2012

Elucidating Genetic and Environmental Influences on Alcohol-Related Phenotypes

Jacquelyn Meyers
Virginia Commonwealth University

Elucidating Genetic and Environmental Influences on

Alcohol-Related Phenotypes

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy at Virginia Commonwealth University

by

Jacquelyn Leigh Meyers, B.S.

Department of Human and Molecular Genetics

Virginia Institute of Psychiatric and Behavioral Genetics

Director: Danielle M. Dick, Ph.D.

Department of Psychiatry

Department of Psychology

Department of Human and Molecular Genetics

Virginia Institute of Psychiatric and Behavioral Genetics

Virginia Commonwealth University

Richmond, Virginia

May, 2012

ACKNOWLEDGEMENT

The author wishes to thank several people that have supported her in preparation for this thesis. First, she would like to thank her advisor Danielle M. Dick for her endless patience, guidance and support. Next, she would like to acknowledge the faculty, post-doctoral trainees, graduate students and staff at the Virginia Institute for Psychiatric and Behavioral Genetics and the Department of Human Genetics for making the past five years of living and training in Richmond a pleasure. Finally, she would like to thank her friends and family for their continuous love and encouragement. She attributes all of her success to both her genes and environment.

TABLE OF CONTENTS

TABLES

FIGURES

List of Figures

ABSTRACT

ELUCIDATATING GENETIC AND ENVIRONMENTAL INFLUENCES ON ALCOHOL RELATED PHENOTYPES

by Jacquelyn L. Meyers, B.S.

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University. 2012.

Major Director: Danielle M. Dick, Ph.D. Psychiatry, Psychology, & Human and Molecular Genetics

Decades of work has led researchers to believe that risk for complex behavioral phenotypes, such as alcohol use disorders, is likely influenced by multiple genes of small effect acting in conjunction with each other and the environment. Currently, the field of psychiatric genetics is developing methodologies for the identification of genetic risk variants that predispose individuals to the development of complex behavioral disorders. Several challenges related to the complex and polygenic nature of these phenotypes, must be considered. This dissertation study attempts to address these important challenges in the context of alcohol use disorders and related phenotypes. A rich twin and family study literature has indicated that 40-70% of the variance in alcohol use disorders (AUDs) is influenced by genetics. Recent attempts to identify specific

genetic risk variants associated with AUDs have been met with limited success. Meanwhile, evidence of the moderating effects of the environment on AUDs has been mounting, providing a strong rationale for examining gene-environment interaction. In the following chapters several studies will be described that integrate established twin methodologies into gene identification projects in an effort to reduce heterogeneity (both phenotypic and genotypic), elucidate environmental constructs that moderate genetic influences, and to enhance statistical power to detect the subtle genetic influences on alcohol related phenotypes.

GLOBAL INTRODUCTION

Evidence supporting significant heritability for a variety of psychiatric and behavioral disorders has led to considerable efforts to identify the specific genes involved. Behavioral disorders are complex genetic traits that are both clinically and genetically heterogeneous. It is expected that there are multiple genetic loci influencing the manifestation of and variation in these behaviors, and that these loci vary in the direction and magnitude of their effects. Further complicating the search for the biological basis of complex disorders is the influence of the environment, varying in importance throughout development. Although disorders such as alcohol dependence are clearly influenced by genetic components, the dissection of these disorders is more complicated than that originally mapped out by single gene traits. Several challenges related to the complex and polygenic nature of these phenotypes, including statistical power and heterogeneity, must be considered. This dissertation study attempts to address these important challenges in the context of alcohol use disorders and related phenotypes. The first aim of this dissertation study is to conduct a series of twin analyses aimed at understanding the genetic architecture across alcohol consumption and problems. The second aim of this study is to elucidate environments that mask or exacerbate the genetic influence on alcohol phenotypes. The final aim of this study is to identify genetic risk variants for alcohol consumption and problems.

Alcohol Dependence

Genetic studies of alcohol dependence provide an excellent example of the challenges posed by complex behavioral and psychiatric disorders. There are a variety of societal problems, such as job loss and the deconstruction of families, which arise from alcohol use and related behavioral disorders (Kriegbaum et al., 2011), so there is great demand for research in this area. Decades of twin and family studies have demonstrated that there are critical genetic and environmental components in the inheritance of substance use disorders (Kaprio et al., 1987; Heath et al., 1991; McGue et al., 1992; Kendler et al., 1994; Prescott et al., 2001; Ystrom et al., 2011) and modern advances in genetics are making it possible to identify specific variants that may predispose an individual to these disorders. We now know that there is no "gene for alcoholism" but rather a multitude of genes, each with subtle effects. These genes are likely to interact epistatically with each other as well as with their biological and external environments to make an individual more susceptible to the development of these complex disorders. As our understanding of substance use becomes more refined, we see that dependence has a complex development that starts with initiation of use (Dawson et al., 2008).

Twin studies provide an estimation of a trait's heritability in a population; that is, what proportion of phenotypic variation is due to genetic variation underlying the trait. Twin studies accomplish this by comparing phenotypic similarity between monozygotic twins, who share all of their genetic variation, with dizygotic twins, who share (on average) half of their genetic variation. Measures of heritability are a function of the specific population. Heritability estimates of substance use disorders are likely to vary among substances (and the measure of substance use), populations, age, and sex. A 2005 meta-analysis of twin studies has shown that the heritability of all addictive substances ranges from 40% to 60% (Goldman et al., 2005). A recently published large male twin study, reported that after accounting for errors of measurement, the heritability of lifetime history of AD increased from 55 to 71% (Ystrom et al., 2011).

Alcohol dependence is a phenotypically and genetically heterogeneous disorder. DSM-IV (American Psychological Association, 1994) alcohol dependence is currently diagnosed by the presence of any three of the following seven criteria: (1) tolerance; (2) withdrawal; (3) taking the substance in larger amounts than intended; (4) persistent desire or unsuccessful efforts to cut down on the substance; (5) spending a great deal of time obtaining or recovering from the effects of the substance; (6) giving up important recreational, social, or occupational activities as a result of the substance; and (7) continued use of the substance despite physical or psychological problems caused by the substance. These alcohol dependence criteria represent a diversity of physiological and societal consequences of alcohol use. It would seem likely that (1) tolerance and (2) withdrawal may represent a more physiological response to alcohol and employ a host of alcohol metabolism genes, while (6) giving up important recreational, social, or occupational activities as a result of the substance may represent more psychological behavioral disinhibition, which may employ a different set of genes. Cohesive categories of symptoms designed to represent the disorder have been created for the purpose of characterizing disorders and developing a successful treatment plan. However, recent twin studies (Kendler et al., 2012) provide support that our biology does not necessarily respect these same categories. Further, the use of the DSM alcohol dependence diagnosis in gene finding studies creates a research design which tests if one gene is associated with seven heterogeneous symptoms. Recently, quantitative measures of alcohol consumption and problems have gained more attention. Several twin studies (Whitfield et al., 2008, Grant et al., 2009, Kendler et al, 2010, Dick et al., 2011) have examined the relationship between quantitative measures of alcohol consumption (frequency of use, frequency of intoxication, maximum drinks in a 24-hour-period) and problems (DSM AD, Rutgers Alcohol Problem Index, Michigan Alcohol Screening Test). While the results from these studies provide varying estimates of genetic correlation, as a set they suggest that there is both shared and unique genetic liability for alcohol consumption and problems. In addition, large gene finding projects are beginning to utilize quantitative measures of alcohol consumption (Schumann et al., 2011, Baik et al., 2011). Several chapters in this dissertation will utilize alternative, biologically informed, quantitative measures of consumption and problem drinking, to test hypotheses related to the etiology of alcohol dependence.

The Externalizing Spectrum

Epidemiologic studies find that individuals rarely abuse a single substance (Swendsen et al., 2012). Instead, polysubstance abuse and dependence is normative, with high rates of comorbidity across various drug classes. In addition, individuals with substance use disorders also exhibit higher rates of other behavioral disorders (Slutske et al., 1998, Krueger et al., 2001; Krueger et al., 2002, Krueger et al., 2005; Hasin et al., 2011). Twin studies suggest that this comorbidity is due at least in part to a shared genetic etiology underlying susceptibility to different types of substance use and other psychopathologies (Kendler et al., 2003, Hicks et al., 2004, Kendler et al., 2011, Hicks et al., 2011). In 2003, Kendler and colleagues used the Virginia Twin Registry sample to identify common genetic factors underlying substance use disorders and externalizing/internalizing behavioral disorders (eg, conduct disorder, generalized anxiety disorder), and found that one common genetic factor accounted for 34% of the variance in alcohol dependence and 42% of the variance in abuse/dependence on other drugs (Kendler et al., 2003). This factor also loaded onto adult antisocial behavior and conduct disorder. These results suggest a common genetic factor for both substance dependence/abuse and general externalizing psychopathologies.

A number of other studies (Kendler et al., 2006, Dick et al., 2010, Dick et al., 2011, Edwards et al., 2012) lend further support to the premise that shared genetic factors influence externalizing disorders. Kendler's 2006 study also reported that a latent externalizing factor, constructed of measures of conduct disorder, adult antisocial behavior, alcohol and drug abuse/dependence, and disinhibitory personality traits, is highly heritable (80%-85%) (Kendler et al., 2006). Thus, this latent externalizing factor appears to be more heritable than the individual disorders themselves, which show individual heritabilities of approximately 50% (Goldman et al., 2005). A final piece of evidence suggesting a shared genetic liability across externalizing psychopathology comes from the electrophysiological literature in which a number of electrophysiological endophenotypes thought to represent markers of genetic vulnerability are shared across the spectrum of externalizing disorders, including alcohol dependence, other forms of substance dependence, childhood externalizing disorders, and adult antisocial personality disorder (Dick et al., 2005; Gilmore et al., 2010). In summary, there has been much evidence to suggest that adolescent externalizing behavior (including drug, alcohol, and behavior problems) may be an early manifestation of risk to a spectrum of externalizing disorders (Dick et al., 2008). Thus, to consider each of these disorders in isolation may lead us to miss important etiological clues. This early indication of genetic risk for adult alcohol problems can be exploited in longitudinal samples that assess behavior problems and drinking behavior from adolescence into adulthood. Several chapters in this dissertation will utilize longitudinal reports of adolescent behavior problems and alcohol consumption.

Identification of Specific Genes Influencing Complex Traits

Candidate genes may be chosen based on our knowledge of their involvement in specific biological pathways or systems. For example, genes that are part of the dopaminergic system are considered candidate genes for drug addiction, at least in part because of the role of dopamine in the reward pathway. Early studies focusing on functional candidates (e.g., ALDH2, ADH1B) for alcohol related phenotypes were quite successful (Gelernter & Kranzler, 2009). The influence of genetic polymorphisms at loci encoding acetaldehyde and alcohol dehydrogenases on risk for AD in specific populations is well established, and the mechanism tractable. Alcohol is metabolized to acetaldehyde, a toxic intermediary, by alcohol dehydrogenases; acetaldehyde is metabolized primarily by acetaldehyde dehydrogenases, the most relevant of which is encoded by ALDH2. Acetaldehyde produces a “flushing reaction” characterized by a set of uncomfortable symptoms including flushing of the skin, lightheadedness, palpitations, and nausea. A variant that reduces or eliminates ALDH function (occurring mostly in Asian populations) is protective against AD (because clearance of acetaldehyde is impeded), and ADH variants that increase function (and the production of acetaldehyde) may also be protective (Thomasson et al. 1991; Hasin et al. 2002; Konishi et al. 2003). A meta-analysis (Luczak et al. 2006) showed that subjects heterozygous for a null ALDH2 allele have only about one-fourth the risk for alcohol dependence as those with two functional alleles.

Candidate genes also arise from previous implications of involvement with a trait from the linkage literature. Two different regions of chromosome 4 have been implicated in genome-wide linkage scans for alcohol risk variants. These two regions include an ADH gene cluster, which maps to the long arm of chromosome 4, and a GABAA receptor subunit gene cluster, which maps to the short arm of the chromosome. ADH4 (Luo et al. 2005a, b, 2006; Edenberg et al. 2006) is one of several disease-influencing loci in this cluster. Edenberg et al. (1999) demonstrated that the ¡75A allele, at a promoter polymorphic site in ADH4, has promoter activity that is more than twice that of the ¡75C allele (Luo et al. 2005 a, b, Luo et al. 2006). Other candidate genes from this region that are implicated in alcohol related phenotypes include ADH2 (Luczak et al., 2006), GABRA2 (Edenberg et al., 2006, Covault et al., 2004, Fehr et al., 2006), and GABRG1 (Ittiwut et al., 2008; Covault et al., 2008; Enoch et al., 2009). Other candidate genes initially implicated by linkage studies include the muscarinic acetycholine receptor M2, CHRM2 (Wang et al., 2004), a class of opioid receptors OPRM1 (Luo et al., 2003; Zhang et al. 2006), OPRD1, OPRK1 (Gelernter et al., 2007), and the dopamine receptor, DRD2 (Blum et al., 1991), which is likely related to the effects observed with ANKK1, NCAM1, and TTC12 (Neville et al., 2004).

While the candidate gene strategy has been successful in a number of studies, it is largely limited by the scope of our understanding of human biology. The technological advances that have made it feasible to genotype genome-wide representative SNPs via SNP chips (Illumina/Affymetrix), has made the advent of genome wide association studies a solution to some of the limitations of the candidate gene approach. The genome wide approach has created a more agnostic study design that scans a large number of individual genomes and provides a genetic comparison of affected cases to unaffected controls. This strategy removes the biases of a priori gene selection that is driven by previous implication in the literature, and creates a design for identifying novel genetic variants involved in human behavior and disease. While this study design has great potential for success, there are a number of challenges that it creates. In 2007, The Wellcome Trust Case-Control consortium published a collaborative study that examined 2,000 cases of seven common complex diseases and a shared set of 3,000 controls in a general population in the United Kingdom[2]. Of the seven diseases studied, the most prolific results came for Crohn's disease (9 SNPs) and Type I Diabetes (7 SNPs), the least prolific results came from Hypertension (0 SNPs), Bipolar disorder (1 SNP) and coronary-artery disease (1 SNP). One of the questions posed by the field was what contributed to the limited success of hypertension and Bipolar Disorder, two of the most common health concerns examined in this study.

In their 1996 paper, Risch and Merikangas (Risch & Merikangas, 1996) detail the statistical power issues that genome wide research provides us with. Extraordinarily large sample sizes are required to detect the subtle genetic variants that we believe to be underlying complex genetic traits. With odds ratios on the order of 1-1.5, complex traits are in sharp contrast Mendelian traits with large odds-ratios. One of the possible explanations for the failure to detect genetic variants for hypertension, bipolar and coronary-artery disease, is that more subjects are required to detect statistically significant variants. Another possible explanation for the failure to detect genetic variants is the control sample. One consequence of using a shared control group (for which detailed phenotyping for all traits of interest is not available) relates to the potential for misclassification bias: a proportion of the controls is likely to have the disease of interest and therefore might meet the criteria for inclusion as a case (and some others will develop it in the future). If 5% of controls meet the definition of cases at the same age, the loss of power is approximately the same as that due to a reduction of the sample size by 10%. This is particularly relevant with hypertension and coronary artery disease, for which it is estimated that 30% of the population is affected. Genomic association is contingent upon an empirical measure of the phenotype. Hypertension is a chronic medical condition in which an individual's blood pressure is elevated. In this study, Hypertension was defined by blood pressure over 140 mmHg, where normal blood pressure ranges between 90 and 119 mmHg. Pre-hypertension ranged between 120 and 140 mmHg. A binary definition status forces an arbitrary cut-off value of a continuous measure measurement, in this case being blood pressure. This creates a loss of power both in discarding useful data on “borderline” individuals and by creating a potentially inaccurate definition of control subjects, who may have some of the common genetic variants involved in blood pressure levels. The DSM-IV is the primary diagnostic system used by clinicians and in many genetic studies of psychiatric disorders, including Bipolar Disorder. The use of a standardized DSM criterion has many advantages including (1) decades of research focused on the reliability and validity of measures, (2) convenience of a standard measure that is widely used and therefore conducive to collaborative efforts as well as the potential for (3) direct comparison to achieve replication. This is especially useful in large-scale genetic efforts, where multiple sites are often needed to collect the required number of affected families to achieve reasonable power to detect genes in association with complex traits. While the uses of DSM diagnosis provide advantages, many have argued that they are not ideal for genetic studies. The stated priority of the DSM[1] is to “provide a helpful guide to clinical practice” (DSM-IV, p. xv), with a secondary goal of facilitating research. While the DSM’s primary goal is clinical utility, its application in research has become a standard. These diagnoses are based on patterns of human behavior and are not necessarily biologically informed. Therefore, it may be more appropriate to use measures that are biologically informed when searching for genetic variants associated with complex human disease.

Despite the analytic challenges that conducting GWAS on alcohol dependence poses, multiple GWAS of alcohol related phenotypes are now underway. In 2009, the first genome wide association study (GWAS) on alcohol dependence (AD) was published (Treutlein et al., 2009). This study included 487 German male inpatients with alcohol dependence as defined by the DSM-IV and an age at onset younger than 28 years, and 1,358 population-based control individuals. This study also included a follow-up sample of 1,024 German male inpatients and 996 age-matched male controls. This initial GWAS implicated two novel intergenic single nucleotide polymorphisms (SNPs) that reached stringent genome wide significance thresholds required to correct for multiple testing (rs7590720, rs1344694). Since then, several alcohol dependence GWAS have been reported and are detailed in Chapter 3. From 2010-2011, six large GWA studies were published (Lind et al., 2010, Bierut et al., 2010, Edenberg et al., 2010, Kendler et al., 2011, Heath et al., 2011, Wang et al, 2011), none of which reported genome wide significant findings. Thus far, two very large alcohol dependence GWAS have been published in 2012 (Zuo et al., 2012, Frank et al., 2012), both of which have reported genome wide significant findings. Earlier this year, Zuo and colleagues combined the Study of Addiction Genetics and Environment (SAGE) data and Australian family study of alcohol use disorder (OZ-ALC) with the goal of discovering the novel risk loci for alcohol dependence. The authors reported that variants within *KIAA0040* and the *PHF3-PTP4A1* gene complex might harbor a causal variant for AD (Zuo et al., 2012). Frank and colleagues (Frank et al., 2012) conducted an AD GWAS on 1,333 German (inpatient) cases and 2,168 German controls and reported genome-wide significant support for the role of the ADH gene cluster (*ADH1B/ADH1C*). In addition to these AD GWAS reports, several studies have conducted association with alcohol-related phenotypes, such as alcohol consumption. Many studies have suggested that use of a quantitative measure could improve power to detect variants of small effect (Agrawal et al., 2009). In 2010, Joslyn and colleagues conducted a GWAS on level of response to alcohol in 367 individuals and reported no genome wide significant findings. However in 2011, two large studies conducted GWAS on alcohol consumption (Baik et al., 2011, Schumann et al., 2011) and reported genome wide significant findings. Baik and colleagues reported genome wide significant signals in (or near) *C12orf51, CCDC63*, and *MYL2* that were successfully replicated in a sample of Korean male drinkers; rs2074356, located in *C12orf51*, was in high linkage disequilibrium with SNPs in *ALDH2*, but other SNPs were not (Baik et al., 2011). *ALDH2* met genome-wide significance in an alcohol consumption GWAS in a Japanese population based sample (Takeuchi et al., 2011). The Collaborative Study on the Genetics of Alcoholism (COGA) has reported associations with alcohol withdrawal symptoms in *KDM4C* (Wang et al., 2011b). The largest alcohol related GWAS to date examined alcohol consumption in 12 population-based samples of European ancestry, comprising 26,316 individuals, with replication genotyping in an additional 21,185 individuals. SNP rs6943555 in autism susceptibility candidate 2 gene (*AUTS2*) was associated with alcohol consumption at a genome-wide significant level (Schumann et al., 2011). Most recently, Agrawal and colleagues conducted a GWAS on alcohol craving in 3,976 individuals and reported no genome wide significant findings (Agrawal et al., 2012).

