs,* 27:321–333.

6. Smedley A & Smedley BD. (2005). Race as biology is fiction, racisms as a social problem is real: Anthropological and historical perspectives on the social construction of race. *Am Psychol*, 60(1): 16-26.

7. Clark, R., Anderson, N. B., Clark, V. R., & Williams, D. R. (1999). Racism as a stressor for African Americans: A biopsychosocial model. *American psychologist, 54*(10), 805.

8. Williams, D. R., Mohammed, S. A., Leavell, J., & Collins, C. (2010). Race, socioeconomic status, and health: complexities, ongoing challenges, and research opportunities. *Annals of the New York Academy of Sciences, 1186*(1), 69-101.

9. Zapolski, T. C., Pedersen, S. L., McCarthy, D. M., & Smith, G. T. (2014). Less drinking, yet more problems: Understanding African American drinking and related problems. *Psychological bulletin*, *140*(1), 188.

10. McCabe, S. E., Morales, M., Cranford, J. A., Delva, J., McPherson, M. D., & Boyd, C. J. (2007). Race/ethnicity and gender differences in drug use and abuse among college students. *Journal of Ethnicity in Substance Abuse*, *6*(2), 75-95.

11. Abuse, S. Mental Health Services Administration (SAMHSA), 2014. Results from the 2013 National Survey on Drug Use and Health: Summary of National Findings. NSDUH Series H-48, HHS Publication No.(SMA) 14-4863. *Substance Abuse and Mental Health Services Administration, Rockville, MD*.

12. Johnston, L. D., O'Malley, P. M., Bachman, J. G., & Schulenberg, J. E. (2005). *Monitoring the Future national survey results on drug use, 1975-2004. Volume II: College students and adults ages 19-45* (NIH Publication No. 05-5728). Bethesda, MD: National Institute on Drug Abuse, 278 pp.

13. Horton, E. G. (2007). Racial differences in the effects of age of onset on alcohol consumption and development of alcohol-related problems among males from mid-adolescence to young adulthood. *J Ethn Subst Abuse, 6*(1), 1-13. doi: 10.1300/J233v06n01_01.

14. Nasim, A., Belgrave, F. Z., Jagers, R. J., Wilson, K. D., & Owens, K. (2007). The moderating effects of culture on peer deviance and alcohol use among high-risk African-American adolescents. *Journal of Drug Education*, *37*(3), 335-363.

15. Duncan, S. C., Strycker, L. A., & Duncan, T. E. (2012). Alcohol use of African Americans and Whites from ages 9–20: Descriptive results from a longitudinal study. *Journal of ethnicity in substance abuse*, 11(3), 214-225.

16. Kochanek, K., Murphy, S., Anderson, R., & Scott, C. (2004). VDeaths: Final Data for 2002V. 7II. ED7B, 5, I7B.

17. Sartor, C. E., Agrawal, A., Lynskey, M. T., Duncan, A. E., Grant, J. D., Nelson, E. C., . . . Bucholz, K. K. (2013). Cannabis or alcohol first? Differences by ethnicity and in risk for rapid progression to cannabis-related problems in women. *Psychol Med, 43*(4), 813-823. doi: 10.1017/S0033291712001493.

18. French, K., Finkbiner, R., & Duhamel, L. (2002). Patterns of substance use among minority youth and adults in the United States: An overview and synthesis of national survey findings. *Fairfax, VA: Department of Health and Human Services*.

19. David, S. P., Hamidovic, A., Chen, G. K., Bergen, A. W., Wessel, J., Kasberger, J. L., . . . Furberg, H. (2012). Genome-wide meta-analyses of smoking behaviors in African Americans. *Transl Psychiatry, 2*, e119. doi: 10.1038/tp.2012.41.

20. McCarthy, D. M., Pedersen, S. L., Lobos, E. A., Todd, R. D., & Wall, T. L. (2010). ADH1B* 3 and Response to Alcohol in African-Americans. *Alcoholism: Clinical and Experimental Research*, *34*(7), 1274-1281.

21. Crabb, D. W., Matsumoto, M., Chang, D., & You, M. (2004). Overview of the role of alcohol dehydrogenase and aldehyde dehydrogenase and their variants in the genesis of alcohol-related pathology. *Proceedings of the Nutrition Society*, *63*(01), 49-63.

22. Jackson, C. L., Hu, F. B., Kawachi, I., Williams, D. R., Mukamal, K. J., & Rimm, E. B. (2015). Black–White Differences in the Relationship Between Alcohol Drinking Patterns

and Mortality Among US Men and Women. *American journal of public health*, *105*(S3), S534-S543.

23. Horton, E. G. (2007). Racial differences in the effects of age of onset on alcohol consumption and development of alcohol-related problems among males from mid-adolescence to young adulthood. *J Ethn Subst Abuse*, *6*(1), 1-13. doi: 10.1300/J233v06n01_01.

24. Hawkins, J. D., Catalano, R. F., & Miller, J. Y. (1992). Risk and protective factors for alcohol and other drug problems in adolescence and early adulthood: implications for substance abuse prevention. *Psychological bulletin*, *112*(1), 64.

25. Anthenelli, R. M. (2012). Overview: stress and alcohol use disorders revisited. *Alcohol Research-Current Reviews*, *34*(4), 386.

26. Walden, B., McGue, M., Burt, S. A., & Elkins, I. (2004). Identifying shared environmental contributions to early substance use: the respective roles of peers and parents. *Journal of Abnormal Psychology*, *113*(3), 440.

27. Wallace, J. M. & Muroff, J. R. (2002). Preventing Substance Abuse Among African American Children and Youth: Race Differences in Risk Factor Exposure and Vulnerability. *Journal of Primary Prevention, 22*(3), 235-261.

28. Heath, A. C., Madden, P. A. F., Grant, J. D., McLaughlin, T. L., Todorov, A. A. & Bucholz, K. K. (1999). Resiliency factors protecting against teenage alcohol use and smoking: influences of religion, religious involvement and values, and ethnicity in the Missouri Adolescent Female Twin Study. *Twin Research, 2*, 145-155.

29. Peterson, P. L., Hawkins, J. D., Abbott, R. D., & Catalano, R. F. (1994). Disentangling the effects of parental drinking, family management, and parental alcohol norms on current drinking by black and white adolescents. *Journal of Research on Adolescence*, *4*(2), 203-227.

30. Jones-Webb, R. (1998). Drinking patterns and problems among African-Americans: Recent findings. *Alcohol Research and Health*, *22*(4), 260.

31. Lombe, M., Yu, M., Nebbitt, V., & Earl, T. (2011) Understanding Alcohol Consumption and Its Correlates among African American Youths in Public Housing: A Test of Problem Behavior Theory. *Social Work Research*. *35*, 173-182. doi: 10.1093/swr/35.3.173.

32. Wallace, J. M. & Muroff, J. R. (2002). Preventing Substance Abuse Among African American Children and Youth: Race Differences in Risk Factor Exposure and Vulnerability. *Journal of Primary Prevention, 22*(3), 235-261.

33. Wallace, J.M., Jr. (1999). Explaining Race Differences in Adolescent and Young Adult Drug Use: the Role of Racialized Social Systems. Drugs & Society, 14(1–2) US: Haworth Press Inc., 21–36.

34. Chartier, K. G. Hesselbrock, M. N., Hesselbrock, V. M. (2009). Ethnicity and Adolescent Pathways to Alcohol Use. *J. Stud. Alcohol Drugs, 70,* 337-345.

35. Slutske, W. S., Cronk, N. J., Sher, K. J., Madden, P. A., Bucholz, K. K., & Heath, A. C. (2002). Genes, environment and individual differences in alcohol expectancies among female adolescents and young adults. *Psychology of Addictive Behaviors, 16*(4), 308.

36. Stock, M. L., Gibbons, F. X., Gerrard, M., Houlihan, A. E., Weng, C. Y., Lorenz, F. O., & Simons, R. L. (2013). Racial identification, racial composition, and substance use vulnerability among African American adolescents and young adults. *Health Psychol, 32*(3), 237-247. doi: 10.1037/a0030149.