In reviewing the current state of alcohol dependence GWAS findings, fewer than half of the published studies report genome-wide significant findings. At this point, evidence that the genome-wide significant variants implicated in these studies replicate in an independent sample is limited. However, there is some suggestion from this literature that larger sample sizes and quantitative measures of alcohol use may increase the likelihood (via an increase in statistical power) of identifying genome wide significant findings. Several chapters in this dissertation will utilize quantitative measures of alcohol use and problems.

Gene-Environment Interaction

There is an emerging literature documenting how specific environmental factors moderate the importance of genetic effects. A growing number of variables have been shown to moderate the relative importance of genetic effects on substance use and dependence and externalizing behavior. Among the environmental moderators being studied are childhood stressors (emotional, physical, and sexual abuse), availability and access to drugs and alcohol, peer-group antisocial and prosocial behavior, religiosity, parental attitudes toward drugs and alcohol, parental monitoring, and socioregional factors. Religiosity has been shown to moderate genetic influences on alcohol use among females, with genetic factors playing a larger role among individuals without a religious upbringing (Koopmans et al., 1998). Social contact and cotwin dependency have also been shown to moderate twin similarity, with reduced genetic effects and enhanced environmental influences among more codependent pairs (Penninkilampi et al., 2005). Genetic influences on adolescent substance use are also enhanced in environments with lower parental monitoring (Dick et al., 2007). These analyses suggest that when adolescents receive little parental monitoring, it creates an environment that allows for greater opportunity to express genetic predispositions. The moderating effects of peer alcohol use on adolescent drinking has been shown to operate in a similar fashion: among adolescents with a larger number of peers who used alcohol, there was greater expression of genetic predispositions (Dick et al., 2007). These findings may reflect a situation in which environments characterized by low parental monitoring or high peer substance use create opportunity for adolescents to express genetic predispositions. These results support previous findings from the Finnish Twin Studies, which indicated that in neighborhoods in which there is less stability, presumably engendering less community monitoring, there was greater evidence of genetic influence (Rose et al., 2003; Dick et al., 2009). Conversely, in more supervised and restricted environments, there was less opportunity to express genetic predispositions and greater influence of environmental effects. Hicks and colleagues examined the specificity of each of these environmental risk factors on externalizing spectrum disorders, including substance dependence/abuse (Hicks et al., 2009). They concluded that, in the context of environmental adversity, broadly defined, genetic factors become more important in the etiology of externalizing disorders. In addition, their results suggest a general mechanism of environmental influence on externalizing disorders, regardless of the specific form of environmental risk.

These analyses illustrate the importance of incorporating measured aspects of the environment into genetically informative twin models to understand how specific environments act and interact with genetic predispositions. They may also have implications for studying the risk associated with specific genes. For example, a 2009 study aimed to characterize the pathway of risk associated with *GABRA2*, a gene previously associated with adult alcohol dependence, in a community sample of children followed longitudinally from childhood to young adulthood (Dick et al. 2009). Association between *GABRA2* and trajectories of externalizing behavior was tested from adolescence to young adulthood and moderation of genetic effects by parental monitoring was also tested. Two classes of externalizing behavior emerged: a stable, high externalizing class and a moderate, decreasing externalizing-behavior class. The *GABRA2* gene was associated with class membership, with subjects who showed persistent increased trajectories of externalizing behavior more likely to carry the genotype previously associated with increased risk of adult alcohol dependence. A significant interaction with parental monitoring emerged; the association of *GABRA2* with externalizing trajectories diminished with high levels of parental monitoring. In the last decade, candidate-gene x environment studies have received much attention, both positive and negative. Most notorious was Caspi's report that the serotonin transporter (5-HTT) gene moderated the influence of stressful life events on depression (Caspi et al., 2003). This initial report was followed by a plethora of candidate-gene x environment studies producing mixed results and a largely un-interpretable literature. A recent review by Duncan and Keller suggested that most positive candidate-gene x environment findings are false-positives, resulting from low power along with publication bias (Duncan & Keller, 2011).

In summary, decades of research has led researchers to the assumption that risk for complex behavioral phenotypes, such as alcohol use disorders, is likely influenced by multiple genes of small effect acting in conjunction with the environment. Currently, the field of psychiatric genetics is developing effective methodologies for the identification of genetic risk variants that predispose individuals to the development of complex behavioral disorders. Several challenges related to the complex and polygenic nature of these phenotypes, including statistical power and heterogeneity, must be considered. This dissertation study attempts to address these important challenges in the context of alcohol use disorders and related phenotypes. A rich twin and family study literature has indicated that 40-70% (Goldman et al., 2005; Ystrom et al., 2011) of the variance in Alcohol Use Disorders (AUDs) is influenced by genetics. Recent attempts to identify specific genetic risk variants associated with AUDs have been met with limited success. Meanwhile, evidence of the moderating effects of the environment on AUDs has been mounting providing a strong rationale for examining gene-environment interaction. In the following chapters several studies will be described that integrate established twin methodologies into gene identification projects in an effort to reduce heterogeneity, both phenotypic and genotypic, elucidate environmental constructs that moderate genetic influences, and to enhance statistical power to detect the subtle genetic influences on AUDs.

The first aim of this dissertation study is to conduct a series of twin analyses aimed at understanding the genetic architecture across alcohol consumption and problems. The second aim of this study is to elucidate environments that mask or exacerbate the genetic influence on alcohol phenotypes. The final aim of this study is to identify genetic risk variants for alcohol consumption and problems. In the following chapters, I will describe several studies that seek to address these research aims (for each study, the chapter, title, research design, alcohol outcome and age are described below in table 1). In the first chapter of this dissertation, I will describe a study that examined the genetic architecture across several measures of young adult (~age 22) alcohol consumption and problems using twin methodology. In the following chapters, I will describe two studies that put the information gained from this twin study into use in genetic association studies, first with a candidate gene (chapter 2) and then on a genome-wide level (chapter 3). I will then go on to describe three studies that examine gene-environment interaction across development, first using twin methodology (chapter 4) to examine whether three environments moderate the genetic influences on adolescent drinking frequency (ages 14 and 17.5), the second following up on these effects using polygene scores derived from GWAS data (chapter 5), and the third examining weather these gene-environment interaction effects observed in adolescence remain relevant in young adulthood (~age 22) (chapter 6). Finally, I will conclude by describing a study that examines the relevance of genetic influences on alcohol consumption across adolescent development and into young adulthood (chapter 7).

Table 1. Summary of Dissertation Studies

Chapter (*Aim*)	**Study**	**Study Design**	**Outcome**	**Age**
I *(1)*	Measures of Current Alcohol Consumption and Problems: Two Independent Twin Studies Suggest A Complex Genetic Architecture	Twin Study	Alcohol Consumption and Problems	22
II *(3)*	The Association between *DRD2* and Genetically Informed Measures of Alcohol Use and Problems	Genetic Association	Alcohol Consumption and Problems	25
III *(3)*	Finntwin12 GWAS of Alcohol Consumption and Problems	GWAS	Alcohol Consumption and Problems	22
IV *(2)*	Environmental Moderation of Alcohol Use and Behavior Problems in Adolescence: Specificity versus Generality of Environmental Risk Factors	Twin Study	Alcohol Consumption	14, 17.5
V *(2)*	Life Events Moderate Genetic and Environmental Influences on Adolescent Externalizing Disorders	Twin Study and Polygene Score x Environment	Alcohol Consumption	14, 17.5
VI *(2)*	The Interaction between Parental Knowledge in Adolescence and Genetic Risk for Alcohol Dependence Predicts Adult Alcohol Dependence	Twin Study and Polygene Score x Environment	Alcohol Consumption	14, 17.5, 22
VII *(1)*	Genetic Risk for Alcohol and Externalizing Problems across Time	Twin Study	Alcohol Consumption	14, 17.5, 22

References

1. Kriegbaum M, Christensen U, Osler M, Lund R. Excessive drinking and history of unemployment and cohabitation in Danish men born in 1953. Eur J Public Health. 2011 Aug;21(4):444-8.
2. Kaprio et al., 1987; Heath et al., 1991; McGue et al., 1992; Kendler et al., 1994; Kendler et al., 2001; Ystrom et al., 2011
3. Kaprio J, Koskenvuo M, Langinvainio H, Romanov K, Sarna S, Rose RJ. Genetic influences on use and abuse of alcohol: a study of 5638 adult Finnish twin brothers. Alcohol Clin Exp Res. 1987 Aug;11(4):349-56.
4. Heath AC, Meyer J, Jardine R, Martin NG. The inheritance of alcohol consumption patterns in a general population twin sample: II. Determinants of consumption frequency and quantity consumed. J Stud Alcohol. 1991Sep;52(5):425-33.
5. McGue M, Pickens RW, Svikis DS. Sex and age effects on the inheritance of alcohol problems: a twin study. J Abnorm Psychol. 1992 Feb;101(1):3-17.
6. Kendler KS, Neale MC, Heath AC, Kessler RC, Eaves LJ. A twin-family study of alcoholism in women. Am J Psychiatry. 1994 May;151(5):707-15.
7. Prescott CA, Kendler KS. Influence of ascertainment strategy on finding sex differences in genetic estimates from twin studies of alcoholism. Am J Med Genet. 2000 Dec 4;96(6):754-61.
8. Ystrom E, Reichborn-Kjennerud T, Aggen SH, Kendler KS. Alcohol dependence in men: reliability and heritability. Alcohol Clin Exp Res. 2011 Sep;35(9):1716-22.

9. Dawson DA, Goldstein RB, Chou SP, Ruan WJ, Grant BF. Age at first drink and the first incidence of adult-onset DSM-IV alcohol use disorders. Alcohol Clin Exp Res. 2008 Dec;32(12):2149-60.

10. Goldman D, Oroszi G, Ducci F. The genetics of addictions: Uncovering the genes. *Nat Rev Genet* 2005;6(7):521-32.

11. Schumann G, Coin LJ, Lourdusamy A, Charoen P, Berger KH, Stacey D, Desrivières S, Aliev FA, Khan AA, Amin N, Aulchenko YS, Bakalkin G, Bakker SJ, Balkau B, Beulens JW, Bilbao A, de Boer RA, Beury D, Bots ML, Breetvelt EJ, Cauchi S, Cavalcanti-Proença C, Chambers JC, Clarke TK, Dahmen N, de Geus EJ, Dick D, Ducci F, Easton A, Edenberg HJ, Esko T, Fernández-Medarde A, Foroud T, Freimer NB, Girault JA, Grobbee DE, Guarrera S, Gudbjartsson DF, Hartikainen AL, Heath AC, Hesselbrock V, Hofman A, Hottenga JJ, Isohanni MK, Kaprio J, Khaw KT, Kuehnel B, Laitinen J, Lobbens S, Luan J, Mangino M, Maroteaux M, Matullo G, McCarthy MI, Mueller C, Navis G, Numans ME, Núñez A, Nyholt DR, Onland-Moret CN, Oostra BA, O'Reilly PF, Palkovits M, Penninx BW, Polidoro S, Pouta A, Prokopenko I, Ricceri F, Santos E, Smit JH, Soranzo N, Song K, Sovio U, Stumvoll M, Surakk I, Thorgeirsson TE, Thorsteinsdottir U, Troakes C, Tyrfingsson T, Tönjes A, Uiterwaal CS, Uitterlinden AG, van der Harst P, van der Schouw YT, Staehlin O, Vogelzangs N, Vollenweider P, Waeber G, Wareham NJ, Waterworth DM, Whitfield JB, Wichmann EH, Willemsen G, Witteman JC, Yuan X, Zhai G, Zhao JH, Zhang W, Martin NG, Metspalu A, Doering A, Scott J, Spector TD, Loos RJ, Boomsma DI, Mooser V, Peltonen L, Stefansson K, van Duijn CM, Vineis P, Sommer WH, Kooner JS, Spanagel R,

Heberlein UA, Jarvelin MR, Elliott P. Genome-wide association and genetic functional studies identify autism susceptibility candidate 2 gene (AUTS2) in the regulation of alcohol consumption. Proc Natl Acad Sci U S A. 2011 Apr 26;108(17):7119-24.

12. Baik I, Cho NH, Kim SH, Han BG, Shin C. Genome-wide association studies identify genetic loci related to alcohol consumption in Korean men. Am J Clin Nutr. 2011 Apr;93(4):809-16.
13. Swendsen J, Burstein M, Case B, Conway KP, Dierker L, He J, Merikangas KR. Use and abuse of alcohol and illicit drugs in US adolescents: results of the national comorbidity survey-adolescent supplement. Arch Gen Psychiatry. 2012 Apr;69(4):390-8.
14. Slutske WS, Heath AC, Dinwiddie SH, et al. Common genetic risk factors for conduct disorder and alcohol dependence. *J Abnorm Psychol* 1998;107(3):363-74.
15. Krueger RF, Hicks BM, Patrick CJ, et al. Etiologic connections among substance dependence, antisocial behavior, and personality: Modeling the externalizing spectrum. *J Abnorm Psychol* 2002;111(3):411-24.
16. Krueger, R. F. (1999). The structure of common mental dis- orders. *Archives of General Psychiatry, 56,* 921–926.
17. Krueger, R. F., Markon, K. E., Patrick, C. J., & Iacono, W. (2005). Externalizing psychopathology in adulthood: A dimensional-spectrum conceptualization and its implica- tions for DSM-V. *Journal of Abnormal Psycholgy, 114,* 537–550.
18. Krueger, R. F., McGue, M., & Iacono, W. G. (2001). The higher-order structure of

common DSM mental disorders: Internalization, externalization, and their connections to personality. *Personality and Individual Differences, 30,* 1245–1259.

19. Hasin D, Fenton MC, Skodol A, Krueger R, Keyes K, Geier T, Greenstein E, Blanco C, Grant B. Personality disorders and the 3-year course of alcohol, drug, and nicotine use disorders. Arch Gen Psychiatry. 2011 Nov;68(11):1158-67.
20. Hicks, B. M., Krueger, R. F., Iacono, W. G., McGue, M., & Patrick, C. J. (2004). Family transmission and heritability of externalizing disorders: A twin-family study. *Archives of General Psychiatry, 61,* 922–928.
21. Hicks, B. M., Schalet, B. D., Malone, S. M., Iacono, W. G., & McGue, M. (2011). Psychometric and genetic architecture of substance use disorder and behavioral disinhibition measures for gene association studies. *Behavior Genetics, 41,* 459–475.
22. Kendler, K. S., Aggen, S., Knudsen, G. P., Roysamb, E., Neale, M., & Reichborn-Kjennerud, T. (2011a). The structure of genetic and environmental risk factors for syndromal and subsyndromal common DSM-IV Axis I and all Axis II personality disorders. *American Journal of Psychiatry, 168,* 29–39.
23. Kendler, K. S., Gatz, M., Gardner, C., & Pedersen, N. (2006). A Swedish National Twin Study of Lifetime Major Depression. *American Journal of Psychiatry, 163,* 109–114.
24. Kendler, K. S., Myers, J. M., Maes, H. H., & Keyes, C. L. (2011b). The relationship between the genetic and environ- mental influences on common internalizing psychiatric disorders and mental well-being. *Behavior Genetics.*

25. Kendler, K. S., Neale, M. C., Sullivan, P., Corey, L. A., Gardner, C. O., & Prescott, C. A. (1999). A population- based twin study in women of smoking initiation and nicotine dependence. *Psychological Medicine, 29,* 299–308.
26. Kendler, K. S., Prescott, C. A., Myers, J., & Neale, M. C. (2003). The structure of genetic and environmental risk factors for common psychiatric and substance use disor- ders in men and women. *Archives of General Psychiatry, 60,* 929–937.
27. Dick DM, Meyers JL, Latendresse SJ, Creemers HE, Lansford JE, Pettit GS, Bates JE, Dodge KA, Budde J, Goate A, Buitelaar JK, Ormel J, Verhulst FC, Huizink AC. CHRM2, parental monitoring, and adolescent externalizing behavior: evidence for gene-environment interaction. Psychol Sci. 2011 Apr;22(4):481-9.
28. Dick DM, Meyers J, Aliev F, Nurnberger J Jr, Kramer J, Kuperman S, Porjesz B, Tischfield J, Edenberg HJ, Foroud T, Schuckit M, Goate A, Hesselbrock V, Bierut L. Evidence for genes on chromosome 2 contributing to alcohol dependence with conduct disorder and suicide attempts. Am J Med Genet B Neuropsychiatr Genet. 2010 Sep;153B(6):1179-88.
29. Dick DM, Aliev F, Wang JC, Grucza RA, Schuckit M, Kuperman S, Kramer J, Hinrichs A, Bertelsen S, Budde JP, Hesselbrock V, Porjesz B, Edenberg HJ, Bierut LJ, Goate A. Using dimensional models of externalizing psychopathology to aid in gene identification. Arch Gen Psychiatry. 2008 Mar;65(3):310-8.
30. Gelernter J, Kranzler HR. Genetics of alcohol dependence. Hum Genet. 2009 Jul;126(1):91-9.

31. Thomasson HR, Edenberg HJ, Crabb DW, Mai XL, Jerome RE, Li TK, Wang SP, Lin YT, Lu RB, Yin SJ (1991) Alcohol and aldehyde dehydrogenase genotypes and alcoholism in Chinese men. Am J Hum Genet 48:677–681
32. Hasin D, Aharonovich E, Liu X, Mamman Z, Matseoane K, Carr L, Li TK (2002) Alcohol and ADH2 in Israel: Ashkenazis, Sephardics, and recent Russian immigrants. Am J Psychiatry 159:1432–1434
33. Konishi T, Smith JL, Lin KM, Wan YJ (2003) InXuence of genetic admixture on polymorphisms of alcohol-metabolizing enzymes: analyses of mutations on the CYP2E1, ADH2, ADH3 and ALDH2 genes in a Mexican-American population living in the Los Angeles area. Alcohol Alcohol 38:93–94
34. Luczak SE, Glatt SJ, Wall TL (2006) Meta-analysis of ALDH2 and ADH1B with alcohol dependence in Asians. Psychol Bull 132:607–621
35. Luo X, Kranzler HR, Zhao H, Gelernter J (2003) Haplotypes at the OPRM1 locus are associated with susceptibility to substance dependence in European Americans. Am J Med Genet B Neuro- psychiatr Genet 120B:97–108
36. Luo X, Kranzler HR, Zuo L, Yang BZ, Lappalainen J, Gelernter J (2005a) ADH4 gene variation is associated with alcohol and drug dependence in European Americans: results from family-con- trolled and population-structured association studies. Pharmaco- genet Genomics 15:755–768
37. Luo X, Kranzler HR, Zuo L, Lappalainen J, Yang BZ, Gelernter J (2005b) ADH4 gene variation is associated with alcohol and drug dependence in European Americans: results from HWD tests and case-control association studies. Neuropsychopharmacology 31:1085–1095

38. Luo X, Kranzler HR, Zuo L, Blumburg H, Wang S, Gelernter J (2005c) *CHRM2* gene predisposes to alcohol dependence, drug depen- dence, and aVective disorders: results from an extended case–con- trol structured association study. Hum Mol Genet 14:2421–2434

39. Luo X, Kranzler HR, Zuo L, Wang S, Schork NJ, Gelernter J (2006) Diplotype trend regression analysis of the ADH gene cluster and the ALDH2 gene: multiple signiWcant associations with alcohol dependence. Am J Hum Genet 78:973–987

40. Edenberg HJ, Kranzler HR (2005) The contribution of genetics to addiction therapy approaches. Pharmacol Ther 108:86–93

41. Edenberg HJ, Jerome RE, Li M (1999) Polymorphism of the human alcohol dehydrogenase 4 (ADH4) promoter aVects gene expres- sion. Pharmacogenetics 9:25–30

42. Edenberg HJ, Dick DM, Xuei X, Tian H, Almasy L, Bauer LO, Crowe RR, Goate A, Hesselbrock V, Jones K, Kwon J, Li T-K, Nurnberger JI Jr, O'Connor SJ, Reich T, Rice M, Schuckit MA, Porjesz B, Foroud T, Begleiter H (2004) Variations in GABRA2, encoding the □2 subunit of the GABAA receptor, are associated with alcohol dependence and with brain oscillations. Am J Hum Genet 74:705–714

43. Edenberg HJ, Xuei X, Chen HJ, Tian H, Wetherill LF, Dick DM, Almasy L, Bierut L, Bucholz KK, Goate A, Hesselbrock V, Kuperman S, Nurnberger J, Porjesz B, Rice J, Schuckit M, TischWeld J, Begleiter H, Foroud T (2006) Association of

alcohol dehydrogenase genes with alcohol dependence: a comprehensive analysis. Hum Mol Genet 15:1539–1549

44. Covault J, Gelernter J, Jensen K, Anton R, Kranzler HR (2008) Mark- ers in the 5′-region of *GABRG1* associate to alcohol dependence and are in linkage disequilibrium with markers in the adjacent *GABRA2* gene. Neuropsychopharmacology 33(4):837–848
45. Enoch MA, Schwartz L, Albaugh B, Virkkunen M, Goldman D (2006) Dimensional anxiety mediates linkage of GABRA2 haplotypes with alcoholism. Am J Med Genet B Neuropsychiatr Genet 141B(6):599–607
46. Enoch MA, Hodgkinson CA, Yuan Q, Albaugh B, Virkkunen M, Goldman D (2009) GABRG1 and GABRA2 as independent pre- dictors for alcoholism in two populations. Neuropsychopharma- cology 34:1245–1254
47. Fehr C, Sander T, Tadic A, Lenzen KP, Anghelescu I, Klawe C, Dahmen N, Schmidt LG, Szegedi A (2006) ConWrmation of association of the GABRA2 gene with alcohol dependence by subtype-speciWc analysis. Psychiatr Genet 16:9–17
48. Gelernter J, Yu Y, Weiss R, Brady K, Panhuysen C, Yang BZ, Kranzler HR, Farrer L (2006) Haplotype spanning TTC12 and ANKK1, Xanked by the DRD2 and NCAM1 loci, is strongly associated to nicotine dependence in two distinct American populations. Hum Mol Genet 15:3498–3507. doi:10.1093/hmg/ddl426
49. Gelernter J, Gueorguieva R, Kranzler HR, Zhang H, Cramer J, Rosen- heck R, Krystal JH, the V A Cooperative Study #425 Study Group (2007) Opioid receptor gene (OPRM1, OPRK1, and OPRD1) variants and response to naltrexone

treatment for alcohol depen- dence: results from the VA Cooperative Study. Alcohol Clin Exp Res 31(4):555–563

50. Gelernter J, Kranzler HR, Panhuysen C, Weiss RD, Brady K, Poling J, Farrer L (2009) Dense genomewide linkage scan for alcohol dependence in African-Americans: signiWcant linkage on chro- mosome 10. Biol Psychiatry 65:111–115

51. Wang JC, Hinrichs AL, Stock H, Budde J, Allen R, Bertelsen S, Kwon JM, Wu W, Dick DM, Rice J, Jones K, Nurnberger JI Jr, TischWeld J, Porjesz B, Edenberg HJ, Hesselbrock V, Crowe R, Schuckit M, Begleiter H, Reich T, Goate AM, Bierut LJ (2004) Evidence of common and speciWc genetic eVects: association of the muscarinic acetylcholine receptor M2 (CHRM2) gene with alcohol dependence and major depressive syndrome. Hum Mol Genet 13:1903–1911

52. Wang JC, Grucza R, Cruchaga C, Hinrichs AL, Bertelsen S, Budde JP, Fox L, Goldstein E, Reyes O, Saccone N, Saccone S, Xuei X, Bucholz K, Kuperman S, Nurnberger J Jr, Rice JP, Schuckit M, TischWeld J, Hesselbrock V, Porjesz B, Edenberg HJ, Bierut LJ, Goate AM (2008) Genetic variation in the *CHRNA5* gene aVects mRNA levels and is associated with risk for alcohol dependence. Mol Psychiatry 14:501–510

53. Neville MJ, Johnstone EC, Walton RT (2004) IdentiWcation and characterization of ANKK1: a novel kinase gene closely linked to DRD2 on chromosome band 11q23.1. Hum Mutat 23:540–545

54. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007 Jun 7;447(7145):661-78.

55. Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science. 1996 Sep 13;273(5281):1516-7.
56. Zuo L, Zhang CK, Wang F, Li CS, Zhao H, Lu L, Zhang XY, Lu L, Zhang H, Zhang
57. F, Krystal JH, Luo X. A novel, functional and replicable risk gene region for alcohol dependence identified by genome-wide association study. PLoS One. 2011;6(11):e26726.
58. Wang KS, Liu X, Aragam N, Jian X, Mullersman JE, Liu Y, Pan Y. Family-based association analysis of alcohol dependence in the COGA sample and replication in the Australian twin-family study. J Neural Transm. 2011 Sep;118(9):1293-9.
59. Frank J, Cichon S, Treutlein J, Ridinger M, Mattheisen M, Hoffmann P, Herms S, Wodarz N, Soyka M, Zill P, Maier W, Mössner R, Gaebel W, Dahmen N, Scherbaum N, Schmäl C, Steffens M, Lucae S, Ising M, Müller-Myhsok B, Nöthen MM, Mann K, Kiefer F, Rietschel M. Genome-wide significant association between alcohol dependence and a variant in the ADH gene cluster. Addict Biol. 2012 Jan;17(1):171-80. doi: 10.1111/j.1369-1600.2011.00395.x.
60. Rose RJ, Dick DM, Viken RJ, et al. Drinking or abstaining at age 14? A genetic epidemiological study. *Alcohol Clin Exp Res* 2001;25(11):1594-1604.
61. Rose RJ, Dick DM, Viken RJ, et al. Gene-environment interaction in patterns of adolescent drinking: Regional residency moderates longitudinal influences on alcohol use. *Alcohol Clin Exp Res* 2001;25(5):637-43.
62. Rose RJ, Viken RJ, Dick DM, et al. It does take a village: Nonfamilial environments and children’s behavior. *Psychological Science* 2003; 14(3):273–

77.

63. Koopmans JR, Slutske WS, Van Baal GC, et al. The influence of religion on alcohol use initiation: Evidence for genotype X environment interaction. *Behav Genet* 1999;29(6):445-53.

64. Rose RJ, Kaprio J, Williams CJ, et al. Social contact and sibling similarity: Facts, issues, and red herrings*. Behav Genet* 1990;20(6):763-78.

65. Dick DM, Viken R, Purcell S, et al. Parental monitoring moderates the importance of genetic and environmental influences on adolescent smoking. *J Abnorm Psychol* 2007;116(1):213-18.

66. Marshal MP, Chassin L. Peer influence on adolescent alcohol use: The moderating role of parental support and discipline. *Applied Developmental Science* 2000;4:80–8.

67. Dick DM, Rose RJ, Viken RJ, et al. Exploring gene-environment interactions: Socioregional moderation of alcohol use. *J Abnorm Psychol* 2001;110(4):625-32.

68. Hicks BM, South SC, Dirago AC, et al. Environmental adversity and increasing genetic risk for externalizing disorders. *Arch Gen Psychiatry* 2009;66(6):640-48.

69. Dick DM, Latendresse SJ, Lansford JE, et al. Role of GABRA2 in trajectories of externalizing behavior across development and evidence of moderation by parental monitoring. *Arch Gen Psychiatry* 2009;66(6):649-57.

70. Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW, Harrington H, McClay J, Mill J, Martin J, Braithwaite A, Poulton R. Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. Science. 2003 Jul 18;301(5631):386-9.

71. Duncan LE, Keller MC. A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. Am J Psychiatry. 2011 Oct;168(10):1041-9. Epub 2011 Sep 2. Review.

Chapters 1-7

Chapter 1

Measures of Current Alcohol Consumption and Problems: Two Independent Twin Studies Suggest A Complex Genetic Architecture

****This chapter is adapted from the following manuscript:***
Dick DM, Meyers JL, Rose R, Kaprio J, Kendler K.S. Measures of Current Alcohol Consumption and Problems: Two Independent Twin Studies Suggest A Complex Genetic Architecture. Alcohol Clin Exp Res. 2011 Dec;35(12):2152-61

Abstract

Background: Twin studies demonstrate that measures of alcohol consumption show evidence of genetic influence, suggesting they may be useful in gene identification efforts. The extent to which these phenotypes will be informative in identifying susceptibility genes involved in alcohol dependence depends on the extent to which genetic influences are shared across measures of alcohol consumption and alcohol problems. Previous studies have demonstrated that alcohol consumption reported for the period of heaviest lifetime drinking shows a large degree of genetic overlap with alcohol dependence; however, many studies with genetic material assess current alcohol consumption. Further, there are many different aspects of alcohol consumption that can be assessed (e.g., frequency of use, quantity of use, frequency of intoxication, etc).

Methods: Here we use data from two large, independent, population-based twin samples, *Finntwin16* and *The Virginia Adult Twin Study of Psychiatric and Substance Use Disorders*, to examine the extent to which genetic influences are shared across many different measures of alcohol consumption and alcohol problems.

Results: Genetic correlations across current alcohol consumption measures and alcohol problems were high across both samples. However, both samples suggest a complex genetic architecture with many different genetic factors influencing various aspects of current alcohol consumption and problems.

Conclusions: These results suggest that careful attention must be paid to the phenotype in efforts to "replicate" genetic effects across samples or combine samples for meta-analyses of genetic effects influencing susceptibility to alcohol-related outcomes.

Introduction

Alcohol dependence is under substantial genetic influence (Dick et al. 2009), and twin studies demonstrate that measures of alcohol consumption (AC) are under significant genetic influence as well (Dick and Bierut, 2006; Goldman, 1993; Prescott and Kendler, 1999; Rose, 1998). That evidence has fostered studies investigating the extent to which the same genetic factors underlie patterns of consumption and the development of problems. Data from the Australian twin registry indicated moderate correlations (r=0.42 for females and r=0.45 for males) between genetic influences on weekly alcohol consumption and lifetime alcohol problems, and between heavy drinking and alcohol dependence (r=.63) (Heath and Martin, 1994). More recently, Grant and colleagues found a genetic correlation of .97 between a composite alcohol consumption factor score, comprised of drinking measures from the period of heaviest use, and alcohol dependence symptoms (Grant et al., 2009). Similarly, Kendler and colleagues, using data from the Virginia Twin Study of Adult Psychiatric and Substance Use Disorders, found complete overlap between the genetic risk for alcohol dependence and four measures of alcohol consumption at the time of heaviest intake in females; in men, the consumption measures captured 85% of the genetic risk for dependence (Kendler et al., 2010). Both studies concluded that the high genetic overlap between consumption and alcohol dependence suggests that continuous consumption measures may be useful in the discovery of genes contributing to dependence risk.

The extent to which genetic influences on alcohol dependence are shared with genetic influences on measures of alcohol consumption has important implications for gene identification efforts. It is more practical to collect information on alcohol consumption from large samples of individuals than to recruit alcohol dependent probands and appropriate controls and assess psychiatric diagnoses. Measures of alcohol consumption also have attractive statistical properties because analyzing quantitative traits can improve power in association analyses (Agrawal et al., 2009). While a small number of studies are underway with the express purpose of identifying genes involved in alcohol dependence (Edenberg et al., 2005; Prescott et al., 2005), many projects with genetic material have collected data on alcohol consumption, making it possible to use existing datasets for gene identification, replication, and/or meta-analyses. However, the relevance of these findings for understanding predispositions to develop alcohol-related problems hinges on the extent to which genes associated with measures of alcohol consumption also relate to alcohol problems.

One critical aspect that has not been widely addressed in this burgeoning literature is the fact that there are many different ways to assess "alcohol consumption", reflecting the many different aspects and facets of drinking patterns. For example, in the studies reviewed above, measures of alcohol consumption included frequency (weekly and annually), quantity by frequency, maximum drinks in a 24-hour period, frequency of heavy drinking (5+ drinks), and frequency of intoxication. The most recent studies (Grant et al., 2009; Kendler et al., 2010) addressing genetic overlap have used measures of alcohol consumption at the heaviest point of drinking. However, many studies assess current alcohol consumption, rather than lifetime consumption patterns. Here, we use data from two twin studies to conduct an exploratory set of analyses examining the extent to which different measures of past year alcohol consumption share genetic overlap with various indices of alcohol problems. We test the extent to which genetic influences are shared across different measures of consumption, and between these different consumption measures and measures of alcohol related problems.

Methods

FinnTwin16 (FT16)

FT16 is a population-based study consisting of five consecutive birth cohorts of Finnish twins. All twins were identified through Finland's Population Register Center, permitting exhaustive and unbiased ascertainment. Zygosity was determined using a well-validated questionnaire completed by both co-twins at the baseline, as described elsewhere (Kaprio et al., 1991). FT16 consists of twins born 1975-1979 (Kaprio et al., 2002). The five birth cohorts contained 3065 families of twins in which both twins were living and residing in Finland at the age of 16. Details about data collection have previously been published (Kaprio, 2006; Kaprio et al., 2002). Briefly, four waves of postal questionnaires were completed at ages 16, 17, 18.5, and as young adults. Here we analyze data from the most recent questionnaire and focus on alcohol consumption and alcohol problems in adulthood. The average age for the respondent twins at this assessment was 24.4 years (SD=1.50, range 22.8- to 27.2), with a response rate of 88.1%. For ease of presentation, this assessment is referred to as age 25 throughout this paper. Parallel to current practice in gene identification efforts for alcohol dependence, only individuals who had evidence of alcohol exposure were included in twin analyses, so that genetic and environmental influences on the decision to initiate alcohol are not confounded with genetic and environmental influences on alcohol consumption or problems. After exclusion of individuals who had not been exposed to alcohol, data were available for 685 complete pairs of twin brothers (287 MZ and 398 DZ), and 693 complete pairs of twin sisters (378 MZ and 315 DZ).

Measures

Frequency was assessed with the following question: “At the present, how often do you drink alcohol?’’ Response options included: (1) I don’t use alcohol; (2) Once or year or less frequently; (3) 3-4 times a year; (4) About once in two months; (5) About once a month; (6) A couple times a month; (7) About once a week; (8) About twice a week; (9) Daily. Note that responses were reverse-coded from the actual order asked so that higher numbers reflected more drinking across all items used in analyses.

Frequency x Quantity was a composite of two items; the frequency of reported alcohol use in the past 28 days multiplied by the quantity of drinks (drinks defined as 1 beer, 1 glass of wine, or 1 mixed drink containing hard liquor) consumed per drinking day during the past 28 days. Because this measure was highly skewed, with over representation of those who drank on less than one occasion in the past 28 days, we log-transformed this variable.

Frequency of Heavy Drinking was assessed with the following question: “At the present, how often do you within one occasion use more than five bottles of beer, or more than a bottle of wine, or more than half a bottle of hard liquor?” Response options included:

(1) I don't use alcohol; (2) Never; (3) Once or year or less frequently; (4) 3-4 times a year; (5) About once in two months; (6) About once a month; (7) A couple times a month; (8) About once a week; (9) About twice a week; (10) Daily.

Frequency of Intoxication was assessed with the following question: "At the present, how often do you use alcohol to get drunk?" Response options included: (1) I don't use alcohol/Never; (2) Once or year or less frequently; (3) 3-4 times a year; (4) About once in two months; (5) About once a month; (6) A couple times a month; (7) About once a week; (8) About twice a week; (9) Daily.

Maximum Drinks (Max Drinks) was the maximum number of drinks twins reported ever consuming in a 24 hour period, with 1 drink defined as 1 beer, 1 glass of wine, or 1 mixed drink containing hard liquor. Responses ranged from 1-100 (mean= 16.49, SD=9.46).

The Malmo-modified Michigan Alcohol Screening Test (Mm-MAST;(Kristenson and Trell, 1982)) is a 9-item self-report scale of current drinking patterns and problems designed for application in Nordic cultures (Seppa et al., 1999). Representative items include taking a drink before going to a party, increased tolerance over time, and having difficulty not drinking more than one's friends. Our scale added two items more directly overlapping DSM diagnostic criteria: finding it hard to stop after having had a drink and feeling that someone close to you thinks you should drink less. Each of these questions was asked of "current and past drinking habits" and had a "Yes" or "No" response option. For those twins who answered at least 9 of the 11 items, we calculated a MmMAST score by taking the average response (yes/no) across the number of items answered. This scoring method permitted us to retain participants who completed the majority of the items but who may have neglected to answer a few of them.

Alcohol Problem Index (RAPI) is a reliable 22 item scale designed to assess problematic drinking (White and Labouvie, 1989). The RAPI contains items assessing dependence, withdrawal, blackouts, neglect of responsibilities in several domains, shame and/or embarrassment to self or others, and inappropriate behaviors such as fighting. Individuals indicated how often each consequence of alcohol use had happened in the past twelve months using the following five response options: (1) Never/I don't use alcohol, (2) Rarely, (3) Sometimes, or (4) Quite often. For subjects who answered at least 18 of the 22 items, we calculated a RAPI severity score by taking the average response (1-4) across the number of items answered.