37. Flint, J., Greenspan, R. J., Kendler, K. S. (2010). *How Genes Influence Behavior*. New York: Oxford University Press Inc.

38. Neale, M., & Cardon, L. (1992). *Methodology for genetic studies of twins and families* (Vol. 67). Springer Science & Business Media.

39. Prescott, C. A., Caldwell, C. B., Carey, G., Vogler, G. P., Trumbetta, S. L., & Gottesman, II. (2005). The Washington University Twin Study of alcoholism. *Am J Med Genet B Neuropsychiatr Genet, 134B*(1), 48-55. doi: 10.1002/ajmg.b.30124.

40. Sartor, C. E., Agrawal, A., Lynskey, M. T., Bucholz, K. K., Madden, P. A., & Heath, A. C. (2009). Common genetic influences on the timing of first use for alcohol, cigarettes, and cannabis in young African-American women. *Drug Alcohol Depend, 102*(1-3), 49-55. doi: 10.1016/j.drugalcdep.2008.12.013.

41. Fowler, T., Lifford, K., Shelton, K., Rice, F., Thapar, A., Neale, M. C., . . . van den Bree, M. B. (2007). Exploring the relationship between genetic and environmental influences on initiation and progression of substance use. *Addiction, 102*(3), 413-422. doi: 10.1111/j.1360-0443.2006.01694.x.

42. Sullivan, P. F. Kendler, K. S. (1999). The genetic epidemiology of smoking. *Nicotine & Tobacco Research, 1*, 51-57.

43. Heath, A.C, Bucholz, K.K., Madden, P. A. F., Dinwiddie, S. H., Slutske, W. S., Bierut, L. J., … Martin, N. G. (1997). Genetic and environmental contributions to alcohol dependence risk in a national twin sample: Consistency of findings in women and men. *Psychol Med*. 27:1381–1396.

44. McGue, M., Pickens, R. W., & Svikis, D. S. (1992). Sex and age effects on the inheritance of alcohol problems: A twin study. *J Abnorm Psychol*. 101:3–17.

45. Prescott, C. A. & Kendler, K. S. (1999). Age at first drink and risk for alcoholism: a noncausal association. *Alcohol Clin Exp Res, 23*, 101-107.

46. Marigorta, U. M., & Navarro, A. (2013). High trans-ethnic replicability of GWAS results implies common causal variants. *PLoS Genet, 9*(6), e1003566. doi: 10.1371/journal.pgen.1003566.

47. Rosenberg, N. A., Huang, L., Jewett, E. M., Szpiech, Z. A., Jankovic, I., & Boehnke, M. (2010). Genome-wide association studies in diverse populations. *Nat Rev Genet, 11*(5), 356-366. doi: 10.1038/nrg2760.

48. Fu, J., Festen, E. A., & Wijmenga, C. (2011). Multi-ethnic studies in complex traits. *Hum Mol Genet, 20*(R2), R206-213. doi: 10.1093/hmg/ddr386.

49. Thomas, D. C., & Witte, J. S. (2002). Point: population stratification: a problem for case-control studies of candidate-gene associations?. Cancer Epidemiology Biomarkers & Prevention, 11(6), 505-512.

50. Dick, D. M., Latendresse, S. J., & Riley, B. (2011). Incorporating genetics into your studies: A guide for social scientists. Frontiers in psychiatry, 2(17), 1-11.

51. Gabriel, S. B., Schaffner, S. F., Nguyen, H., Moore, J. M., Roy, J., Blumenstiel, B., . . . Altshuler, D. (2002). The structure of haplotype blocks in the human genome. *Science, 296*(5576), 2225-2229. doi: 10.1126/science.1069424.

52. Adeyemo, A., & Rotimi, C. (2010). Genetic variants associated with complex human diseases show wide variation across multiple populations. *Public Health Genomics, 13*(2), 72-79. doi: 10.1159/000218711.

53. Li, M. D., Mangold, J. E., Seneviratne, C., Chen, G. B., Ma, J. Z., Lou, X. Y., & Payne, T. J. (2009). Association and interaction analyses of GABBR1 and GABBR2 with nicotine dependence in European- and African-American populations. *PLoS One, 4*(9), e7055. doi: 10.1371/journal.pone.0007055.

54. Chen, L. S., Saccone, N. L., Culverhouse, R. C., Bracci, P. M., Chen, C. H., Dueker, N., . . . Bierut, L. J. (2012). Smoking and genetic risk variation across populations of European, Asian, and African American ancestry--a meta-analysis of chromosome 15q25. *Genet Epidemiol, 36*(4), 340-351. doi: 10.1002/gepi.21627.

55. Heath, A. C. & Nelson, E. C. (2002). Effects of the Interaction Between Genotype and Environment Research into the Genetic Epidemiology of Alcohol Dependence. *Alcohol Res Health*. *26*(3): 193-201.

56. Khoury, M.J., Beaty, T.H., and Cohen, B.H. *Fundamentals of Genetic Epidemiology*. New York: Oxford University Press, 1993.

57. Latvala, A., Dick, D.M., Tuulio-Henriksson, A., Suvisaari, J., Viken, R., Rose, R.J., & Kaprio, J. (2011). Genetic correlation and gene-environment interaction between alcohol problems and education level in young adulthood. *J Stud Alcohol Drugs*. *72*, 210-220.

58. Olfson, E., Edenberg, H. J., Nurnberger, J., Agrawal, A., Bucholz, K. K., Almasy, L. A., ... & Kuperman, S. (2014). An ADH1B Variant and Peer Drinking in Progression to Adolescent Drinking Milestones: Evidence of a Gene-by-Environment Interaction. *Alcoholism: Clinical and Experimental Research*, *38*(10), 2541-2549.

59. Salvatore, J. E., Aliev, F., Bucholz, K., Agrawal, A., Hesselbrock, V., Hesselbrock, M., ... & Edenberg, H. J. (2014). Polygenic risk for externalizing disorders gene-by-development and gene-by-environment effects in adolescents and young adults. *Clinical Psychological Science*, 2167702614534211.

60. Office of Juvenile Justice and Delinquency Prevention. (2012). Effects and Consequences of Underage Drinking. Retrieved from http://www.ojjdp.gov/pubs/237145.pdf.

61. Centers for Disease Control and Prevention. (2002). Youth Risk Behavior Survey. Available at: www.cdc.gov/yrbs.

62. Horton EG. Racial Differences in the Effects of Age of Onset on Alcohol Consumption and Development of Alcohol-Related Problems Among Males from Mid-Adolescence to Young Adulthood. *J Ethn Subst Abuse*. 2007; 6(1): 1-13. doi: 10.1300/ J233v06n01_01.

63. French K, Finkbiner R, & Duhamel L. Patterns of substance use among minority youth and adults in the United States: An overview and synthesis of national survey findings (National Evaluation Data Services [NEDS] Technical Report, Center for Substance Abuse Treatment, Substance Abuse and Mental Health Services Administration, contract no. 270-97-7016) Fairfax, VA: Caliber Associates. 2002.

64. David SP, Hamidovic A, Chen GK, et al. Genome-wide meta-analyses of smoking behaviors in African Americans. *Transl Psychiatry*, 2012*;* 2: e119. doi: 10.1038/tp.2012.41.

65. Nebbitt V, Lombe M. Environmental correlates of depressive symptoms among African American adolescents living in public housing. *J Hum Behav Soc Environ*. 2007; 15: 435–454. doi: 10.1300/J137v15n02_24.

66. Substance Abuse and Mental Health Services Administration. (2009). Results from the 2008 National Survey on Drug Use and Health: National findings (Office of Applied Studies, NSDUH Series H-36, HHS Publication No. SMA 09-4434). Rockville, MD.

67. Green JG, McLaughlin KA, Berglund PA, et al. Childhood adversities and adult psychopathology in the National Comorbidity Survey Replication (NCS-R) I: Associations with first onset of DSM-IV disorders. Arch Gen Psychiatry. 2010; 67(2): 113-134. doi:10.1001/archgenpsychiatry.2009.186.

68. Sapolsky RM, Romero LM, Munck AU. How do glucocorticoids influence stress responses? Integrating permissive, suppressive, stimulatory and preparative actions. *Endocr Rev*. 2000; 28: 55-89. doi: 10.1210/er.21.1.55.

69. Dube SR, Miller JW, Brown DW, et al. Adverse childhood experiences and the association with ever using alcohol and initiating alcohol use during adolescence. *Journal of Adolescent Health*. 2006; 38(4): 444.e1-444.e10. doi: 10.1016/j.jadohealth.2005.06.006.

70. Pilowsky, D. J., Keyes, K. M., & Hasin, D. S. (2009). Adverse childhood events and lifetime alcohol dependence. *American journal of public health*, *99*(2), 258-263.

71. Wills TA, Sandy JM, Yaeger AM, Cleary SD, Shinar O. Coping dimensions, life stress, and adolescent substance use: A latent growth analysis. *J Abnorm. Psychol*. 2001; 110: 309-323.

72. King KM, Chassin L. Adolescent Stressors, Psychopathology, and Young Adult Substance Dependence: A Prospective Study. *J Stud Alcohol Drugs*. 2008; 69(5): 629-638.

73. Havey JM, Dodd DK. Children of alcoholics, negative life events, and early experimentation with drugs. *J. School Psychol*. 1995; 33*:* 305-317.

74. Lombe M, Yu M, Nebbitt V, Earl T. Understanding Alcohol Consumption and Its Correlates among African American Youths in Public Housing: A Test of Problem Behavior Theory. *Social Work Research*. 2011; 35: 173-182. doi: 10.1093/swr/35.3.173.

75. Doherty EE, Robertson JA, Green KM, Fothergill KE, Ensminger ME. A longitudinal study of substance use and violent victimization in adulthood among a cohort of urban African Americans. *Addiction*. 2012; 107: 339-348. doi: 10.1111/j.1360-0443.2011.03665.x.

76. Kendler KS, Schmitt E, Aggen SH, Prescott CA. Genetic and environmental influences on alcohol, caffeine, cannabis, and nicotine use from early adolescence to middle adulthood. *Arch Gen Psychiatry*. 2008; 65(6): 674-682.

77. Maes HH, Woodard CE, Murrelle L, et al. Tobacco, alcohol and drug use in eight- to sixteen-year-old twins: the Virginia twin study of adolescent behavioral development. *J Stud Alcohol*. 1999; 60: 293-305.

78. Poelen EAP, Derks EM, Engels RCME, et al. The Relative Contribution of Genes and Environment to Alcohol Use in Eary Adolescents: Are Similar Factors Related to Initiation of Alcohol Use and Frequency of Drinking? *Alcohol Clin Exp Res*. 2008; 32(6): 975-982.

79. Rose RJ, Dick DM, Viken RJ, et al. Drinking of abstaining at age 14? A genetic epidemiological study. *Alcohol Clin Exp Res.* 2001; 25:594-1604.

80. Viken RJ, Kapiro J, Koskenvuo M, et al. Longitudinal analyses of the determinants of drinking and of drinking to intoxication in adolescent twins. *Behav Genet. 1999;* 29:455-461.

81. Covault J, Tennen H, Armeli S, et al. Interactive Effects of the Serotonin Transporter 5-HTTLPR Polymorphism and Stressful Life Events on College Student Drinking and Drug Use. *Biol Psychiatry*. 2007; 61: 609-616.

82. Kaufman J, Yang BZ, Douglas-Palumberi H, et al. Genetic and environmental predictors of early alcohol use. *Biol Psychiatry*. 2007; 61(11): 1228-1234.

83. Madrid GA, MacMurray J, Lee JW, Anderson BA, Comings DE. Stress as a mediating factor in the association between the DRD2 *TaqI* polymorphism and alcoholism. *Alcohol*. 2001; 23: 117-122.

84. Bau CH, Almeida S, Hutz MH. The TaqI A1 allele of the dopamine D2 receptor gene and alcoholism in Brazil: association and interaction with stress and harm avoidance on severity prediction. *Am J Med Genet.* 2000; 96(3): 302-306.28. Hansson AC, Cippitelli A, Sommer WH, et al. Variation at the rat Crhr1 locus and sensitivity to relapse into alcohol seeking induced by environmental stress. *Proc Natl Acad Sci USA.* 2006; 103(41): 15236-15241.

85. Hansson AC, Cippitelli A, Sommer WH, et al. Variation at the rat Crhr1 locus and sensitivity to relapse into alcohol seeking induced by environmental stress. *Proc Natl Acad Sci USA.* 2006; 103(41): 15236-15241.

86. Blomeyer D, Treutlein J, Esser G, Schmidt M, Schumann G, Laucht M. Interaction between CRHR1 Gene and Stressful Life Events Predicts Adolescent Heavy Alcohol Use. *Biol Psychiatry*. 2008; 63(2): 146-151.

87. Blomeyer D, Buchmann AF, Lascorz J, et al. Association of PER2 genotype and Stressful Life Events with Alcohol Drinking in Young Adults. *PLoS ONE*. 2013; 8(3): e59136.

88. Ambroggi F, Turiault M, Milet A, et al. Stress and addiction: glucocorticoid receptor in dopaminoceptive neurons facilitates cocaine seeking. *Nat Neurosci*. 2009; 12: 247–249.

89. Stevens A, Ray DW, Zeggini E, et al. Glucocorticoid Sensitivity is Determined by a Specific Glucocorticoid Receptor Haplotype. *J Clin Endocrinol Metab*. 2004; 89(2):892-897.

90. DeRijk RH, Schaaf M, de Kloet ER. Glucocorticoid receptor variants: clinical implications. *J Steroid Biochem Mol Biol*. 2002; 81(2):103-122.

91. Lupien SJ, McEwen BS, Gunnar MR, Heim C. Effects of stress throughout the lifespan on the brain, behaviour, and cognition. *Nat Rev Neurosci*. 2009; 10: 435-445. doi:10.1038/nrn2639.

92. Ray L. Stress-Induced and Cue-Induced Craving for Alcohol in Heavy Drinkers: Preliminary Evidence of Genetic Moderation by the OPRM1 and CRH-BP Genes. *Alcohol Clin Exp Res*. 2011; 35(1): 166-174.

93. Enoch MA, Shen PH, Ducci F, et al. Common Genetic Origins for EEG, Alcoholism and Anxiety: The Role of CRH-BP. *PLoS ONE*. 2008; *3*(10): e3620. doi:10.1371/journal.pone.0003620.

94. Kertes DA, Kalsi G, Prescott CA, et al. Neurotransmitter and Neuromodulator Genes Associated with a History of Depressive Symptoms in Individuals with Alcohol Dependence. *Alcohol Clin Exp Res.* 2011; 35(3): 496-505.

95. Shanahan, M. J., & Hofer, S. M. (2005). Social context in gene–environment interactions: Retrospect and prospect. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *60*(Special Issue 1), 65-76.

96. Bolland, J. M., Lian, B. E., & Formichella, C. M. (2005). The origins of hopelessness among inner-city African-American adolescents. *American journal of community psychology*, *36*(3-4), 293-305.

97. Bolland, J. M., Bryant, C. M., Lian, B. E., McCallum, D. M., Vazsonyi, A. T., & Barth, J. M. (2007). Development and Risk Behavior Among African American, Caucasian, and Mixed-race Adolescents Living in High Poverty Inner-city Neighborhoods. *American Journal of Community Psychology*, *40*(3-4), 230-249.

98. Donenberg, G. R., Emerson, E., Bryant, F. B., Wilson, H., & Weber-Shifrin, E. (2001). Understanding AIDS-risk behavior among adolescents in psychiatric care: Links to psychopathology and peer relationships. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*(6), 642-653.

99. National Institute on Drug Abuse. (1995). Prevalence of Drug Use in the DC Metropolitan Area Adult and Juvenile Offender Populations: 1991 Rockville, MD: US Department of Health and Human Services.