Because of the limitations of the genetic statistical analysis program, we were unable to simultaneously analyze both continuous and ordinal variables; thus, we collapsed the drinking measures into four categories (once individuals who had indicated that they do not use alcohol were removed). An alcoholic drink was defined as "<u>one</u> bottle of beer, <u>one</u> glass of wine or <u>one</u> shot of liquor" across all questions. For drinking frequency, frequency of heavy drinking, and frequency of intoxication, these categories were (1) About 1- 4 times a year, (2) About once in two months, (3) About 1-2 times a month, (4) About 1-2 times a week. Maximum Drinks, the MmMAST, and RAPI scores were each collapsed into five levels using the SAS System's univariate quintiles procedure, where the first level contains those individuals lowest on problem drinking and the fifth level contains those highest on problem drinking (SAS, 2002-2003).

Virginia Adult Twin Study of Psychiatric and Substance Use Disorders (VATSPSUD)

Participants in this study derive from two inter-related studies of Caucasian same-sex twin pairs who participated in VATSPSUD (Kendler, 2006). All subjects for the VATSPSUD were ascertained from the population-based Virginia Twin Registry formed from a systematic review of birth certificates in the Commonwealth of Virginia. Female-female twin pairs (FF), from birth years 1934-1974, became eligible if both members previously responded to a mailed questionnaire in 1987-1988, the response rate to which was approximately 64%. Zygosity was determined by discriminate function analyses using standard twin questions validated against DNA genotyping in 496 pairs (Kendler and Prescott, 1999). All female-female data on AC and AD used in this report were collected at the fourth wave of interviews (FF4), conducted in 1995-1997. For this wave, we succeeded in interviewing 85% of the sample who had responded to the previous questionnaire. Data on the male-male (MM) pairs, birth years 1940-1974, came from a sample initially ascertained directly from registry records, which contained all twin births. The first interview (MM1) was completed largely by phone in 1993-1996 and obtained a 72% response rate. This was followed by a second wave of interviews (MM2), conducted in 1994-1998 with a follow up response rate of 83%. Data on AC and AD were collected at both of these waves. We used the measures of drink frequency, regular quantity, maximum quantity and AD from MM1 because of the larger sample size, but frequency of intoxication was only assessed at MM2 and so those data were used. The mean (SD) age of the twins was 36.3 (8.2) at the FF4 interview and 35.5 (9.1) at the MM1 interview. Note, that the FT16 sample is age standardized (~age 25) and differs in this sense from the wide age range covered in the VATSPSUD sample. The VATSPSUD alcohol section began by asking about any lifetime alcohol use. In our FF4, MM1 and MM2 interviews, 8.0, 5.0 and 4.3% of participants respectively denied any lifetime alcohol use and were excluded from all subsequent analyses. After excluding abstainers, the total sample size on which we had data for AC and AD was 5,073 and consisted of 1,766 complete pairs and 893 twins whose cotwins did not participate. By zygosity, the numbers of complete pairs were: monozygotic (MZ) male twins 613; dizygotic (DZ) male 435; MZ female 440 and DZ female 278.

Measures

Frequency was assessed by the following question: “In a typical month over the last year, how often do you drink alcohol?” Response options included: (1) 1-3, (2) 4-9, (3) 10-15, (4) 16-27 and (5) 28-30 days per month.

Regular Quantity was assessed with the following question on drinking habits in the past year: “On those days when you drank, how many drinks did you usually have in a day?” Response options included: (1) 1-2, (2) 3, (3) 4-5, (4) 6-9 and (5) ≥ 9 drinks/day.

Frequency of Intoxication was assessed with the following question: “During the past year, how often did you use alcohol to get drunk?” Response options were: (1) 1-2, (2) 3-5, (3) 6-7, (4) 8 and (5) 9-11 times/year.

Maximum Drinks was assessed with the following question: “What is the largest number of drinks you had on any single day during the past year?” Response options were: (1) 1-5, (2) 6-9, (3) 10-12, (4)13-20, and (5) ≥ 21 drinks/day.

DSM-IV AD Symptoms were assessed for *lifetime* in the interviews based on seven DSM-IV criteria (American Psychological Association, 1994), and was the only VATSPSUD measure that did not reflect current alcohol problems.

Multivariate Cholesky

A multivariate Cholesky model was used to estimate genetic and environmental influences across the measures of consumption/problem drinking (Neale and Cardon, 1992). Analyses were conducted separately using the measures available in each sample. The Cholesky model allows us to evaluate (1) the magnitude of genetic and environmental influences on each phenotype and (2) the extent to which these influences contribute to the covariation between the phenotypes. Phenotypic variance was decomposed into three components: variance due to additive genetic factors (a2); variance due to shared environmental factors (c2); and variance due to non-shared environmental, or individual-specific, factors (e2). Calculation of variance accounted for by each of these factors is performed by comparing monozygotic twin correlations to dizygotic twin correlations. Genetic influences correlate 1.0 between monozygotic (MZ) twins, who share all of their genetic variation identical-by-descent, and 0.5 between dizygotic (DZ) twins, who share, on average, 50% of their segregating genes, as do ordinary siblings. Common/shared environmental effects, as defined in biometrical twin modeling, refer to all environmental influences that make siblings more similar to one another. By definition, these influences correlate 1.0 between both MZ and DZ twins. Unique/nonshared environmental influences are uncorrelated between co-twins and have the effect of decreasing the covariance between siblings. When data on multiple phenotypes are available, these models can be extended to evaluate the extent to which genetic and environmental contributions to the disorders are shared. This is calculated by comparing cross-twin, cross-trait correlations, with the logic extended from the basic twin model that comparison of the cross-twin, cross-trait correlations between MZs and DZs provides information about the extent to which a2, c2, and e2 contribute to the phenotypic correlations between traits.

The full model (depicted in Figure 1 for Finntwin16 and Figure 2 for the VATSPSUD) calculated variance components separately by sex. Thresholds for each variable were adjusted by age to account for the variability in age in the samples. Additional models were tested to evaluate goodness-of-fit in which estimates of the variance components were constrained to be equal across sex. Estimates were obtained from observed twin data using maximum likelihood estimation in the software program Mx (Neale et al., 1999). Model fit was evaluated by Akaike's Information Criterion (AIC), and the probability (p) value associated with the χ^2 statistic. Lower AIC values indicate an optimal balance between explanatory power and parsimony. Additionally, nonsignificant χ^2 values (p >. 05) indicate a good fit. We compared nested alternative models by the change in chi-square between models, which is used to evaluate the significance of dropping parameters. A significant change in χ^2 ($p < .05$) for the difference in degrees of freedom of the models indicates that the model with fewer degrees of freedom should be adopted, because the gain in degrees of freedom of the alternate model caused a significant decrease in fit. Missing data were handled by reading raw data into Mx and fitting to the observed and unobserved data vectors using full information maximum likelihood estimation.

Results

FinnTwin16

Table 2 details the phenotypic correlations across the different measures of alcohol consumption and problem drinking. Polychoric correlations were computed on only one twin from each pair, chosen randomly. Table 3 shows the MZ and DZ twin correlations for each of the measures. The results of the series of models fit are shown in Table 4.

Table 2. FinnTwin16 Phenotypic Correlations between Measures of Alcohol Consumption and Problems

Measure	**Freq**	**Freq x Quant**	**Freq of Heavy**	**Freq of Intox**	**Max Drinks**	**MAST**	**RAPI**
Frequency	1						
Freq x Quant	.77	1					
Freq Heavy	.73	.79	1				
Freq Intox	.73	.80	.91	1			
Max Drinks	.46	.53	.56	.53	1		
MAST	.33	.41	.44	.45	.39	1	
RAPI	.23	.31	.34	.35	.26	.47	1

Note: all correlations significant at p<0.001

Table 3. FinnTwin16 MZ and DZ Correlations between Measures of Alcohol Consumption and Problems

Measure	**MZ Females**	**DZ Females**	**MZ Males**	**DZ Males**
Frequency	.59	.43	.75	.47
Freq x Quant	.45	.30	.61	.37
Freq Heavy	.54	.34	.64	.42
Freq Intox	.64	.38	.65	.45
Max Drinks	.55	.35	.65	.29
MAST	.55	.34	.63	.52
RAPI	.43	.23	.52	.25

Note: all correlations significant at p<0.001

We initially fit a full Cholesky model including full A, C, and E matrices separately for each sex (AIC=5967.906, DF=16618) (Model I in Table 3). Next we tested a model in which we constrained all parameters to be equal in males and females (Model II). The AIC decreased and the χ^2 change was non-significant for the change in degrees of freedom, indicating that the more parsimonious model constraining males and females to be equal provided a better fit. We next tested a model including full A and E matrices, and dropping the full C Matrix representative of all shared environmental influences (Model III). The AIC decreased and the χ^2 change was nonsignificant for the change in degrees of freedom, indicating that the more parsimonious model dropping all shared environmental influences on the measures provided a better fit. Models IV – VI are submodels that test for a reduced number of genetic factors. We systematically tested the significance of each genetic factor and each pathway in the following sequence: (1) tested the significance of the entire A matrix; (2) tested the significance of each latent genetic factor; (3) tested the significance of each individual genetic pathway. Each of the pathways retained in the Best Fitting Model is by definition significant.

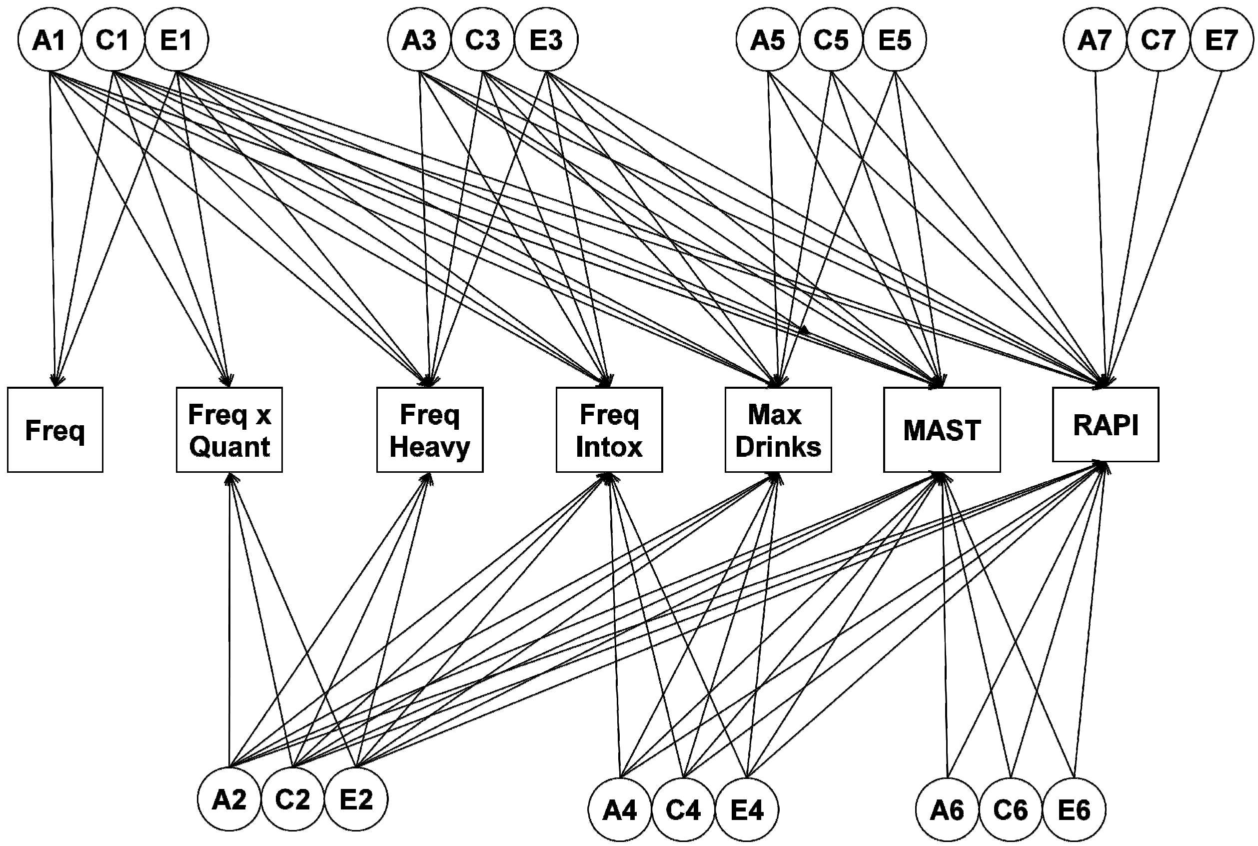
Table 4. FinnTwin16 Model Fitting Results

Model			**Δ Fit**			
		Compared to Model	**Δ X^2**	**Probability**	**Δ DF**	**Δ AIC**
I*	Full Model	--	--	--	--	--
II	Sexes equated	I	16.60	0.96	84	39.39
III	C Matrix dropped	II	60.05	0.98	28	107.95
IV	A1	III	337.39	0.00	21	+127.39
V	A1 + A2	III	216.36	0.00	15	+18.36
VI	A1 + A2 + A3	III	145.48	0.00	10	+145.48
VII^	**A1 + A2 + A3 + A4**	**III**	**111.60**	**0.12**	**6**	**78.36**

Fit of Model I: − 2LL = 39203.91, df = 16618, AIC = 5967.91

^ Best fitting model.

Figure 1. FinnTwin16 Full Twin Model



Model IV allows for only one latent genetic factor (A1 in Figure 1), Model V allows for two latent genetic factors (A1 and A2), and Model VI allows for three latent genetic factors (A1, A2, and A3). For each of these submodels, the AIC increased and the χ^2 change was significant for the change in degrees of freedom, indicating that these models provided a worse fit to the data. The best-fitting model (Model VII; shown in Figure 3), obtained by systematically dropping parameters based on order of magnitude until no further pathways could be dropped without causing a significant decrease in fit, allowed for four latent genetic factors. Additionally, this model dropped the individual pathway from the third latent genetic factor (A5 in figure 1) loading onto the RAPI. This model indicates that genetic variance across the measures of alcohol consumption and problems are accounted for by multiple latent genetic factors. The genetic correlations, computed for each pair of variables as the covariance of the two measures divided by the square root of the product of the variances of each of the measures, are shown in Table 4. They range from .45 (frequency of alcohol use with max drinks) to .99 (frequency of heavy drinking and frequency of intoxication).

VATSPSUD

Table 5 details the phenotypic correlations across the different measures of current alcohol consumption and lifetime symptoms of problem drinking. Polychoric correlations were computed on only one twin from each pair, chosen randomly. Note that while FT16 phenotypic correlations ranged from 0.25-0.75, VATSPSUD phenotypic correlations were somewhat higher ranging from 0.53-0.84. Table 6 shows the MZ and DZ twin correlations for each of the measures. We fit a series of models paralleling those fit in the FT16 data, as described above. The results of those models are shown in Table 7. Constraining all parameters to be equal in males and females (Model II), and dropping the full C Matrix (representing all shared environmental influences; Model III) provided better fits to the data, as indicated by decreases in the AIC and a nonsignificant χ^2 change. A systematic series of fitting submodels to test the significance of the individual genetic factors/pathways resulted in the best fitting Model VII, shown in Fable 4. Parallel to the results from the FinnTwin16 data, this model contained multiple latent genetic factors across the measures of alcohol consumption and alcohol problems. Genetic correlations for this sample are shown in Table 8, and range from .76 (drinking frequency and quantity) to .96 (drinking quantity and max drinks).

Table 5. VATSPSUD Phenotypic Correlations between Measures of Alcohol Consumption and Problems

Measure	**Drinking Frequency**	**Drinking Quantity**	**Frequency of Intoxication**	**Max Drinks**	**DSM-IV AD Sx**
Frequency	1				
Quantity	.53	1			
Freq of Intoxication	.73	.76	1		
Max Drinks	.68	.84	.79	1	
DSM AD Symptoms	.73	.70	.80	.79	1

Note: all correlations significant at p<0.001

Table 6. VATSPSUD MZ and DZ Correlations between Measures of Alcohol Consumption and Problems

Measure	**MZ Females**	**DZ Females**	**MZ Males**	**DZ Males**
Frequency	.56	.34	.46	.29
Quantity	.39	.24	.42	.24
Freq of Intoxication	.48	.29	.46	.29
Max Drinks	.48	.30	.53	.34
DSM AD Symptoms	.47	.27	.48	.24

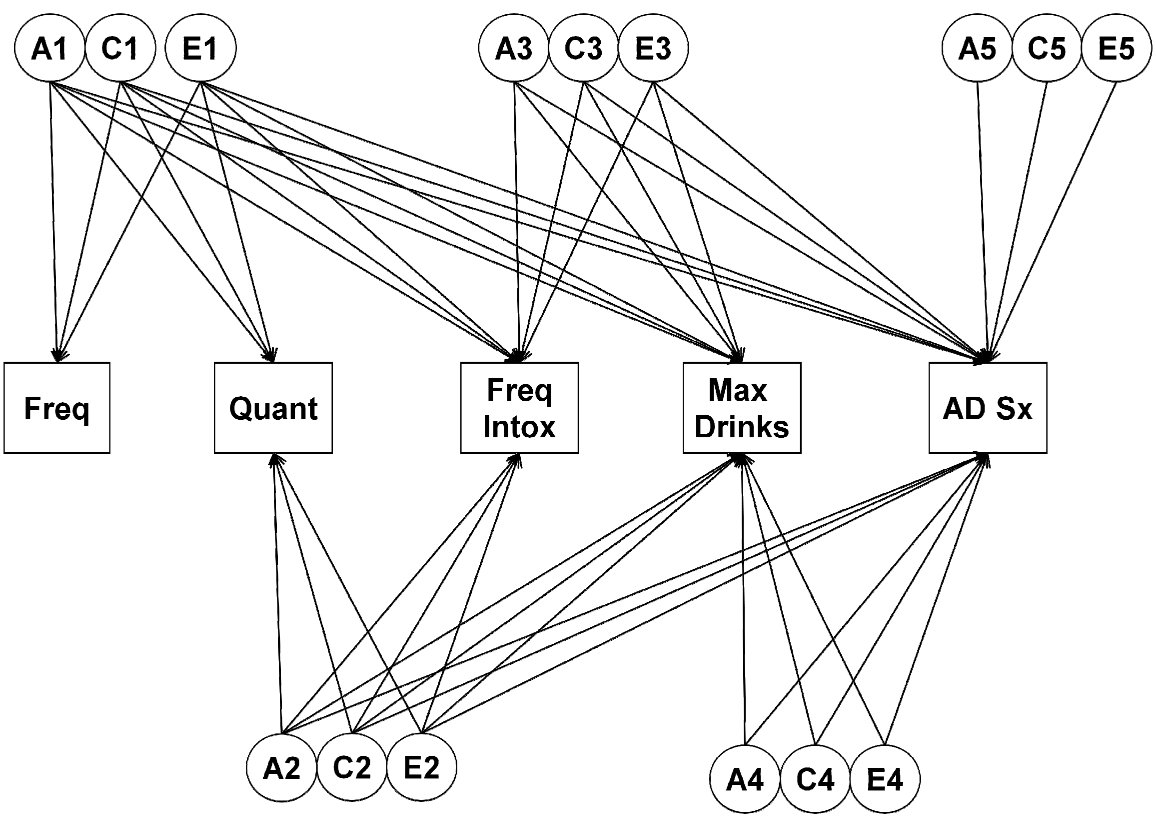
Note: all correlations significant at p<0.001

Table 7. VATSPSUD Model Fitting Results

Model			Δ Fit			
		Compared to Model	**Δ X^2 units**	**Probability**	**Δ DF**	**Δ AIC**
I*	Full Model	--	--	--	--	--
II	Sexes equated	I	9.42	0.86	45	20.58
III	C Matrix dropped	II	34.42	0.87	15	55.57
IV	A1	III	220.71	0.00	10	23.72
V	A1+A2	III	199.32	0.00	6	47.36
VI	A1+A2+A3	III	185.44	0.00	3	56.44
VI	A1+A2+A3+A4	III	74.08	0.05	1	58.09
VII^	**A1+A2+A3+A4+A5**	**III**	**35.96**	**0.90**	**3**	**60.04**

* Fit of Model I: − 2LL = 43147.81, df = 17540, AIC = 8067.81; All subsequent models are compared to Model I.
^ Best fit model.