100. Institute of Behavioral Science. (1991). *Denver youth survey- Youth interview schedule*: University of Colorado.

101. Needle, R., Fisher, D. G., Weatherby, N., Chitwood, D., Brown, B., Cesari, H., ... & Braunstein, M. (1995). Reliability of self-reported HIV risk behaviors of drug users. *Psychology of addictive behaviors*, *9*(4), 242.

102. Weatherby, N. L., Needle, R., Cesari, H., Booth, R., McCoy, C. B., Watters, J. K., ... & Chitwood, D. D. (1994). Validity of self-reported drug use among injection drug users and crack cocaine users recruited through street outreach. *Evaluation and Program planning*, *17*(4), 347-355.

103. Watters JK. (1994). Street Youth at Risk for AIDS (final report). Rockville, MD: National Institute on Drug Abuse.

104. Attar, B. K., Guerra, N. G., & Tolan, P. H. (1994). Neighborhood disadvantage, stressful life events and adjustments in urban elementary-school children. *Journal of Clinical Child Psychology*, *23*(4), 391-400.

105. Gorman–Smith, D., & Tolan, P. (1998). The role of exposure to community violence and developmental problems among inner-city youth. *Development and psychopathology*, *10*(01), 101-116.

106. van den Oord EJ, Sullivan PF, Jiang Y, Walsh D, O'Neill FA, Kendler KS, Riley BP. Identification of a high-risk haplotype for the dystrobrevin binding protein 1 (DTNBP1) gene in the Irish study of high-density schizophrenia families. *Mol Psychiatry*. 2003b; 8: 499–510.

107. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005; 21(2): 263-265.

108. Nyholt DR. A Simple Correction for Multiple Testing for Single-Nucleotide Polymorphisms in Linkage Disequilibrium with Each Other. *Am J Hum Genet*. 2004; 74(4): 765-769.

109. Aliev F, Latendresse SJ, Bacanu SA, Neale MC, Dick DM. Testing for measured gene-environment interaction: Problems with the use of cross-product terms and a regression model reparameterization solution. *Behavior Genetics*, (under review).

110. Belsky J, Bakermans-Kranenburg MJ, van IJzendoorn MH. For Better and For Worse Differential Susceptibility to Environmental Influences. *Curr Dir Psychol Sci*. 2007; 16: 300–304.

111. Wang B, You ZB, Rice KC, Wise RA. Stress-induced relapse to cocaine seeking: roles for the CRF_2 receptor and CRF-binding protein in the ventral tegmental area of the rat. *Psychopharmacology.* 2007; 193: 293-294.

112. Johnston LD, O'Malley PM, Bachman JG, Schulenberg JE. *Monitoring the Future national survey results on drug use, 1975-2004. Volume II: College students and adults ages 19-45* (NIH Publication No. 05-5728). Bethesda, MD: National Institute on Drug Abuse, 2005; 1-278.

113. Sartor, CE, Nelson EC, Lynskey MT, Madden PA, Heath AC, Bucholz KK. Are There Differences Between Young African-American and European-American Women in the Relative Influences of Genetics Versus Environment on Age at First Drink and Problem Alcohol Use? *Alcohol Clin Exp Res*. 2013; 37(11): 1939-1946. doi: 10.1111/acer.12185.

114. Biafora F, Zimmerman R. (1998). Developmental patterns of African American adolescent drug use. In: Vega WA, Gil AG, eds. *Drug use and ethnicity in early adolescence.* New York, NY: Plenum; 1998: 149-175.

115. Caldwell CH, Sellers RM, Bernat DH, Zimmerman MA. Racial identity, parental support, and alcohol use in a sample of academically at-risk African American high school students. *American Journal of Community Psychology,* 2004; 34(1/2): 71-82.

116. Youth Risk Behavior Survey. (2010). *Healthy living*. Retrieved from http://www.cdc.gov/Features/ RiskBehavior/.

117. Rosenberg NA, Huang L, Jewett EM, Szpiech ZA, Jankovic I, Boehnke M. Genome-wide association studies in diverse populations. *Nat Rev Genet.* 2010; 11(5): 356-366. doi: 10.1038/nrg2760.

118. Brody GH, Chen Y, Beach SRH, et al. Differential Sensitivity to Prevention Programmin: A Dopaminergic Polymorphism-Enhanced Prevention Effect on Protective Parenting and Adolescent Substance Use. *Health Psychology.* 2014; 33(2):182-191.

119. Brody GH, Beach SRH, Philibert RA. Prevention Effects Moderate the Association of 5-HTTLPR and Youth Risk Behavior Initiation: Gene x Environment Hypotheses Tested via a Randomized Prevention Design. *Child Dev*. 2009; 80(3): 645-661.

120. Beach SRH, Brody G, Lei MK, Philibert RA. Differential Susceptibility to Parenting among African American Youths: Testing the DRD4 Hypothesis. *J Fam Psychol*. 2010; 24(5):

513-521.

121. Meyers, J. L., Salvatore, J. E., Vuoksimaa, E., Korhonen, T., Pulkkinen, L., Rose, R. J., ... & Dick, D. M. (2014). Genetic influences on alcohol use behaviors have diverging developmental trajectories: a prospective study among male and female twins. *Alcoholism: Clinical and Experimental Research*, *38*(11), 2869-2877.

122. Kendler, K. S., Gardner, C., & Dick, D. M. (2011). Predicting alcohol consumption in adolescence from alcohol-specific and general externalizing genetic risk factors, key environmental exposures and their interaction. *Psychological medicine*, *41*(07), 1507-1516.

123. Knopik, V. S., Heath, A. C., Madden, P. A., Bucholz, K. K., Slutske, W. S., Nelson, E. C., ... & Martin, N. G. (2004). Genetic effects on alcohol dependence risk: re-evaluating the importance of psychiatric and other heritable risk factors. *Psychological medicine*, *34*(08), 1519-1530.

124. Rietschel, M., & Treutlein, J. (2013). The genetics of alcohol dependence. *Annals of the New York Academy of Sciences*, *1282*(1), 39-70.

125. Edenberg, H. J. (2007). The genetics of alcohol metabolism: role of alcohol dehydrogenase and aldehyde dehydrogenase variants. *Alcohol Research & Health*, *30*(1), 5-14.

126. Samochowiec, J., Samochowiec, A., Puls, I., Bienkowski, P., & Schott, B. H. (2014). Genetics of alcohol dependence: a review of clinical studies. *Neuropsychobiology*, *70*(2), 77-94.

127. Prescott, C. A., Sullivan, P. F., Kuo, P. H., Webb, B. T., Vittum, J., Patterson, D. E. E. A., ... & Robinson, V. P. (2006). Genomewide linkage study in the Irish affected sib pair study of alcohol dependence: evidence for a susceptibility region for symptoms of alcohol dependence on chromosome 4. *Molecular psychiatry*, *11*(6), 603-611.

128. Foroud, T., Edenberg, H. J., Goate, A., Rice, J., Flury, L., Koller, D. L., ... & Li, T. K. (2000). Alcoholism susceptibility loci: confirmation studies in a replicate sample and further mapping. *Alcoholism: Clinical and Experimental Research*, *24*(7), 933-945.

129. Kuo, P. H., Neale, M. C., Riley, B. P., Webb, B. T., Sullivan, P. F., Vittum, J., ... & Kendler, K. S. (2006). Identification of Susceptibility Loci for Alcohol-Related Traits in the Irish Affected Sib Pair Study of Alcohol Dependence. *Alcoholism: Clinical and Experimental Research*, *30*(11), 1807-1816.

130. Bierut, L. J., Agrawal, A., Bucholz, K. K., Doheny, K. F., Laurie, C., Pugh, E., ... & Hinrichs, A. L. (2010). A genome-wide association study of alcohol dependence. *Proceedings of the National Academy of Sciences*, *107*(11), 5082-5087.

131. Zuo, L., Gelernter, J., Zhang, C. K., Zhao, H., Lu, L., Kranzler, H. R., ... & Deng, H. W. (2012). Genome-wide association study of alcohol dependence implicates KIAA0040 on chromosome 1q. *Neuropsychopharmacology*, *37*(2), 557-566.

132. Wang, K. S., Liu, X., Zhang, Q., Pan, Y., Aragam, N., & Zeng, M. (2011). A meta-analysis of two genome-wide association studies identifies 3 new loci for alcohol dependence. *Journal of psychiatric research*, *45*(11), 1419-1425.

133. Sterling, R., Henderson, G. E., & Corbie-Smith, G. (2006). Public willingness to participate in and public opinions about genetic variation research: a review of the literature. *American journal of public health*, *96*(11), 1971-1978.

134. Schulz, A., Caldwell, C., & Foster, S. (2003). "What are they going to do with the information?" Latino/Latina and African American perspectives on the Human Genome Project. *Health education & behavior*, *30*(2), 151-169.

135. Murphy, E., & Thompson, A. (2009). An exploration of attitudes among black Americans towards psychiatric genetic research. *Psychiatry*, *72*(2), 177-194.

136. Royal, C., Baffoe-Bonnie, A., Kittles, R., Powell, I., Bennett, J., Hoke, G., ... & Mason, T. (2000). Recruitment experience in the first phase of the African American Hereditary Prostate Cancer (AAHPC) study. *Annals of epidemiology*, *10*(8), S68-S77.

137. Hoyo, C., Reid, M. L., Godley, P. A., Parrish, T., Smith, L., & Gammon, M. (2002). Barriers and strategies for sustained participation of African-American men in cohort studies. *Ethnicity & Disease*, *13*(4), 470-476.

138. Haga, S. B. (2010). Impact of limited population diversity of genome-wide association studies. *Genetics in Medicine*, *12*(2), 81-84.

139. Schnittker, J., Freese, J., & Powell, B. (2000). Nature, nurture, neither, nor: Black-White differences in beliefs about the causes and appropriate treatment of mental illness. *Social Forces*, *78*(3), 1101-1132.

140. Campbell, M. C., & Tishkoff, S. A. (2008). African genetic diversity: implications for human demographic history, modern human origins, and complex disease mapping. *Annual review of genomics and human genetics*, *9*, 403.

141. Gonder, M. K., Mortensen, H. M., Reed, F. A., de Sousa, A., & Tishkoff, S. A. (2007). Whole-mtDNA genome sequence analysis of ancient African lineages. *Molecular biology and evolution*, *24*(3), 757-768.

142. Gelernter, J., Kranzler, H. R., Sherva, R., Almasy, L., Koesterer, R., Smith, A. H., ... & Wodarz, N. (2014). Genome-wide association study of alcohol dependence: significant findings in African-and European-Americans including novel risk loci. *Molecular psychiatry*, *19*(1), 41-49.

143. Steele, R. G., Forehand, R., Armistead, L., & Brody, G. (1995). Predicting alcohol and drug use in early adulthood: The role of internalizing and externalizing behavior problems in early adolescence. *American Journal of Orthopsychiatry*, *65*(3), 380.

144. Englund, M. M., Egeland, B., Oliva, E. M., & Collins, W. A. (2008). Childhood and adolescent predictors of heavy drinking and alcohol use disorders in early adulthood: a longitudinal developmental analysis. *Addiction*, *103*(s1), 23-35.

145. Edwards, A. C., & Kendler, K. S. (2012). Twin study of the relationship between adolescent attention-deficit/hyperactivity disorder and adult alcohol dependence. *Journal of studies on alcohol and drugs*, *73*(2), 185-194.

146. Dick, D. M., Bierut, L., Hinrichs, A., Fox, L., Bucholz, K. K., Kramer, J., ... & Tischfield, J. (2006). The role of GABRA2 in risk for conduct disorder and alcohol and drug dependence across developmental stages. *Behavior genetics*, *36*(4), 577-590.

147. Iacono, W. G., Malone, S. M., & McGue, M. (2008). Behavioral disinhibition and the development of early-onset addiction: common and specific influences. *Annu. Rev. Clin. Psychol.*, *4*, 325-348.

148. Young-Wolff, K. C., Kendler, K. S., & Prescott, C. A. (2012). Interactive effects of childhood maltreatment and recent stressful life events on alcohol consumption in adulthood. *Journal of studies on alcohol and drugs*, *73*(4), 559-569.

149. Williams, D. R., & Collins, C. (2001). Racial residential segregation: a fundamental cause of racial disparities in health. *Public health reports*, *116*(5), 404.

150. Carroll, G. (1998). Mundane extreme environmental stress and African American families: A case for recognizing different realities. *Journal of Comparative Family Studies*, 271-284.

151. Williams, D. R., Neighbors, H. W., & Jackson, J. S. (2003). Racial/ethnic discrimination and health: findings from community studies. *American journal of public health*, *93*(2), 200-208.

152. Dawson, D. A. (2000). Alcohol consumption, alcohol dependence, and all-cause mortality. *Alcoholism: Clinical and Experimental Research*, *24*(1), 72-81.

153. Achenbach, T.M. & Rescorla, L.A. (2001). Manual for the ASEBA school-age forms & profiles: an integrated system of mult-informant assessment. Burlington: University of Vermont, Research Center for Children, Youth & Families.

154. Anderson, C. A., Pettersson, F. H., Clarke, G. M., Cardon, L. R., Morris, A. P., & Zondervan, K. T. (2010). Data quality control in genetic case-control association studies. *Nature protocols*, *5*(9), 1564-1573.

155. Laurie, C. C., Doheny, K. F., Mirel, D. B., Pugh, E. W., Bierut, L. J., Bhangale, T., ... & Gabriel, S. B. (2010). Quality control and quality assurance in genotypic data for genome-wide association studies. *Genetic epidemiology*, *34*(6), 591-602.

156. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., ... & Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, *81*(3), 559-575.

157. Thornton, T., Tang, H., Hoffmann, T. J., Ochs-Balcom, H. M., Caan, B. J., & Risch, N. (2012). Estimating kinship in admixed populations. *The American Journal of Human Genetics*, *91*(1), 122-138.

158. Delaneau, O., Marchini, J., & Zagury, J. F. (2012). A linear complexity phasing method for thousands of genomes. *Nature methods*, *9*(2), 179-181.

159. Delaneau, O., Zagury, J. F., & Marchini, J. (2013). Improved whole-chromosome phasing for disease and population genetic studies. *Nature methods*, *10*(1), 5-6.

160. Howie, B. N., Donnelly, P., & Marchini, J. (2009). A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet*, *5*(6), e1000529.

161. Holland, P. M., Abramson, R. D., Watson, R., & Gelfand, D. H. (1991). Detection of specific polymerase chain reaction product by utilizing the 5'----3'exonuclease activity of Thermus aquaticus DNA polymerase. *Proceedings of the National Academy of Sciences*, *88*(16), 7276-7280.

162. Evans, D. M., Visscher, P. M., & Wray, N. R. (2009). Harnessing the information contained within genome-wide association studies to improve individual prediction of complex disease risk. *Human molecular genetics*, *18*(18), 3525-3531.

163. Hicks, B. M., South, S. C., DiRago, A. C., Iacono, W. G., & McGue, M. (2009). Environmental adversity and increasing genetic risk for externalizing disorders. *Archives of general psychiatry*, *66*(6), 640-648.

164. Meyers, J. L., Cerdá, M., Galea, S., Keyes, K. M., Aiello, A. E., Uddin, M., ... & Koenen, K. C. (2013). Interaction between polygenic risk for cigarette use and environmental exposures in the Detroit neighborhood health study. *Translational psychiatry*, *3*(8), e290.