Figure 2. VATSPSUD Full Model



In summary, the best fitting model across both samples indicated that a single latent genetic factor cannot explain the genetic influences on all consumption and problem measures. Rather, several latent genetic factors are needed (Figures 3 and 5). The first (A1) loads most heavily on the frequency items, but retains considerable influence across the other items. A second latent genetic factor (A2) loads more heavily on the heavier drinking items but again retains considerable influence on all items. Additional latent genetic factors are more specific to other consumption measures, with both samples showing some latent genetic influences specific to measures of alcohol problems (unshared with any of the measures of consumption).

Figure 3. FinnTwin16 Best Fitting Model: Additive Genetic Pathways

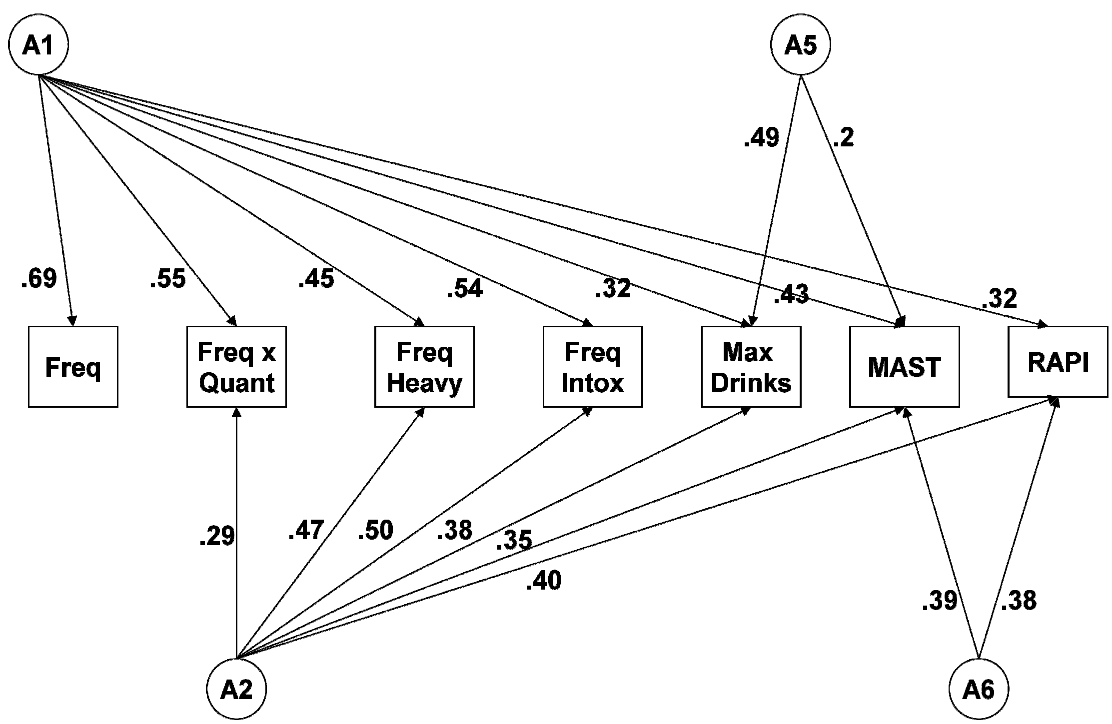
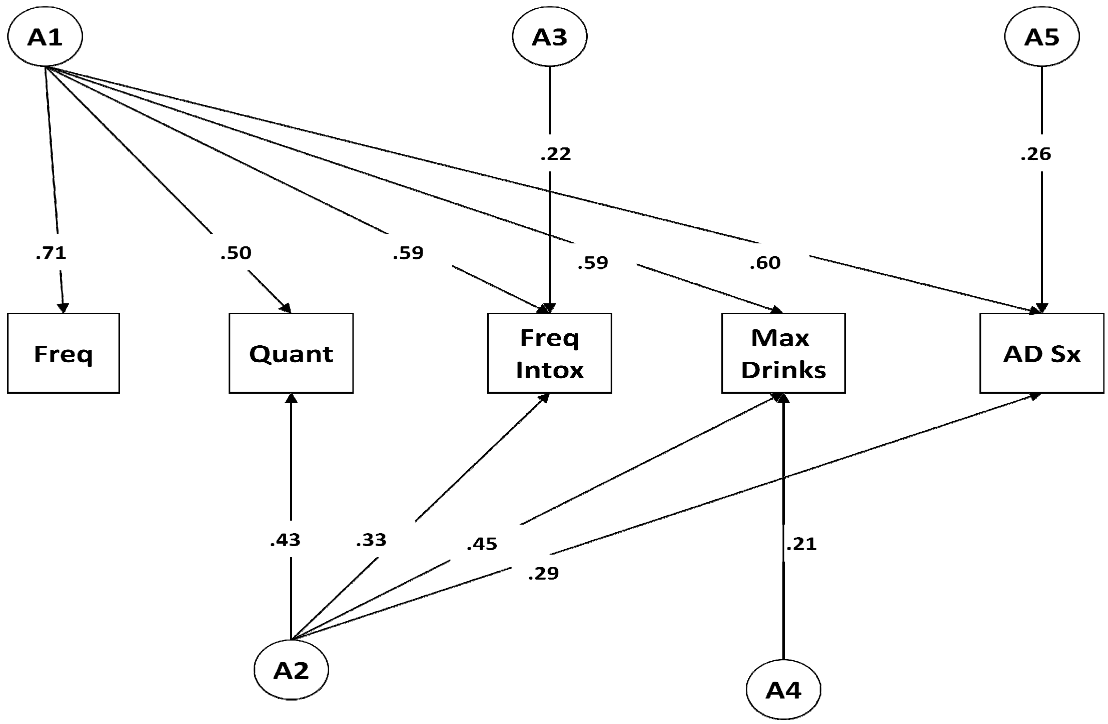


Figure 4. VATSPSUD Best Fitting Model: Additive Genetic Pathways



The goal of these analyses was to examine the underlying genetic architecture across measures of consumption and alcohol problems; accordingly, we did not test any models in which we dropped any component of the E matrix for either sample. Path estimates for the E parameters from the best-fitting models for the FinnTwin16 and VATSPSUD samples are shown in Figures 4 and 6, respectively.

Discussion

The initial genome-wide association studies have taught us that very large sample sizes will be necessary to identify genes of small effect (Wellcome Trust Case Control Consortium, 2007), as are assumed involved in psychiatric and substance use disorders. Failure to identify robust genetic effects reaching genome-wide significance has led to large-scale meta-analytic efforts (McMahon et al., 2010). But often the increase in sample size comes with a reduction in phenotypic specificity, because different assessment measures or outcomes have been used across different samples. Rather than assuming that different measures are influenced by the same genetic factors, twin studies provide a method to explicitly evaluate these relationships. In this study, we examined the genetic architecture across different measures of current alcohol consumption and problems in two independent twin samples from two different cultures: FinnTwin16 and the VATSPSUD. Previous analyses found a large proportion of overlap in the genetic factors that influence alcohol dependence and measures of alcohol consumption during the heaviest period of drinking. Our analyses also suggest considerable overlap of genetic influences across different indices of current drinking and different measures of alcohol problems, across both samples, as evidenced by genetic correlations ranging from .45 to .99. Across both samples, frequency of intoxication and quantity of alcohol use were more strongly genetically correlated with alcohol problems than frequency of use. The Kendler et al (2010) study of lifetime indices of consumption also found that drinking frequency had the lowest shared genetic overlap with alcohol problems. The Grant et al 2009 study only evaluated a composite consumption factor score, making it impossible to evaluate differential informativeness of various drinking indices. However, the available data from this study and the Kendler study suggest that quantity of alcohol consumption and frequency of heavy drinking or intoxication have greater shared genetic overlap with alcohol problem measures than measures of the frequency of alcohol use, which likely reflects a number social factors as well. Overall, genetic correlations were higher in the VATSPSUD sample, which may reflect the somewhat older mean age of the sample (36 versus 24 years of age) and more stabilized drinking patterns as individuals move further into adulthood. This suggests that meta-analytic studies may want to test for heterogeneity across samples according to age when using studies assessing consumption to replicate genetic findings originally identified with alcohol dependence, as drinking indices among slightly older adults may be more genetically correlated with alcohol problems than among younger adults, for whom drinking patterns are still more transitional.

Despite high genetic correlations, across both samples the genetic architecture is complex. A single latent genetic factor influencing all the consumption measures did not provide a good fit to the data in either sample. Rather, there are several different genetic factors that influence different measures of alcohol consumption. This indicates that there is not complete overlap across measures of alcohol consumption and alcohol problems, and there are different genetic influences impacting different indices of drinking. This has implications for gene identification studies in the area of alcohol dependence. It suggests that there are valid reasons why genetic findings may not “replicate” across studies that have assessed different aspects of alcohol use and dependence. In practice, this has already been seen in candidate gene studies, where genes have been associated with aspects of alcohol use, but not with alcohol dependence diagnoses (Dick et al., 2005; Foroud et al., 2007). Meta-analytic efforts that combine different indices of alcohol use and alcohol problems may enhance power to detect genetic influences that are shared across these measures, but they may miss some genetic influences specific to different aspects of alcohol use.

These findings should be interpreted in the context of several limitations. Although we believe that the demonstration of similar effects across two independent samples is a strength of the study, we note that the exact measures of alcohol use and alcohol problems collected in the two projects differed. Even when the construct was the same (e.g., drinking frequency), the exact wording of the item and response options varied across the samples. Differential reliabilities and distributional properties of the items could have influenced the emergent genetic factor structures. Differences in psychometric properties across the samples likely contributed to some of the observed sample variability. We believe that the convergence of results across these studies is notable, given that the samples contained slightly different measures of current consumption and different indices of problem drinking, covered different age ranges (the FT16 sample was limited to young adults while the VATSPUD sample covered a much broader age range of adults), and come from different drinking cultures. Another potential limitation of this study was choice of statistical model. In this manuscript, we chose to use a cholesky decomposition model. However, other models such as an independent pathway model and common pathway model could have been used to test this research question.

In summary, our analyses are consistent across two independent twin samples in finding fairly high genetic correlations across current alcohol consumption measures and alcohol problems. This is true across several different indices of consumption (frequency of drinking, quantity of alcohol use, frequency of heavy drinking/drunkenness) and using different measures of alcohol related problems (MMAST, RAPI, DSMIV symptom counts). Frequency of drinking appears to be the least genetically correlated with other measures of alcohol (less so than quantity of alcohol use/frequency of heavy drinking or drunkenness), suggesting there is more unique environmental variance on this aspect of alcohol use. This suggests that this measure may be least likely to “replicate” genetic effects identified with alcohol dependence. Both samples indicate that there is not a single genetic factor responsible for the phenotypic overlap between different measures of consumption and problem use. Accordingly, combining studies using different indices of alcohol use and problems may help increase power to identify shared genetic influences, but may introduce noise if the gene under study is more specific to a particular aspect of alcohol consumption. Creating multivariate genetic factor scores that take into account the extent to which different indices of alcohol use are reflective of the underlying genetic predisposition allows researchers to capitalize on all available information, while taking into account the differential informativeness of various indices of use. This illustrates one of the ways in which twin studies remain informative in the evolving era of gene identification.

References

Agrawal, A., Grant, J. D., Littlefield, A., Waldron, M., Pergadia, M. L., Lynskey, M. T., Madden, P. A., Todorov, A., Trull, T., Bucholz, K. K., Todd, R. D., Sher, K., & Heath, A. C. 2009, "Developing a quantitative measure of alcohol consumption for genomic studies on prospective cohorts", *J.Stud.Alcohol Drugs*, vol. 70, no. 2, pp. 157-168.

Association, A. P. 1994, *Diagnostic and statustical manual of mental disorders*, Fourth Edition edn, American Psychiatric Association, Washington, DC.

Dick, D. M. & Bierut, L. 2006, "The Genetics of Alcohol Dependence", *Current Psychiatry Report*, vol. 8, pp. 151-157.

Dick, D. M., Edenberg, H. J., Xuei, X., Goate, A., Hesselbrock, V., Schuckit, M., Crowe, R., & Foroud, T. 2005, "No association of the GABA-A receptor genes on chromosome 5 with alcoholism in the Collaborative Study on the Genetics of Alcoholism sample", *American Journal of Medical Genetics (Neuropsychiatric Genetics)*, vol. 132B, pp. 24-28.

Dick, D. M., Prescott, C., & McGue, M. 2009, "The Genetics of Substance Use and Substance Use Disorders," Y.-K. Kim, ed., Springer.

Edenberg, H., Bierut, L., Boyce, P., Cao, M., Cawley, S., Chiles, R., Doheny, K. F., Hansen, M., Hinrichs, A. L., Jones, K., Kelleher, M., Kennedy, G. C., Liu, G., Marcus, G., McBride, C., Murray, S. S., Oliphant, A., Pettengill, J., Porjesz, B., Pugh, E. W., Rice, J. P., Rubano, T., Shannon, S., Steeke, R., Tischfield, J. A., Tsai, Y. Y., Zhang, C., & Begleiter, H. 2005, "Description of the data from the Collaborative Study on the

Genetics of Alcoholism (COGA) and single-nucleotide polymorphism genotyping for Genetic Analysis Workshop 14", *BMC Genetics*, vol. 6, no. Suppl 1, p. S2.

Foroud, T., Wetherill, L. F., Liang, T., Dick, D. M., Hesselbrock, V., Kramer, J., Nurnberger, J., Schuckit, M., Carr, L., Porjesz, B., Xuei, X., & Edenberg, H. J. 2007, "Association of Alcohol Craving With alpha-Synuclein (SNCA)", *Alcoholism: Clinical and Experimental Research*, vol. 31, no. 4, pp. 537-545.

Goldman, D. 1993, "Recent developments in alcoholism:genetic transmission", *Recent Dev.Alcohol*, vol. 11, pp. 231-248.

Grant, J. D., Agrawal, A., Bucholz, K. K., Madden, P. A., Pergadia, M. L., Nelson, E. C., Lynskey, M. T., Todd, R. D., Todorov, A. A., Hansell, N. K., Whitfield, J. B., Martin, N. G., & Heath, A. C. 2009, "Alcohol consumption indices of genetic risk for alcohol dependence", *Biol.Psychiatry*, vol. 66, no. 8, pp. 795-800.

Heath, A. C. & Martin, N. G. 1994, "Genetic influences on alcohol consumption patterns and problem drinking: results from the Australian NH&MRC twin panel follow-up survey", *Ann.N.Y.Acad.Sci.*, vol. 708, pp. 72-85.

Kaprio, J. 2006, "Twin studies in Finland 2006", *Twin.Res.Hum.Genet.*, vol. 9, no. 6, pp. 772-777.

Kaprio, J., Pulkkinen, L., & Rose, R. J. 2002, "Genetic and environmental factors in health-related behaviors: studies on Finnish twins and twin families", *Twin Research*, vol. 5, pp. 358-365.

Kaprio, J., Rose, R. J., Romanov, K., & Koskenvuo, M. 1991, "Genetic and environmental determinants of use and abuse of alcohol: The Finnish twin cohort studies", *Alcohol & Alcoholism*, vol. Suppl. 1, pp. 131-136.

Kendler, K. S. & Prescott, C. A. 1999, "A population-based twin study of lifetime major depression in men and women", *Arch.Gen.Psychiatry*, vol. 56, no. 1, pp. 39-44.

Kendler, K. S. &. P. C. A. 2006, *Genes, Environment, and Psychopathology:Understanding the Causes of Psychiatric and Substance Use Disorders.* Guilford Press, New York.

Kendler, K. S. M. J. D. D. &. P. C. A. The Relationship between Genetic Influences on Alcohol Dependence and on Patterns of Alcohol consumption. Alcoholism: Clinical and Experimental Research . 2010.

Ref Type: In Press

Kristenson, H. & Trell, E. 1982, "Indicators of alcohol consumption: Comparisons between a questionnaire (Mm-MAST), interviews and serum y-Glutamyl transferase (GGT) in a health survey of middle-aged males", *British Journal of Addiction*, vol. 77, pp. 297-304.

McMahon, F. J., Akula, N., Schulze, T. G., Muglia, P., Tozzi, F., Detera-Wadleigh, S. D., Steele, C. J., Breuer, R., Strohmaier, J., Wendland, J. R., Mattheisen, M., Muhleisen, T. W., maier, W., Nothen, M. M., Cichon, S., Farmer, A., Vincent, J. B., Holsboer, F., Preisig, M., & Rietschel, M. 2010, "Meta-analysis of genome-wide association data identifies a risk locus for major mood disorders on 3p21.1", *Nat.Genet.*

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. Mx: Statistical Modeling. [5th Edition]. 1999. Box 126 MCV, Richmond, VA 23298, Department of Psychiatry.

Ref Type: Computer Program

Neale, M. C. & Cardon, L. R. 1992, *Methodology for genetic studies of twins and families* Kluwer Academic Publishers, Dordrecht.

Office of National Drug Control Policy 2004, *The Economic Costs of Drug Abuse in the United States, 1992-2002.*, Executive Office of the President, Washington, D.C..

Prescott, C., Sullivan, P. F., Myers, J., Patterson, D., Devitt, M., Halberstadt, L. J., Walsh, D., & Kendler, K. S. 2005, "The Irish Affected Sib Pair Study of Alcohol Dependence: study methodology and validation of diagnosis by interview and family history", *Alcoholism: Clinical and Experimental Research*, vol. 29, pp. 417-429.

Prescott, C. A. & Kendler, K. S. 1999, "Genetic and environmental contributions to alcohol abuse and dependence in a population-based sample of male twins", *Am.J.Psychiatry*, vol. 156, no. 1, pp. 34-40.

Rose, R. J. 1998, "A developmental behavior-genetic perspective on alcoholism risk", *Alcohol Health Res.World*, vol. 22, no. 2, pp. 131-143.