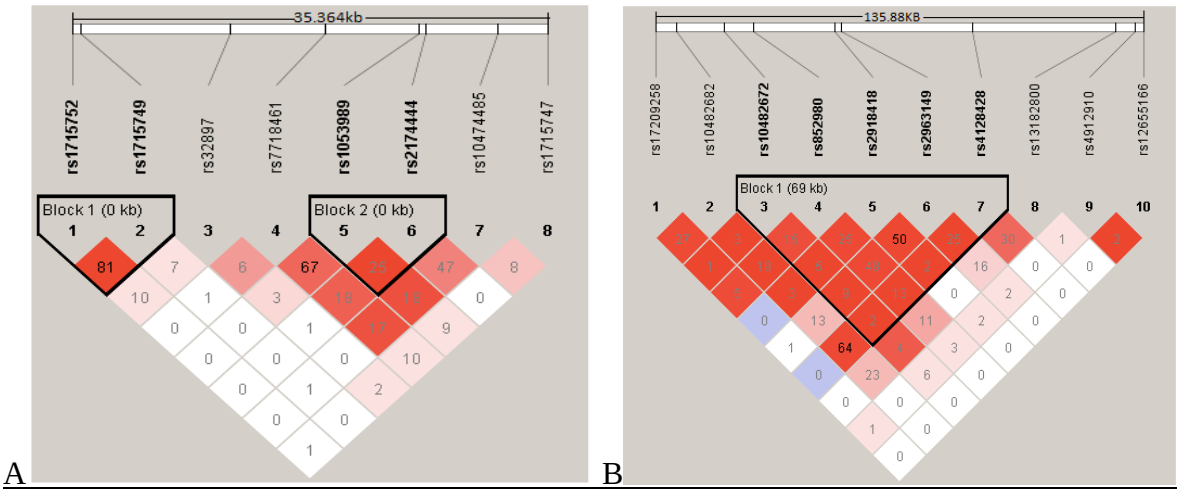
165. D'Avanzo C, Dunn P, Murdock J, Naegle M. Developing culturally informed strategies for substance-related interventions. In: Naegle MA, D'Avanzo CE, editors. Addictions And Substance Abuse: Strategies For Advanced Practice Nursing. Upper Saddle River, NJ: Prentice Hall Health; 2001. pp. 59–74.

166. Tishkoff, S. A., & Williams, S. M. (2002). Genetic analysis of African populations: human evolution and complex disease. *Nature Reviews Genetics*, *3*(8), 611-621.

167. Gibson, G. (2012). Rare and common variants: twenty arguments. *Nature Reviews Genetics*, *13*(2), 135-145.

168. Dudbridge, F. (2013). Power and predictive accuracy of polygenic risk scores. *PLoS Genet*, *9*(3), e1003348.

Appendix A.1



A B

Figure 1. Haploview plot of linkage disequilibrium structure (r^2) across the genotyped single nucleotide polymorphisms in A)*CRHBP* and B)*NR3C1* using African Americans from the GENI sample.

Appendix A.2

Table 1. Moderated Multiple Regression of CRHBP and NR3C1 SNPs and Total Stressful Life Events Predicting Alcohol Use/Misuse in the Whole Sample

	SNP Position	Relative Position	Minor Allele	Percentage of Subjects with Copies of Minor Allele			Alcohol Use/Misuse			
							SNP Main Effect[a]		SNP x SLE Interaction[b]	
				0	1	2	B	*p*	B	*p*
SNP										
CRHBP										
rs1715752	chr5:76274929	Promoter	T	36.3	46.5	17.2	0.05	0.44	-0.04	0.05
rs1715749	chr5:76275603	Promoter	T	31.5	47.4	21.1	0.02	0.73	-0.06	0.006*
rs32897	chr5:76286728	Intron 3	C	49.6	42.0	8.4	0.17	0.01	-0.01	0.55
rs7718461	chr5:76293804	Intron 5	A	56.4	36.7	6.8	-0.14	0.04	-0.01	0.70
rs1053989	chr5:76300791	3' UTR	C	56.1	37.2	6.6	-0.14	0.05	-0.01	0.69
rs2174444	chr5:76301278	Downstream	T	33.2	50.2	16.6	0.02	0.73	-0.02	0.44
rs10474485	chr5:76306609	Downstream	A	35.4	49.7	14.9	0.03	0.61	0.01	0.60
rs1715747	chr5:76310293	Downstream	T	52.3	39.2	8.6	-0.08	0.26	-0.02	0.41
NR3C1										
rs17209258	chr5:142653590	Intron 7	G	91.0	8.8	0.2	-0.13	0.40	0.02	0.75
rs10482682	chr5:142659590	Intron 5	T	75.8	22.5	1.7	-0.03	0.76	0.02	0.47
rs10482672	chr5:142672726	Intron 3	A	64.6	31.5	3.9	-0.14	0.08	-0.02	0.36
rs852980	chr5:142681049	Intron 2	G	29.0	55.7	15.4	-0.05	0.47	-0.01	0.78
rs2918418	chr5:142703566	Intron 2	C	69.7	27.1	3.2	0.02	0.84	-0.01	0.73
rs2963149	chr5:142705277	Intron 2	T	50.7	42.0	7.3	0.03	0.68	-0.01	0.70
rs4128428	chr5:142742006	Intron 2	C	82.7	16.3	0.9	0.02	0.88	0.02	0.67
rs13182800	chr5:142781673	Intron 1	T	67.6	27.5	4.9	0.08	0.30	0.03	0.30

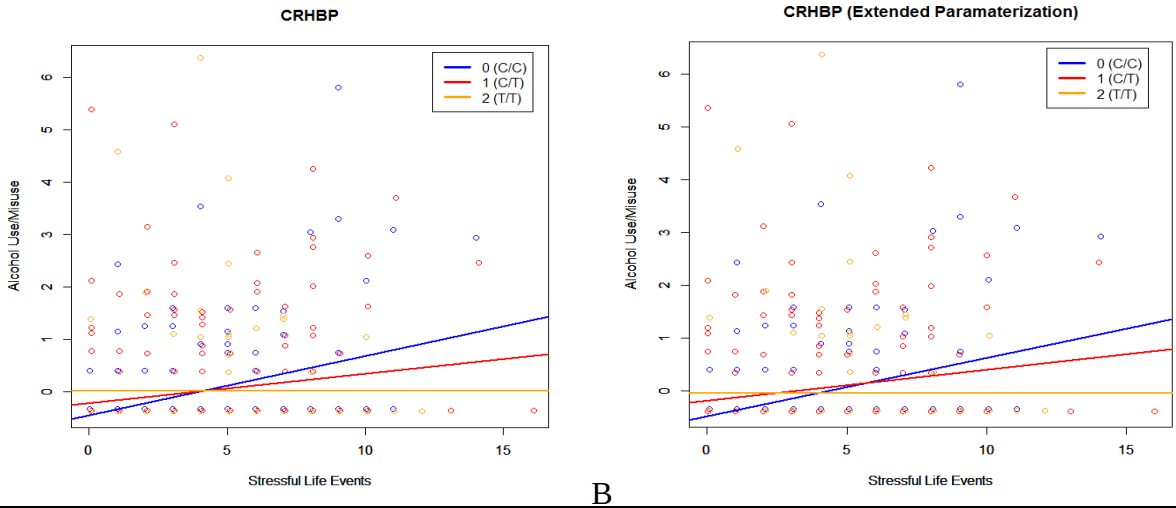
rs4912910	chr5:142787083	Intron 1	G	42.2	46.2	11.6	-0.09	0.17	-0.01	0.54
rs12655166	chr5:142789465	Intron 1	C	91.0	8.4	0.6	-0.03	0.84	0.04	0.49

[a] Covariates accounted for in calculating the main effect were child age and sex.

[b] Covariates accounted for in calculating the interaction effects were sex, child age, gene-specific SNP, and SLE. The SLE variable was centered on its mean for the analyses.

*$p \leq 0.007$ (Nyholt correction p-value for *CRHBP*).

Appendix A.3



A B

Figure 2. Regression Plots. Evidence of the interaction effects for the significant SNP, rs1715749, in *CRHBP* of the (A) normal regression model and in the (B) extended parameterization model, including a comparison of the predicted values based on the regression equation to the raw data. . The x-axis represents the raw values of SLE to aid in interpretation. However, mean centered values of SLE were use in the analyses.

Appendix A.4

Table 2. Moderated Multiple Regression of CRHBP and NR3C1 SNPs and Total Stressful Life Events Predicting Alcohol Use/Misuse in Individuals Endorsing Alcohol Use in the Past 12 Months

	SNP Position	Relative Position	Minor Allele	Percentage of Subjects with Copies of Minor Allele			Alcohol Use/Misuse			
							SNP Main Effect[a]		SNP x SLE Interaction[b]	
				0	1	2	B	*p*	B	*p*
SNP										
CRHBP										
rs1715752	chr5:76274929	Promoter	T	31.6	50.6	17.7	0.02	0.84	-0.08	0.04
rs1715749	chr5:76275603	Promoter	T	27.1	52.9	20.0	-0.01	0.95	-0.09	0.02*
rs32897	chr5:76286728	Intron 3	C	46.2	41.1	12.7	0.18	0.11	-0.08	0.06
rs7718461	chr5:76293804	Intron 5	A	60.4	32.7	6.9	-0.26	0.03	0.01	0.76
rs1053989	chr5:76300791	3' UTR	C	59.1	33.3	7.5	-0.26	0.03	0.01	0.77
rs2174444	chr5:76301278	Downstream	T	34.6	50.3	15.1	0.11	0.33	-0.06	0.14
rs10474485	chr5:76306609	Downstream	A	35.2	48.4	16.4	0.06	0.59	-0.03	0.43
rs1715747	chr5:76310293	Downstream	T	58.2	34.8	7.0	-0.07	0.56	0.02	0.69
NR3C1										
rs17209258	chr5:142653590	Intron 7	G	92.4	7.6	0.0	-0.22	0.47	0.1	0.3
rs10482682	chr5:142659590	Intron 5	T	77.1	19.7	3.2	-0.06	0.69	0.05	0.39
rs10482672	chr5:142672726	Intron 3	A	66.3	29.4	4.4	-0.24	0.08	-0.02	0.59
rs852980	chr5:142681049	Intron 2	G	28.2	59.0	12.8	-0.07	0.57	-0.04	0.44
rs2918418	chr5:142703566	Intron 2	C	69.8	27.7	2.5	0.04	0.8	-0.02	0.61
rs2963149	chr5:142705277	Intron 2	T	49.4	43.8	6.9	0.05	0.71	-0.05	0.22

rs4128428	chr5:142742006	Intron 2	C	82.9	14.6	2.5	-0.02	0.93	0.01	0.86
rs13182800	chr5:142781673	Intron 1	T	63.5	31.4	5.0	0.07	0.6	0.02	0.64
rs4912910	chr5:142787083	Intron 1	G	44.7	48.4	6.9	-0.01	0.93	-0.01	0.92
rs12655166	chr5:142789465	Intron 1	C	90.0	9.4	0.6	-0.14	0.56	0.07	0.42

[a]Covariates accounted for in calculating the main effect were child age and sex.

[b] Covariates accounted for in calculating the interaction effects were sex, child age, gene-specific SNP, and SLE. The SLE variable was centered on its mean for the analyses.

*$p \leq 0.05$

Appendix A.5

Table 3. Means and Standard Deviations for Items Contributing to Alcohol Use/Misuse Factor Score, Externalizing Behavior Composite Score, and Stressful Life Events

	Alcohol Use/Misuse Factor Score Items[a]			Externalizing Behaviors Composite Score	Stressful Life Events
	Frequency	Quantity	Binge Drinking[b]		
Mean (SD)	1.39 (0.96)	0.48(1.14)	1.17 (0.73)	11.42 (8.50)	3.80 (3.06)
Range	0-7	1-6	1-7	0-47	0-16

[a] The number of participants who reported ever use of alcohol was 163 (30.2%).

[b] Binge drinking defined as having 5 or more drinks in a row within a couple of hours.

Appendix A.6

Table 4. Regression of Polygenic Risk Scores Predicting Alcohol Use/Misuse in African American Adolescents [a]

	Intercept	B	*p*	R^2	R^2 without Main Effect	ΔR^2 [b]
p-value thresholds						
0.005	-2.9496	0.0000600	0.0019**	0.0533	0.04296	0.0103
0.01	-2.9499	0.0000596	0.002**	0.0532	0.04296	0.0103
0.05	-2.6345	0.0000185	0.0009**	0.0508	0.04296	0.0078
0.10	-2.6307	0.0000184	0.001**	0.0507	0.04296	0.0077
0.20	-2.4294	0.0000105	0.0037**	0.0491	0.04296	0.0061
0.30	-2.4043	0.0000088	0.0028**	0.0491	0.04296	0.0062
0.50	-2.3384	0.0000070	0.0059**	0.0485	0.04296	0.0055

[a] Covariates accounted for in calculating the main effect were child age and sex. The main effect of SLE on predicting alcohol use and misuse was significant (R^2=0.04, b= 0.06, p<0.001).

[b] ΔR^2 =R^2 (from the full model, including main effect, child age, and sex) - R^2 (from the regression calculated including only child age and sex).

*$p \leq 0.05$

** Bonferroni-adjusted significance level for multiple testing of $p \leq 0.001$ (Bonferroni p-value significance threshold of 0.05/42 was p<0.001).

Appendix A.7

Table 5. Regression of Polygenic Risk Scores and Stressful Life Events Predicting Alcohol Use/Misuse in African American Adolescents [a]

	Intercept	B	*p*	R^2	R^2 without Interaction Term	ΔR^2 [b]
p-value thresholds						
0.005	-3.1570	0.000016	0.026*	0.0936	0.0891	0.0045
0.01	-3.1589	0.000017	0.0235*	0.0937	0.08907	0.0047
0.05	-2.7973	0.000005	0.0188*	0.0910	0.08746	0.0035
0.10	-2.7930	0.000005	0.0196*	0.0908	0.08735	0.0035
0.20	-2.5746	0.000002	0.1762	0.0867	0.08534	0.0014
0.30	-2.5522	0.000002	0.1757	0.0865	0.08515	0.0014
0.50	-2.4904	0.000002	0.0997	0.0865	0.08445	0.002

[a] Covariates accounted for in calculating the interaction effects were sex, child age, p-value threshold specific polygenic risk score, and SLE. The SLE variable was centered on its mean for the analyses.
[b] $\Delta R^2 = R^2$ (from the full model including all covariates) - R^2 (from the full model excluding the interaction term).
*$p \leq 0.05$

Appendix A.8

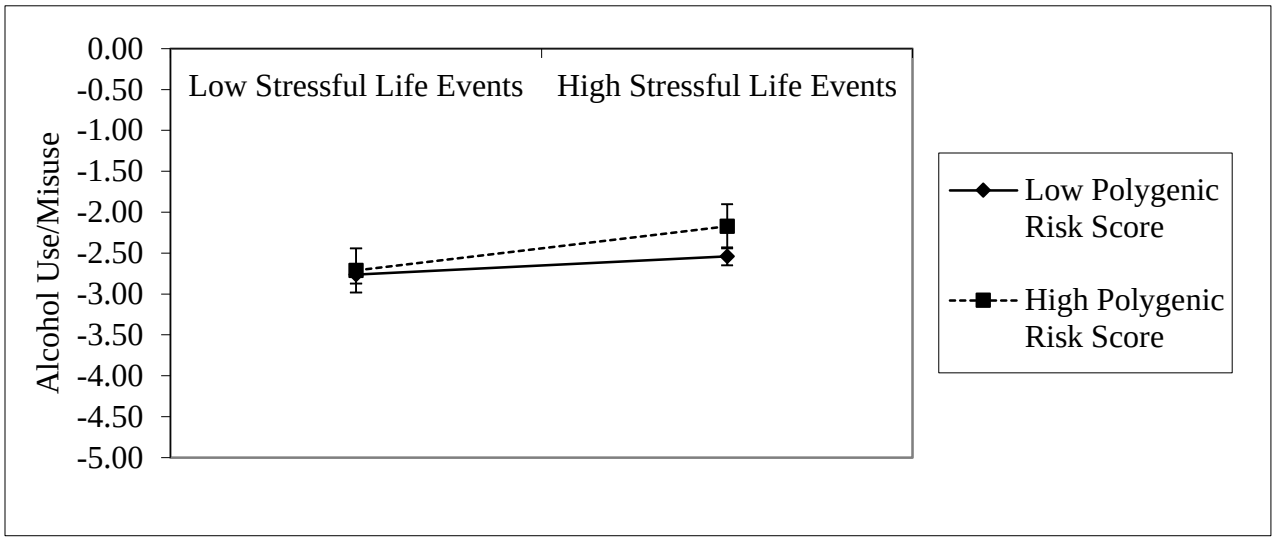


Figure 3. Significant interactions between polygenic risk score and SLE predicting alcohol use/misuse in African American adolescents at p-value threshold of $p \leq 0.01$.
Note: Moderated multiple regression analyses indicated significant interactions at p-value thresholds of 0.005, 0.05, and 0.10, and displayed the same shape of interaction. The interaction using the $p<0.01$ threshold accounted for the greatest proportion of variance.