Seppa, K., Pitkajarvi, T., & Sillanaukee, P. 1999, "Alcohol consumption profile by time in middle-aged men: a longitudinal study based on three different diagnostic instruments", *Alcohol Alcohol*, vol. 34, no. 1, pp. 65-70.

Wellcome Trust Case Control Consortium 2007, "Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls", *Nature*, vol. 447, no. 7145, pp. 661-678.

White, H. R. & Labouvie, E. W. 1989, "Toward the assessment of adolescent problem drinking", *Journal of Studies on Alcohol*, vol. 50, pp. 30-37.

Chapter 2

The Association between *DRD2/ANKK1* and Genetically Informed Measures of Alcohol Use and Problems

*****This chapter is adapted from the following manuscript:*** *Meyers JL, Nyman E, Loukola A, Rose D, Kaprio J, Dick DM. The Association between DRD2/ANKK1 and Genetically Informed Measures of Alcohol Use and Problems.* Under review in *Addiction Biology.*

Abstract

Background: In 1991, Blum and colleagues first reported an association between *DRD2* and alcoholism. While there have been subsequent replications of this genetic association, there have also been numerous studies that failed to detect an association between *DRD2* and alcohol dependence. We propose that one aspect contributing to this inconsistency is the variation in alcohol phenotype used across studies.

Methods: Within the population based Finnish twin sample, FinnTwin16, we previously performed multivariate twin analyses to extract latent genetic factors which account for the variation across seven measures of alcohol consumption (frequency of drinking, frequency x quantity, frequency of heavy drinking, frequency of intoxication, and maximum drinks in a 24 hour period) and problems (the Rutgers Alcohol Problem Index-RAPI and the Mälmö-modified Michigan Alcohol Screen Test - MmMAST). In the present study, we examined the association between thirty-one *DRD2/ANKK1* SNPs and the genetic factor scores generated by twin analyses. We focus on two of the genetic factors: a general alcohol consumption and problems factor score which represents shared genetic variance across alcohol measures, and an alcohol problems genetic factor score which loads onto the two indices of problematic drinking (MAST and RAPI).

Results: After correction for multiple testing across SNPs and phenotypes, of the thirty-one SNPs genotyped across *DRD2/ANKK1,* one SNP (rs10891549) showed significant association with the general alcohol consumption and problems factor score ($p=0.004$), and four SNPs (rs10891549, rs1554929, rs6275, rs6279) showed significant association with the alcohol problems genetic factor score ($p=0.005$, $p=0.005$, $p=0.003$, $p=0.003$).

Conclusions: In this study, we provide additional positive evidence for the association between *DRD2/ANKK1* and alcohol outcomes, including frequency of drinking and drinking problems. Additionally, post hoc analyses indicate stronger association signals using genetic factor scores than individual measures, which suggest that accounting for the genetic architecture of the alcohol measures reduces genetic heterogeneity in alcohol dependence outcomes in this sample and enhances the ability to detect association.

Introduction

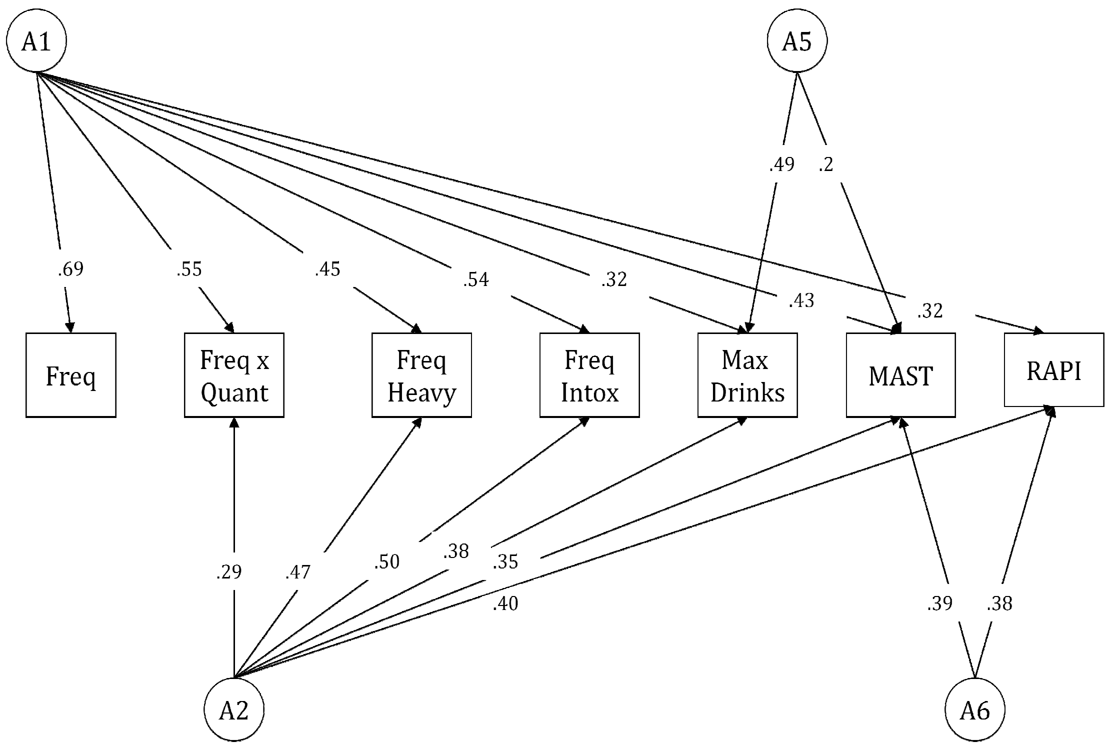
Alcohol consumption and problems are complex human behaviors that are influenced by both genetic and environmental risk factors (Kendler et al., 1992; Kendler et al., 1994). One strong candidate gene for alcohol-related outcomes is the dopamine receptor D2 gene (*DRD2*). In 1989, it was hypothesized that the rewarding effects of alcohol are mediated through the mesolimbic dopamine system (Wise and Rompre, 1989). The association between *DRD2* and alcoholism was first reported by Blum and colleagues, who found that an increased frequency of the Taq1A1 restriction fragment length polymorphism was observed in postmortem brain tissue from severe alcoholics (as compared to nonalcoholic controls) (Blum et al., 1991). Since this initial report, there has been an extensive literature examining the relationship between *DRD2* and alcohol-related outcomes. While there have been subsequent replications of this genetic association (Blum et al., 1991; Comings et al., 1991; Parsian et al., 1991; Amadeo et al., 1993; Noble et al., 1994; Higuchi et al., 1994; Neiswanger et al., 1995; Hietala et al., 1997; Kono et al., 1997; Ishiguro et al., 1998; Noble, 2003; Foley et al., 2004; Konishi et al., 2004), there have also been numerous studies across a variety of samples, populations, and study designs which fail to find an association between *DRD2* and alcohol outcomes (Arinami et al., 1993; Bolos et al., 1990; Chen et al., 1996, 1997, 2001; Cook et al., 1992; Cruz et al., 1995; Edenberg et al., 1998; Gelernter and Kranzler, 1999; Gelernter et al., 1991; Goldman et al., 1992, 1997; Lee et al., 1999; Lobos and Todd, 1998; Lu et al., 1996; Parsian et al., 2000; Sander et al., 1995, 1999; Schwab et al., 1991; Suarez et al., 1994; Turner et al., 1992; Waldman et al., 1999). Critics have proposed that much of this mixed literature resulted from the limitations of early genetic studies including small sample sizes and limited ability to tag all regions of a gene. However, results from more recent genetic association studies remain inconsistent with both positive (Hack et al., 2010, Filbey et al., 2011; Landgren et al., 2011; Van der Zwaluw et al., 2011; Bhaskar et al., 2011) and negative (Kasiakogia-Worlley et al., 2011; Creemers et al., 2011, Heath et al., 2011, Wang et al., 2011, Luo et al., 2011, Schumann et al., 2011) evidence for association between *DRD2* and alcohol problems. Interpreting this literature is further complicated by the 2004 discovery that the Taq1A polymorphism that had been most extensively studied was actually located 10 kb downstream from *DRD2* in a neighboring gene, *ankyrin repeat and kinase domain containing 1* (*ANKK1*) (Neville et al., 2004). The Taq1A variant is located within an exon of *ANKK1*, causing a non-synonymous coding change that may affect the substrate binding specificity of the gene product. It has been hypothesized that *ANKK1* may be involved in the dopaminergic reward pathway through signal transduction (Neville et al., 2004). There have been many reviews of the *DRD2* literature that provide detailed analysis of the variation across these genetic association studies (Goldman, 1998; Noble et al., 2000, Le Foll et al., 2009). However, little attention has been given to variability in the measurement of alcohol problems across these studies.

Many of the aforementioned studies used standard measures of alcohol use and/or problems including the Diagnostic and Statistical Manual of Mental Disorders (DSM) criteria, the Alcohol Dependence Scale (ADS), the Alcohol Expectancy Scale (AES), and the Alcohol Use Disorders Identification Test (AUDIT). Measures of alcohol problems vary by scientific field, setting (clinical vs. research), historical trend (DSM-III vs. DSM-IV), and availability. However, there is evidence to suggest that genetic association results may vary as a function of the alcohol measure used in the analysis. In 2002, Connor and colleagues tested the association between *DRD2* and a variety of alcohol phenotypes, finding association with certain alcohol phenotypes (alcohol quantity, alcohol consumed per week, alcohol dependence scale score) and not others (frequency of alcohol use). This is an example of how even when using an identical sample and method in genetic association analyses the measure of the phenotype can affect the results.

Twin studies provide a method for examining the genetic relationship between different measures of alcohol use and problems. While some twin studies indicate that the genetic correlation between measures of regular alcohol consumption and problems is strong (Grant et al., 2009; Kendler et al., 2010), there is also evidence that there are genetic risk factors unique to alcohol problems (Dick et al. 2011). Additionally, recent twin studies examining the genetic relationship between the DSM-IV alcohol dependence criteria have indicated that the seven items are not genetically homogeneous (Kendler et al, 2011). Therefore, different measures of alcohol use and problems may be mediated by different genetic factors. This has implications for gene identification studies in that there are valid reasons why true genetic findings may not replicate across studies that have assessed different aspects of alcohol use and dependence.

We previously reported analyses conducted within the Finnish population-based twin sample, FinnTwin16, to examine the genetic architecture across seven measures of alcohol consumption (frequency of drinking, frequency x quantity, frequency of heavy drinking, frequency of intoxication, and maximum drinks in a 24 hour period) and problems (the Rutgers Alcohol Problem Index-RAPI and the Mälmö-modified Michigan Alcohol Screen Test - MmMAST) (Dick et al., 2011). Our results yielded a model suggesting four latent factors that account for the genetic variance across the measures of alcohol consumption and measures of problems. The first two latent genetic factors loaded onto all of the drinking measures (consumption and problems), the third latent genetic factor loaded exclusively onto maximum drinks in a 24 hr period and the MmMAST, and the fourth latent genetic factor loaded onto the two indices of problems (the MmMAST and the RAPI). Using comparable measures of alcohol consumption and problems, data from an independent twin sample, the Virginia Adult Twin Study of Psychiatric and Substance Use Disorders, also indicated a parallel genetic architecture (Dick et al., 2011). This previously reported model from the Finntwin16 sample is depicted in Figure 4 from chapter 1 (also depicted below for reference).

Figure 4. Best Fitting Model of the Genetic Architecture of Measures of Alcohol Consumption and Problems in the Full Finntwin16 Sample (previously described in chapter 1)



In the present study, we extended these twin study results to examine the relationship between these measures of alcohol use/problems and *DRD2/ANKK1*. We hypothesized that examining association with genetic factor scores (previously implicated by the twin analyses within the same sample) would decrease the genetic heterogeneity and consequently increase power to detect genetic association between *DRD2/ANKK1* and alcohol outcomes. We were primarily interested in the shared genetic variance across all alcohol measures (Figure 4. latent genetic factor A1) and the shared genetic variance across the two indices of problematic alcohol use (Figure 4. latent genetic factor A6). Additionally, we conducted post hoc analyses of the association between *DRD2/ANKK1* and multiple measures of both alcohol consumption and problems in an effort to evaluate whether using genetic factor scores was an improvement upon using individual measures of alcohol consumption and problems.

Methods

Sample

Details regarding Finntwin16 (FT16) and data collection have been previously described in chapter 1 previous Finnish Twin Study publications (Kaprio et al., 2002; Kaprio et al., 2006). In this chapter, we focus on assessments of alcohol consumption and alcohol problems in young adulthood. The average age for the respondent twins at this assessment was 24.4 years (SD=1.50, range 22.8-27.2). Of these individuals, genotypic data was collected on 602 subjects, 36.0% were monozygotic (MZ) twins (n=216), 63.5% were dizygotic (DZ) twins (n=382).

Measures

Measures of alcohol consumption and problems are described in detail in chapter 1. Briefly, consumption measures included: *Frequency* (how often do you drink alcohol at all?), *Frequency x Quantity* (the frequency of reported use in the past 28 days multiplied by the quantity of drinks consumed per drinking day during the past 28 days; drinks defined as 1 beer, 1 glass of wine, or 1 mixed drink containing hard liquor equivalent to 10 grams of ethanol), *Frequency of Heavy Drinking* (at the present, how often do you within one occasion consume more than five bottles of beer, or more than a bottle of wine, or more than half a bottle of hard liquor?), *Frequency of Intoxication* (how often do you use alcohol to get drunk?), and *Max Drinks* (the maximum number of drinks twins reported ever consuming in a 24 hour period). Alcohol problem measures included: *The Mälmö -modified MAST* (Mm-MAST), a 9-item self-report scale of drinking patterns and problems designed for application in Nordic cultures) and the 22 items from the *Rutgers Alcohol Problem Index* (RAPI), a reliable scale designed to assess problematic drinking. Parallel to current practice in gene identification efforts for alcohol dependence, only individuals who had evidence of alcohol exposure were included in twin analyses, so that genetic/environmental influences on the decision to initiate alcohol are not confounded with genetic/environmental influences on alcohol consumption or problems. Altogether 2% of the sample had never had a full alcoholic beverage and were excluded from analyses. All measures were coded so that higher scores indicated more frequent drinking or more drinking problems.

Twin Modeling

The twin model we employed has been described in chapter 1. Briefly, a multivariate Cholesky model was fit to the measures of alcohol consumption and problems in order to estimate (1) the magnitude of genetic and environmental influences on each phenotype and (2) the extent to which these influences contributed to the covariation between the phenotypes. Using the statistical software package Mx (Neale and Cardon, 1992), we generated individual scores for each subject weighted by the loadings implicated by the genetic architecture from the best fitting twin model. When the best fitting model (Figure 1) from the full sample (n=2,500) was fit in the genotyped subset (n=602), there was not a significant decrease in model fit (χ^2=3.28, *p*=1.00). Thus, we moved the two strongest genetic factors forward in creating individual genetic factor scores for each person within the genotyped sample; (1) A general factor which loads onto measures of alcohol consumption and problems and (2) an alcohol problems factor which loads onto the Mm-MAST and the RAPI. This genetic factor score is similar to a phenotypic factor score in that it encompasses all shared variance across various measures. It differs in that it incorporates genetic information gained from twin data, therefore partitioning this shared variance into shared genetic variance across various measures. Thus, if an individual has an increased score on the specific alcohol measures that are loaded on by the latent genetic factor (e.g., Mm-MAST and RAPI), that individual will also to have an increased score on the genetic factor score (e.g., Alcohol Problems Genetic Factor, which loads onto Mm-MAST and RAPI).

Genotyping

A total of 602 individuals were genotyped using Sequenom's homogeneous Mass Extend (hME) and iPLEX Gold technology (Sequenom, San Diego, CA, USA). Thirty-one tagging single-nucleotide polymorphisms (SNPs) in *DRD2/ANKK1* were selected based on the HapMap Project (http://www.hapmap.org) and NCBI (http://www.ncbi.nlm.nih.gov) databases. The selected variants were bi-allelic and had a minor allele frequency (MAF) >10% in the Caucasian population. The ability to amplify the flanking regions of each SNP was determined by using the applications SNPper (http://www.snpper.chip.org) and RealSNP (http://www.realsnp.com), which define the most reliable regions for designing primers and the quality of the amplicons, respectively. All tagging SNPs failing during the procedure were replaced by newly generated tagging SNPs proposed by Haploview (Barrett, Fry, Maller, & Daly, 2005). The PCR and extension primers were designed using Sequenom's MassARRAY Assay

Design software (version 2.0). SNPs were genotyped in 384-well plates according to manufacturer's instructions. For quality controls, each plate contained at least eight water controls and 22 duplicate samples. PCR reactions were performed in a total reaction volume of 5µl using 20ng of genomic DNA. The alleles were automatically called by Sequenom's Mass ARRAY Typer Analyzer software and verified by two independent persons. Further marker-specific quality controls included a call rate >80% and a Hardy-Weinberg equilibrium (HWE) p-value >0.01 (estimated using unrelated individuals). Mendelian errors were excluded using PedCheck (O'Connell & Weeks, 1998).

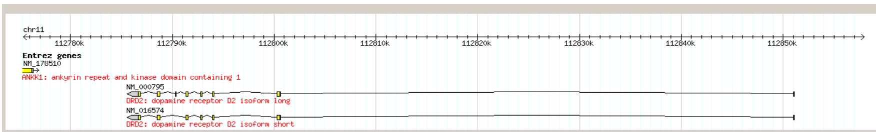
Once data were cleaned for quality control, genotypic data was available on 580 individuals of Finnish descent. An analysis of the population structure of the sample indicated a single ethnicity factor; thus all individuals were included in association analyses. Information on the genotyped SNPs, including chromosomal location and minor allele frequency is provided in Table 8. These thirty-one SNPs represent five different haplotype blocks across *DRD2/ANKK1* (Figure 2). These SNPs are correlated (r^2 range from .21-.93) yet represent five independent signals across *DRD2/ANKK1* as indicated by a Nyholt correction for related SNPs (Nyholt et al., 2004).

Genetic association analyses

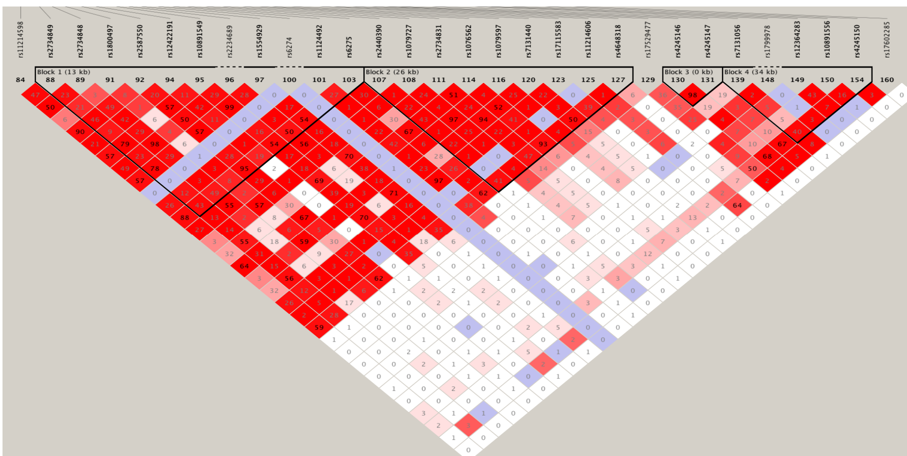
Linear regression was used to analyze the association between each of the SNPs and each of the genetic factor scores. The degree of relatedness (~50% for DZ twins and ~100% for MZ twins) was accounted for in the models using the GENMOD command in SAS 8.2 (SAS Institute, 2008). All p-value results from the association analyses were corrected for the number of independent tests conducted; the Nyholt correction indicated a significant threshold of $p<0.005$. Male and female data were collapsed in the genotypic analyses in order to maximize power to detect genetic association and to mirror the best fitting model from the twin analyses. Additionally, we conducted post hoc analyses of the association between *DRD2/ANKK1* and the seven individual measures of alcohol consumption and problems in order to test whether using genetic factor scores would result in different conclusions than had we analyzed multiple individual measures of alcohol use/problems. When evaluating results for the seven alcohol phenotypes, the Nyholt correction indicated a significant threshold of a $p<0.001$ to take into account the additional tests.

Figure 5. LD structure of *DRD2/ANKK1*

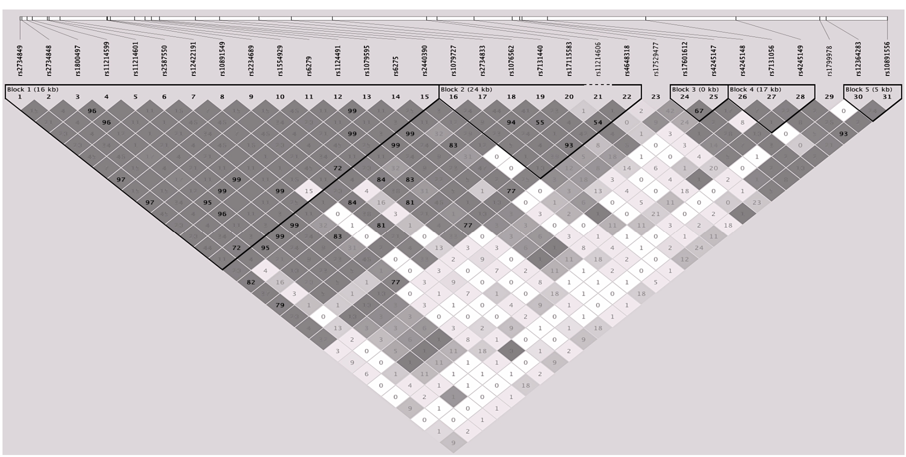
A)



B)



C)



Legend: Location of (A) and correlations between (B and C) the single-nucleotide polymorphisms (SNPs) genotyped in the *DRD2/ANKK1* gene complex (B) in the CEPH (Centre d'Etude du Polymorphisme Humain) data obtained from the HapMap database (The International HapMap Consortium, 2003) and (C) in the Finntwin16 data, Shading indicates the degree of correlation as measured by *D′* (Hedrick & Kumar, 2001); darker shading indicates higher correlations, and white shading indicates that markers are unlinked or uncorrelated. The numbers inside the diamonds are R^2 values, another measure of correlation between SNPs. The black triangles grouping subsets of SNPs indicate blocks of SNPs that are highly correlated (as defined by criteria detailed in Gabriel et al., 2002). Not all SNPs genotyped in the Finntwin16 sample were available in the HapMap database; in these cases, proxy SNPs that were the SNPs most highly correlated with the genotyped SNPs are listed. In the Finntwin16 sample, the LD blocks were similar to those in the HapMap CEPH data, and the somewhat stronger LD between markers is in agreement with previous findings from the Finnish population (Service et al., 2006).

Results

Twin Analyses

The phenotypic correlations across the measures of alcohol consumption and problems ranged from .45-.99 and were virtually identical to those previously reported in the full sample (Dick et al. 2011). Polychoric correlations were computed on only one twin from each pair, chosen randomly. MZ and DZ twin correlations for each of the measures were described previously (Dick et al. 2011). For the first genetic factor score (General Alcohol Consumption and Problems), scores ranged from -2.50 to 4.25 (mean=0, SD= 0.86). For the second genetic factor score (Alcohol Problems), scores ranged from -0.28 to 1.54 (mean=0, SD=0.52).

Genetic Association Analyses

Recall that the Nyholt threshold for a significant p-value for the two genetic factor scores is p<0.005. Of the thirty-one SNPs genotyped across *DRD2/ANKK1,* one SNP (rs10891549) showed significant association with the general alcohol consumption and problems factor score (p=0.004). Four SNPs (rs10891549, rs1554929, rs6275, rs6279) showed significant association with the alcohol problems genetic factor score (p=0.005, p=0.005, p=0.003, p=0.003, respectively). These results are detailed in Table 8. In addition, we conducted post hoc analyses in which we examined the association between *DRD2/ANKK1* SNPs and the individual seven phenotypic measures of alcohol consumption and problems. These results are detailed in Table 9. Recall that the Nyholt corrected p-value for the seven alcohol outcomes is p<0.001. Using this criterion, none of the *DRD2/ANKK1* SNPs were significantly associated with any of the individual alcohol measures.

Table 8. Linear Regression of *DRD2/ANKK1* SNPs on Genetic Factor Scores

DRD2 SNP Information						Genetic Factor Scores			
Chr	Gene	SNP	Base Pair Location	Alleles Major; Minor	MAF	Alcohol Consumption and Problems		Alcohol Problems (MAST and RAPI)	
						Beta	*p-value*	*Beta*	*p-value*
11	*ANKK1*	rs2734849	112775370	A;G	0.282	0.094	0.047	0.127	0.006
11	*ANKK1*	rs2734848	112775584	T;C	0.220	-0.040	0.401	-0.040	0.391
11	*ANKK1*	rs1800497	112776038	G;A	0.330	-0.003	0.945	-0.035	0.451
11	*DRD2*	rs11214599	112776570	C;T	0.330	-0.007	0.886	-0.043	0.353
11	*DRD2*	rs11214601	112777972	C;T	0.330	-0.004	0.936	-0.041	0.373
11	*DRD2*	rs2587550	112778135	A;G	0.120	-0.096	0.042	-0.103	0.026
11	*DRD2*	rs12422191	112779220	G;A	0.900	0.001	0.981	0.034	0.460
11	*DRD2*	**rs10891549**	112783657	T;C	0.235	**0.098**	**0.004**	**0.130**	**0.005**
11	*DRD2*	rs2234689	112783693	C;G	0.220	0.040	0.401	0.040	0.391
11	*DRD2*	**rs1554929**	112783974	C;T	0.235	0.098	0.039	**0.130**	**0.005**
11	*DRD2*	**rs6279**	112786283	C;G	0.118	-0.096	0.042	**-0.103**	**0.003**
11	*DRD2*	rs1124491	112787300	G;A	0.330	0.005	0.914	-0.042	0.367
11	*DRD2*	rs1079595	112787879	A;C	0.330	-0.004	0.936	-0.041	0.373
11	*DRD2*	**rs6275**	112788687	G;A	0.117	-0.099	0.038	**-0.102**	**0.003**
11	*DRD2*	rs2440390	112792088	C;T	0.080	-0.051	0.285	-0.014	0.757
11	*DRD2*	rs1079727	112794392	T;C	0.030	0.006	0.906	-0.035	0.444
11	*DRD2*	rs2734833	112798130	A;G	0.241	-0.098	0.038	-0.108	0.019
11	*DRD2*	rs1076562	112801218	G;A	0.095	-0.107	0.024	-0.087	0.060
11	*DRD2*	rs7131440	112805120	T;C	0.254	-0.104	0.028	-0.105	0.023
11	*DRD2*	rs17115583	112814112	G;A	0.043	-0.091	0.056	-0.081	0.081
11	*DRD2*	rs11214606	112815079	C;T	0.010	-0.007	0.875	-0.012	0.794
11	*DRD2*	rs4648318	112818599	T;C	0.103	-0.105	0.026	-0.074	0.111
11	*DRD2*	rs17529477	112822277	G;A	0.033	0.052	0.267	0.042	0.359
11	*DRD2*	rs17601612	112822955	G;C	0.063	0.025	0.595	0.035	0.446
11	*DRD2*	rs4245147	112823217	T;C	0.099	0.033	0.494	0.068	0.143
11	*DRD2*	rs4245148	112825629	C;T	0.060	0.033	0.491	0.089	0.053
11	*DRD2*	rs7131056	112834984	C;A	0.226	0.078	0.100	0.040	0.391
11	*DRD2*	rs4245149	112843567	G;A	0.052	-0.070	0.141	-0.079	0.087
11	*DRD2*	rs1799978	112851561	A;G	0.050	-0.044	0.255	0.019	0.684
11	*DRD2*	rs12364283	112852165	A:G	0.011	-0.021	0.655	0.000	0.997
11	*DRD2*	rs10891556	112857971	G;T	0.052	-0.073	0.126	-0.072	0.120

Note: SNPs that passed Nyholt threshold for significant association ($p<0.005$) are bolded. The reference build used in this table was HapMap Data Release 28 Phase II+III, August10, on NCBI B36 assesmbly dbSNP b126. The major allele frequencies (MAF) presented in this table were calculated using only one individual per family.

Table 9. Linear Regression of *DRD2/ANKK1* SNPs on Individual Measures of Alcohol Consumption and Problems

SNP	Alcohol Measures (p-values)						
	Frequency of Drinking	Frequency x Quantity	Frequency of Heavy Drinking	Frequency of Intoxication	Max Drinks 24 hr. Period	Michigan Alcohol Screen Test	Rutgers Alcohol Problem Index
rs2734849*	.016	.443	.032	.032	.278	.012	.007
rs2734848*	.119	.637	.379	.329	.455	.378	.522
rs1800497*	.662	.668	.839	.925	.802	.650	.337
rs11214599	.816	.593	.706	.729	.718	.512	.278
rs11214601	.777	.654	.749	.776	.667	.538	.293
rs2587550	.005	.695	.045	.036	.335	.027	.046
rs12422191	.404	.541	1.00	.732	.739	.225	.981
rs10891549	.012	.441	.028	.024	.230	.009	.006
rs2234689	.119	.637	.379	.329	.455	.378	.522
rs1554929	.012	.441	.028	.024	.230	.009	.006
rs6279	.005	.695	.045	.036	.335	.027	.046
rs1124491	.793	.640	.723	.759	.683	.549	.275
rs1079595	.777	.654	.749	.776	.667	.538	.293
rs6275	.004	.616	.042	.032	.407	.026	.053
rs2440390	.221	.806	.262	.108	.407	.750	.718
rs1079727	.566	.783	.885	.916	.756	.632	.430
rs2734833	.046	.345	.027	.013	.226	.034	.015
rs1076562	.010	.473	.010	.011	.261	.058	.073
rs7131440	.045	.294	.018	.008	.210	.034	.021
rs17115583	.039	.332	.063	.043	.579	.094	.084
rs11214606	.937	.927	.893	.752	.642	.816	.755
rs4648318	.014	.575	.028	.013	.311	.103	.126
rs17529477	.184	.388	.411	.090	.229	.424	.239
rs17601612	.632	.835	.853	.534	.482	.515	.327
rs4245147	.444	.800	.912	.586	.348	.209	.101
rs4245148	.309	.298	.927	.782	.343	.073	.080
rs7131056	.037	.160	.087	.075	.702	.550	.368
rs4245149	.023	.429	.258	.152	.729	.115	.102
rs1799978	.530	.528	.263	.154	.325	.357	.768
rs12364283	.568	.448	.671	.656	.811	.879	.935
rs10891556	.017	.434	.228	.129	.743	.140	.147

*Located in *ANKK1*

Conclusions

Two-decades of genetic studies have left the relationship between *DRD2/ANKK1* and alcoholism indeterminate. Many reasons have been put forth to explain the mixed association results. Among them, poor DNA extraction techniques, population stratification, and failure to properly screen controls for drug and alcohol disorders. Previous reviews of this literature have detailed the variability and limitations of these studies (Goldman, 1998). A 2000 review by Noble (Noble, 2000) focused on sample size, types of alcoholics analyzed, and the nature of comparative controls employed in a variety of previously published studies. He reviewed several samples each of which used varying measures of alcoholism (The Michigan Alcoholism Screening Test, the presence or absence of medical complications of alcoholism, alcohol consumption, Severity of Alcohol Dependence Questionnaire (SADQ), and the DSM-III-R criteria). In this paper, we focus on the variability in the measure of the phenotype used across this literature in an effort to understand how this variability may effect the conclusions one would draw about the evidence for association with *DRD2/ANKK1*.

The 36 studies published between 1991 and 2011(Table 10), have yielded both positive and negative evidence of association across a variety of alcohol phenotypes. If more weight is placed on the recently published studies (Dick et al., 2004; Hack et al., 2011; Creemers et al., 2011; Schumann et al., 2011), which are presumably better powered to detect genetic association in that they use larger sample sizes and test a greater number of markers across *DRD2/ANKK1* gene, and considering the publication bias that leaves many null results unreported, there is little evidence of association between *DRD2/ANKK1* and alcohol phenotypes. It does appear however, that most of the studies that used quantitative/continuous measures of alcohol use and problems provide positive evidence of genetic association between *DRD2/ANKK1* and alcohol related traits. This may reflect the fact that using quantitative measures can increase power to detect genetic association (Waldman et al., 1999, Kuo et al., 2010). However, it is of note that the largest of the aforementioned studies (Schumann et al., 2011), a meta-analyses of alcohol consumption GWAS on over 21,000 individuals, did not produce a genome wide significant variant in either *DRD2* or *ANKK1*. The association with *DRD2/ANKK1* appears to be contingent upon the specific measure of the phenotype, specific SNPs, and specific population used in a study. This is consistent with the implications of our twin studies that indicate that different genetic factors may contribute to risk for different measures of the "same" outcome (Dick et al., 2011). Moreover, while two measures of alcohol problems can both be valid and widely used, they are not necessarily genetically homogenous.

In the present study, we modeled the genetic architecture of the alcohol outcomes available in the Finntwin16 sample in an attempt to examine more genetically homogenous alcohol phenotypes. We found modest evidence of association between *DRD2/ANKK1* SNPs and both genetically informed measures of alcohol consumption and problems. As rs10891549 and rs1554929 are highly correlated (r^2=.98) and rs6275 and rs6279 are highly correlated (r^2=0.87), there were two true independent signals detected in this sample. The first of these signals (rs10891549/rs1554929) is highly correlated with the SNPs within the *ANKK1* gene, and may be indirectly associated with *ANKK1*, the original locus detected in association with alcohol problems. The association between the rs10891549/rs1554929 locus was found with both general alcohol consumption and problems in this sample. The second signal (rs6275/rs6279) may be potentially functional as rs6275 and rs6279 are non-synonymous polymorphisms that are located on the 3'UTR and may have a regulatory effect. This locus was only significantly associated with alcohol problems in the Finntwin16. Perhaps multiple independent signals within the *DRD2/ANKK1* gene complex are differentially associated with alcohol outcomes; this may provide some explanation of the inconsistent genetic association findings.

In an effort to assess the utility of the genetic factor score, we also examined the association between *DRD2/ANKK1* SNPs and the individual phenotypic measures of alcohol consumption and problems. As the inclusion of seven outcomes required a more stringent statistical test correction, no SNP passed the significance threshold put forth to correct for the multiple tests conducted. These results may suggest that we are indeed reducing genetic heterogeneity in the alcohol measures using the genetic factor scores. Additionally, we increase power to detect association in reducing the number of phenotypes examined (we correct for the analysis of two factor scores versus seven measures of alcohol consumption and problems). Thus, one can increase power to detect genetic association by (1) reducing the number of tests conducted, and (2) modeling the genetic architecture of the trait/disorder within your sample.

In summary, we provide modest evidence for the association between *DRD2/ANKK1* and alcohol use/ problems. In capturing the genetic heterogeneity across alcohol measures in genetic factor scores, we found association between *DRD2/ANKK1* SNPs with both regular and problematic drinking. It should be noted that the β values associated with each significant *DRD2/ANKK1* SNP range from 0.001- 1.30, indicating that a very small portion of the variation in alcohol behavior is accounted for by *DRD2/ANKK1* SNPs. In this study, we also demonstrated how to maximize the information obtained by twin analyses and molecular analyses within the same sample. By reducing the genetic heterogeneity inherent in the alcohol phenotype and the number of phenotypes analyzed, we detect a genetic association between *DRD2/ANKK1* and alcohol use and problems, which would have been deemed nonsignificant had we not incorporated the genetic architecture across the traits.