Appendix A.9

Table 6. Regression of Polygenic Risk Scores Predicting Externalizing-Behavior Composite in African American Adolescents [a]

	Intercept	B	*p*	R^2	R^2 without Main Effect	ΔR^2 [b]
p-value thresholds						
0.005	8.8844	-0.000281	0.016*	0.0181	0.0121	0.0060
0.01	8.9189	-0.000283	0.015*	0.0181	0.0121	0.0060
0.05	7.1903	-0.000068	0.150	0.0165	0.0121	0.0044
0.10	7.1930	-0.000069	0.144	0.0166	0.0121	0.0045
0.20	6.4327	-0.000039	0.226	0.0162	0.0121	0.0041
0.30	6.3510	-0.000039	0.142	0.0166	0.0121	0.0045
0.50	6.0655	-0.000030	0.174	0.0164	0.0121	0.0043

[a] Covariates accounted for in calculating the main effect were child age and sex. The main effect of SLE on predicting externalizing behaviors was significant ($R^2 = 0.13$, b=1.01, p<.0001).

[b] $\Delta R^2 = R^2$ (from the full model, including main effect, child age, and sex) - R^2 (from the regression calculated including only child age and sex).

*$p \leq 0.05$

Table 7. Regression of Polygenic Risk Scores and Stressful Life Events Predicting Externalizing-Behavior Composite in African American Adolescents [a]

	Intercept	R^2	B	*p*	R^2 without Interaction Term	ΔR^2 [b]
p-value thresholds						
0.005	7.373	0.1491	0.000162	0.014*	0.1428	0.0063
0.01	7.399	0.1491	0.000161	0.015*	0.1428	0.0063
0.05	5.763	0.1529	0.000076	0.0004**	0.1405	0.0124
0.10	5.775	0.1526	7.49E-05	0.001**	0.1405	0.0121
0.20	5.386	0.1502	4.41E-05	0.004**	0.1405	0.0097
0.30	5.245	0.1505	3.71E-05	0.002**	0.1407	0.0098
0.50	4.997	0.1514	3.33E-05	0.001**	0.1406	0.0108

[a] Covariates accounted for in calculating the interaction effects were sex, child age, p-value threshold specific polygenic risk score, and SLE. The SLE variable was centered on its mean for the analyses.

[b] $\Delta R^2 = R^2$ (from the full model including all covariates) - R^2 (from the full model excluding the interaction term).

*$p \leq 0.05$

** Bonferroni-adjusted significance level for multiple testing of $p \leq 0.001$ (Bonferroni p-value significance threshold of 0.05/42 was $p<0.001$).

Appendix A.11

Table 8. Polygenic Risk Scores from a European American Sample Predicting Alcohol Use/Misuse and Externalizing Behavior Composite in African American Adolescents

	Polygenic Score Main Effect Predicting Alcohol Use/Misuse[a]			Polygenic Score Main Effect predicting Externalizing Behavior Composite		
	R^2	B	*p*	R^2	B	*p*
p-value thresholds						
0.005	0.002	-0.0000117	0.520	0.003	0.000029	0.864
0.01	0.004	-0.0000355	0.018*	0.003	-0.000054	0.674
0.05	0.008	-0.0000287	0.005*	0.004	-0.000054	0.413
0.10	0.004	-0.0000148	0.041*	0.003	-0.000015	0.801
0.20	0.002	-0.0000056	0.281	0.003	-0.000004	0.940
0.30	0.002	-0.0000045	0.351	0.003	-0.000010	0.848
0.50	0.001	-0.0000028	0.533	0.003	0.000011	0.812

[a] Covariates accounted for in calculating the main effect were child age and sex.
*$p \leq 0.05$

Appendix A.12

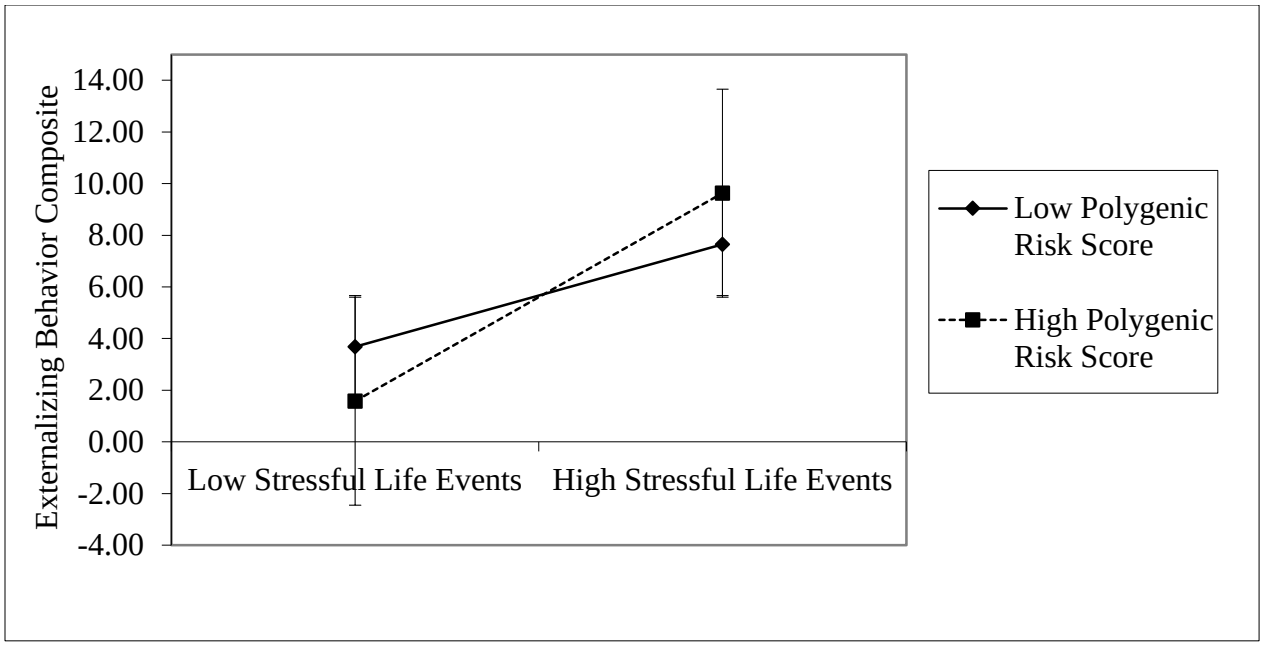


Figure 4. Significant interaction between polygenic risk score and SLE predicting externalizing–behavior composite at p-value threshold of $p \leq 0.05$.
Note: Moderated multiple regression analyses indicated significant interactions using polygenic risk scores computed across all SNP p-value thresholds. The interaction using the $p<0.05$ threshold accounted for the greatest proportion of variance. The nature of the interactions from all other p-value thresholds indicated a similar cross-over effect. Illustrative low and high values (1 *SD* above and below the mean) for polygenic scores and stressful life events are shown.