Table 10. Previously Published Studies on the Genetic Association between *DRD2/ ANKK1* and Alcohol Phenotypes

Study	**Measure of the Phenotype**	**Study Design**	**Sample Size**	**SNPS**	**Evidence of Association**
Blum et al.,1991	Severe alcoholics (post mortem samples)	Case/Control	96 cases (52 severe)	Taq1 A1	Positive
Comings et al., 1991	Michigan Alcohol Screen Test** x stress exposure	Cross-sectional	309 Honduran males	Taq1 A1	Positive
Gelernter et al., 1991	DSM-III-R Alcohol Dependence	Case/Control	44 white cases; 68 controls	Taq1 A1	Negative
Turner et al., 1992	DSM-III-R Alcohol Dependence; AD+medical complications	Cross-sectional	47 white males	Taq1 A1	Negative
Amadeo et al.,1993	DSM-III-R Alcohol Dependence	Case/Control	69 French Polynesian cases; 57 controls	Taq1 A1	Positive (combination of ADH2 and DRD2)
Arinami et al., 1993	DSM-III-R Alcohol Dependence; Greater severity	Case/Control	70 Japanese cases; 100 Japanese controls (unscreened)	Taq1 A1	Positive
Bolos et al.,1990	DSM-III-R Alcohol Dependence	Case/Control	40 white cases; 127 controls	Taq1 A1	Negative
Higuchi et al.,1994	DSM-III-R Alcohol Dependence; Greater severity (Feigner Criteria)	Case/Control	280 Japanese cases; 289 controls	Taq1 A1 (+)	Positive
Noble, 1994	SADQ (Severity)	Case/Control	73 cases; 80 controls	Taq1 A1	Positive
Suarez et al.,1994	Medical complications from Alcoholism	Case/Control	88 white cases; 89 controls	Taq1 A1 (+)	Negative
Geijer et al., 1994	DSM-III-R Alcohol Dependence	Case/Control	74 cases; 81 controls	Taq1 A1/B1	Negative
Cruz et al., 1995	Alcohol Withdrawal Symptoms	Case/Control	38 Mexican cases; 38 controls	Taq1 A1	Negative
Lu et al., 2001	DSM-III-R Alcohol Dependence;	Case/Control	34 cases with CD, 63 cases without	Taq1 A1/B1	Positive

	Conduct Disorder (CD)		CD; 85 controls		
Hietala et al., 1997	SADQ (Severity); MAST	Case/Control	70 Finnish male cases; 50 controls	Taq1 A1	Positive
Kono et al., 1997	DSM-III-R Alcohol Dependence; Early onset	Case/Control	100 Japanese cases; 93 controls	Taq1 A1	Positive
Ishiguro et al., 1998	DSM-III-R Alcohol Dependence	Case/Control	209 Japanese cases; 152 controls	Taq1 A1	Positive
Lobos and Todd, 1998	DSM-III-R Alcohol Dependence; Severity (Feigner Criteria)	Case/Control	55 cases; 80 controls	5 SNPs (6 haplotypes)	Negative
Edenberg et al., 1998	DSM3-R AD and Feigner Criteria	Linkage	433 cases; 401 controls	Taq1 A1	Negative
Sander et al., 1999	DSM3-R AD; Family history of Alcoholism	Case/Control	310 German cases; 196 controls	TaqI A (+)	Negative
Waldman et al., 1999	Quantitative Alcohol Measures**	TDT	433 cases; 401 controls (COGA)	Taq1 A1	Positive
Gelernter & Kranzler, 1999	DSM-III-R Alcohol Dependence	Case/Control	160 EA cases; 136 controls	Taq1 A1/B1	Negative
Lee et al., 1999	DSM-III-R Alcohol Dependence	Case/Control	128 cases; 85 controls	Taq1 A1	Negative
Parsian et al., 2000	Medical complications from alcoholism; Feigner Criteria; Cloninger Criteria	Case/Control	173 cases; 88 controls	TaqI A (+)	Negative
Chen et al., 2001	DSM-IV Alcohol Dependence	Case/Control	203 cases; 213 controls	-141C Ins/Del	Positive
Foley et al., 2004	Alcohol Consumption from medical records**			Taq1 A1/B1	Positive
Konishi et al., 2004	DSM-IV Alcohol Dependence	Case/Control	200 Mexican American cases; 351 controls	TaqI A1/B1	Positive
Dick et al., 2007	DSM-III-R Alcohol Dependence; Feigner Criteria	Family based asociation	219 Caucasian families (n = 1,923) (COGA)	26 single nucleotide polymorphis ms (SNPs) across DRD2/ANK K1	Positive
Hack et al., 2010	DSM-IV Alcohol Dependence;	Case/Control	545 Irish cases; 509 controls	15 DRD2 SNPs (excluding Taq1A1)	Negative
Filbey et al., 2011	Impulsive behavior on the Go/NoGo task Heavy Alcohol Drinking**	Cross-sectional	53 cases	rs1799732	Positive
Van der Zwaluw et al., 2011	Adolescent Binge Drinking	Cross-sectional	282 Dutch adolescent cases	Taq1A	Positive
Bhaskar et al.,	Michigan Alcohol	Case/Control	81 cases; 151	6 DRD2	Positive

2011	Screen Test **		controls	SNPs	
Creemers et al., 2011	Adolescent Regular alcohol use	Cross-sectional	1192 Dutch adolescents	Taq1A1	Negative
Schumann et al., 2011	Alcohol Consumption	Cross-sectional	21,607 drinkers	Affymetrix 500K coverage of DRD2	Negative

** Measure used in the present study

References

Agrawal, A., Grant, J. D., Littlefield, A., Waldron, M., Pergadia, M. L., Lynskey, M. T., Madden, P. A., Todorov, A., Trull, T., Bucholz, K. K., Todd, R. D., Sher, K., & Heath, A. C. 2009, "Developing a quantitative measure of alcohol consumption for genomic studies on prospective cohorts", *J.Stud.Alcohol Drugs*, vol. 70, no. 2, pp. 157-168.

Amadeo S, Abbar M, Fourcade ML, Waksman G, Leroux MG, Madex A, Selin M, Champiat J-C, Brethome A, Leclaire Y, Castelnau D, Venisse J-L, Mallet J (1993) D2 dopamine receptor gene and alcoholism. J Psychiatr Res 27:173–179.

Arinami T, Itokawa M, Komiyama T, Mitsushio H, Mori H, Mifune H, Hamaguchi H, Toru M (1993) Association between severity of alcoholism and the A1 allele of the dopamine D2 receptor gene TaqI A RFLP in Jap- anese. Biol Psychiatry 33:108–114.

American Psychiatric Association (1994) Diagnostic and Statistical Manual of Mental Disorders. 4th ed. American Psychiatric Association, Washington, DC.

Barrett, J. C., Fry, B., Maller, J., & Daly, M. J. (2005). Haploview: Analysis and visualization of LD and haplotype maps. Bioinformatics, 21(2), 263-265.

Begleiter H, Reich T, Hesselbrock V, Porjesz B, Li TK, Schuckit M, Edenberg HJ, Rice J (1995) The collaborative study on the genetics of alcoholism. Alcohol Health Res World 19:228–236.

Blomqvist O, Gelernter J, Kranzler HR (2000) Family-based study of DRD2 alleles in alcohol and drug dependence. Am J Med Genet B Neuropsychiatr Genet 96:659–664.

Blum K, Cull JG, Braverman ER, Comings DE (1996) Reward deficiency syn- drome. Am Sci 84:132–145.

Blum K, Noble EP, Sheridan PJ, Finley O, Montgomery A, Ritchie T, Ozk- aragoz T, Fitch RJ, Sadlack F, Sheffield D (1991) Association of the A1 allele of the D2 dopamine receptor gene with severe alcoholism. Arch Gen Psychiatry 48:409–416.

Blum K, Noble EP, Sheridan PJ, Montgomery A, Ritchie T, Jagadeeswaran P, Nogami H, Briggs AH, Cohn JB (1990) Allelic association of human dopamine D2 receptor gene in alcoholism. J Am Med Assoc 263:2055– 2060.

Boehnke M (1991) Allele frequency estimation from pedigree data. Am J Hum Genet 48:22–25.

Bolos AM, Dean M, Lucase-Derse S, Ramsburg M, Brown GL, Goldman D (1990) Population and pedigree studies reveal a lack of association between the D2 receptor gene and alcoholism. JAMA 264:3156–3160.

Bucholz KK, Cadoret R, Cloninger CR, Dinwiddie SH, Hesselbrock VM, Nurnberger JI, Reich T, Schmidt I, Schuckit MA (1994) A new, semi-struc- tured psychiatric interview for use in genetic linkage studies: a report on the reliability of the SSAGA. J Stud Alcohol 55:149–158.

Chen WJ, Chen C-H, Huang J, Hsu Y-PP, Seow S-V, Chen C-C, Cheng ATA (2001) Genetic polymorphisms of the promoter region of dopamine D2 receptor and dopamine transporter genes and alcoholism among four aboriginal groups and Han Chinese in Taiwan. Psychiatr Genet 11:187– 195.

Chen C-H, Chien SH, Hwu HG (1996) Lack of association between TawI A1 allele of dopamine D2 receptor gene and alcohol-use disorders in Atayal natives of Taiwan. Am J Med Genet B Neuropsychiatr Genet 67:488– 490.

Chen WJ, Lu M-L, Hsu Y-PP, Chen C-C, Yu J-M, Cheng ATA (1997) Dop- amine D2 receptor gene and alcoholism among four aboriginal groups and Han in Taiwan. Am J Med Genet B Neuropsychiatr Genet 74:129–136.

Clayton D (1999) A generalization of the transmission ⁄ disequilibrium test for uncertain haplotype transmission. Am J Hum Genet 65:1170–1177.

Comings DE, Blum K (2000) Reward deficiency syndrome: genetic aspects of behavioral disorders. Prog Brain Res 126:325–341.

Comings DE, Comings BG, Muhleman D, Dietz G, Shahbahrami B, Tast D, Knell E, Kocsis P, Baumgarten R, Kovacs BW (1991) The dopamine D2 receptor locus as a modifying gene in neuropsychiatric disorders. JAMA 266:1793–1800.

Connor JP, Young RM, Lawford BR, Ritchie TL, Noble EP.(2002) D(2) dopamine receptor(DRD2) polymorphism is associated with severity of alcohol dependence. Eur Psychiatry. 17(1):17-23.

Cook BL, Wang ZW, Crowe RR, Hauser R, Freimer M (1992) Alcoholism and the D2 receptor gene. Alcohol Clin Exp Res 16:806–809.

Crabbe JC, Phillips TJ (1998) Genetics of alcohol and other abused drugs. Drug Alcohol Depend 51:61–71.

Cruz C, Camarena B, Mejia JM, Paez F, Eroza V, de la Fuente JR, Kershe- nobich D, Nicolini H (1995) The dopamine D2 receptor gene TaqI A1 poly- morphism and alcoholism in a Mexican population. Arch Med Res 26:421– 426.

Dick, DM, Meyers, JL, Rose, D, Kaprio, J, Kendler, K.S. (2011) Measures of Current Alcohol Consumption and Problems: Two Independent Twin Studies Suggest A Complex Genetic Architecture. Alcohol Clin Exp Res 35(12): 2152-61.

Edenberg HJ, Foroud T, Koller DL, Goate A, Rice J, Van Eerdewegh P, Reich T, Cloninger CR, Nurnberger JI, Kowalczuk M, Wu B, Li TK, Conneally PM, Tischfield JA, Wu W, Shears S, Crowe R, Hesselbrock V, Schuckit M, Porjesz B, Begleiter H (1998) A family-based analysis of the association of the dopamine D2 receptor (DRD2) with alcoholism. Alcohol Clin Exp Res 22:505–512.

Finckh U, Rommelspacher H, Kuhn S, Dufeu P, Otto G, Heinz A, Delttling M, Giraldo-Valasquez M, Pelz J, Graf K-J, Harms H, Sander T, Schmidt LG, Rolfs A (1997) Influence of the dopamine D2 receptor (DRD2) geno- type on neuroadaptive effects of alcohol and the clinical outcome of alcohol- ism. Pharmacogenetics 7:271–281.

Foley PF, Loh E-W, Innes DJ, Williams SM, Tannenberg AEG, Harper CG, Dodd PR (2004) Association studies of neurotransmitter gene polymor- phisms in alcoholic Caucasians. Ann N Y Acad Sci 1025:39–46.

Gelernter J, Goldman D, Risch N (1993) The A1 allele at the D2 dopamine receptor gene and alcoholism. JAMA 269:1673–1677.

Gelernter J, Kranzler HR (1999) D2 dopamine receptor gene (DRD2) allele and haplotype frequencies in alcohol dependent and control subjects: no association with phenotype or severity of phenotype. Neuropsychopharma- cology 20:640–649.

Gelernter J, O’Malley S, Risch N, Kranzler HR, Krystal J, Merikangas K, Kennedy JL, Kidd KK (1991) No association between an allele at the D2 dopamine receptor gene (DRD2) and alcoholism. JAMA 266:1801–1807.

Gelernter J, Yu Y, Weiss R, Brady K, Panhuysen C, Yang BZ, Kranzler HR, Farrer L (2006) Haplotype spanning TTC12 and ANKK1, flanked by the DRD2 and NCAM1 loci, is strongly associated to nicotine dependence in two distinct American populations. Hum Mol Genet 15:3498–3507.

Goldman D, Brown GL, Albaugh B, Robin R, Goodson S, Trunzo M, Akh- tar L, Lucas-Derse S, Long JC, Linnoila M, Dean M (1993) DRD2 dop- amine receptor genotype, linkage disquilibrium, and alcoholism in American Indians and other populations. Alcohol Clin Exp Res 17:199– 204.

Goldman D, Dean M, Brown GL, Bolos AM, Tokola R, Virkkunen M, Lin- noila M (1992) D2 dopamine receptor genotype and cerebrospinal fluid homovanillic acid, 5-hydroxyindoleacetic acid and 3-methoxy-4-hydroxy- phenylglycol in alcoholics in Finland and the United States. Acta Psychiatr Scand 86:351–357.

Goldman D, Urbanek M, Guenther D, Robin R, Long JC (1997) Linkage and association of a functional DRD2 variant (Ser311Cys) and DRD2 markers to alcoholism, substance abuse and schizophrenia in Southwestern American Indians. Am J Med Genet B Neuropsychiatr Genet 74:386–394.

Gorwood P, Batel P, Gouya L, Courois F, Feingold J, Ades J (2000) Reap- praisal of the association between the DRD2 gene, alcoholism and addiction. Eur Psychiatry 15:90–96.

Grant, J. D., Agrawal, A., Bucholz, K. K., Madden, P. A., Pergadia, M. L., Nelson, E. C., Lynskey, M. T., Todd, R. D., Todorov, A. A., Hansell, N. K., Whitfield, J. B., Martin, N. G., & Heath, A. C. 2009, "Alcohol consumption indices of genetic risk for alcohol dependence", *Biol.Psychiatry*, vol. 66, no. 8, pp. 795-800.

Heath AC, Whitfield JB, Martin NG, Pergadia ML, Goate AM, Lind PA, McEvoy BP, Schrage AJ, Grant JD, Chou YL, Zhu R, Henders AK, Medland SE, Gordon SD, Nelson EC, Agrawal A, Nyholt DR, Bucholz KK, Madden PA, Montgomery GW. (2011) A quantitative-trait genome-wide association study of alcoholism risk in the community: findings and implications. Biol Psychiatry 15;70 (6):513-8.

Hesselbrock M, Easton C, Bucholz KK, Schuckit M, Hesselbrock V (1999) A validity study of the SSAGA—a comparison with the SCAN. Addiction 94:1361–1370.

Hesselbrock V, Hesselbrock M (1994) Alcoholism and subtypes of antisocial personality disorder. Alcohol Alcohol 2(Suppl):479–484.

Hietala J, Pohjalainen T, Heikkila-Kallio U, West C, Salaspuro M, Syvalahti E (1997) Allelic association between D2 but not D1 dopamine receptor gene and alcoholism in Finland. Psychiatr Genet 7:19–25.

Higuchi S, Muramatsu T, Murayama M, Hayashida M (1994) Association of structural polymorphism of the dopamine D2receptor gene and alcoholism. Biochem Biophys Res Commun 204:1199–1205.

Hill SY, Zezza N, Wipprecht G, Zu J, Neiswanger K (1999) Linkage studies of D2 and D4 receptor genes and alcoholism. Am J Med Genet B Neuro- psychiatr Genet 88:676–685.

Ishiguro H, Arinami T, Akazawa S, Enomoto M, Mitushio H, Fujishiro H, Tada K, Akimoto Y, Mifune H, Shioduka S, Hamaguchi H, Toru M, Shi- buya H (1998) Association study between the -141C Ins/Del and TaqI A FAMILY-BASED ASSOCIATION ANALYSES OF ALCOHOL DEPENDENCE PHENOTYPES 1653 polymorphisms of the dopamine D2 receptor gene and alcoholism. Alcohol Clin Exp Res 22:845–848.

Kaprio, J. 2006, "Twin studies in Finland 2006", *Twin.Res.Hum.Genet.*, vol. 9, no. 6, pp. 772-777.

Kaprio, J., Pulkkinen, L., & Rose, R. J. 2002, "Genetic and environmental factors in health-related behaviors: studies on Finnish twins and twin families", *Twin Research*, vol. 5, pp. 358-365.

Kaprio, J., Rose, R. J., Romanov, K., & Koskenvuo, M. 1991, "Genetic and environmental determinants of use and abuse of alcohol: The Finnish twin cohort studies", *Alcohol & Alcoholism*, vol. Suppl. 1, pp. 131-136.

Kendler, K. S. & Prescott, C. A. 1999, "A population-based twin study of lifetime major depression in men and women", *Arch.Gen.Psychiatry*, vol. 56, no. 1, pp. 39-44.

Kendler, K. S. &. P. C. A. 2006, *Genes, Environment, and Psychopathology:Understanding the Causes of Psychiatric and Substance Use Disorders.* Guilford Press, New York.

Kendler, K. S. M. J. D. D. &. P. C. A. The Relationship between Genetic Influences on Alcohol Dependence and on Patterns of Alcohol consumption. Alcoholism: Clinical and Experimental Research . 2010.

Kidd KK, Morar B, Castiglione CM, Zhao H, Pakstis AJ, Speed WC, Bonne-Tamir B, Lu R-B, Goldman D, Lee C, Nam YS, Grandy DK, Jenkins T, Kidd JR (1998) A global survey of haplotype frequencies and linkage dis- equilibrium at the DRD2 locus. Hum Genet 103:211–227.

Konishi T, Calvillo M, Leng A-S, Lin K-M, Wan Y-JY (2004) Polymor- phisms of the dopamine D2 receptor, serotonin transporter, and GABA-A receptor B3 subunit genes and alcoholism in Mexican Americans. Alcohol 32:45–52.

Kono Y, Yoneda H, Sakai T, Nonomura Y, Inayama Y, Koh J, Sakai J, Ina- da Y, Imamichi H, Asaba H (1997) Association between early-onset alco- holism and the dopamine D2 receptor gene. Am J Med Genet B Neuropsychiatr Genet 74:179–182.

Kristenson, H. & Trell, E. 1982, "Indicators of alcohol consumption: Comparisons between a questionnaire (Mm-MAST), interviews and serum y-Glutamyl transferase (GGT) in a health survey of middle-aged males", *British Journal of Addiction*, vol. 77, pp. 297-304.

Lawford BR, Young RM, Rowell JA, Gibson JN, Feeney GFX, Ritchie TL, Syndulko K, Noble EP (1997) Association of the D2 dopamine receptor A1 allele with alcoholism: medical severity of alcoholism and type of controls. Biol Psychiatry 41:386–393.

Lee JF, Lu RB, Ko HC, Chang FM, Yin S-J, Pakstis AJ, Kidd KK (1999) No association between DRD2 locus and alcoholism after controlling the ADH and ALDH genotypes in Chinese Han population. Alcohol Clin Exp Res 23:592–599.

Le Foll B, Gallo A, Le Strat Y, Lu L, Gorwood P. Genetics of dopamine receptors and drug addiction: a comprehensive review. Behav Pharmacol. 2009 Feb;20(1):1-17.

Li J, Ji L (2005) Adjusting multiple testing in multilocus analyses using the ei- genvalues of a correlation matrix. Heredity 95:221–227.

Lobos EA, Todd RD (1998) Association analysis in an evolutionary context: cladistic analysis of the DRD2 locus to test for association with alcoholism. Am J Med Genet B Neuropsychiatr Genet 81:411–419.

Lu R-B, Ko HC, Chang FM, Castiglione CM, Schoolfield G, Pakstis AJ, Kidd JR, Kidd KK (1996) No association between alcoholism and multiple polymorphisms at the dopamine D2 receptor gene (DRD2) in three distinct Taiwanese populations. Biol Psychiatry 39:419–429.

Martin ER, Monks SA, Warren LL, Kaplan NL (2000) A test for linkage and association in general pedigrees: The Pedigree Disequilibrium Test. Am J Hum Genet 67:146–154.

Matsushita S, Muramatsu T, Murayama M, Nakane J, Higuchi S (2001) Alcoholism, ALDH2*2 allele and the A1 allele of the dopamine D2 receptor gene: an association study. Psychiatry Res 104:19–26.

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. Mx: Statistical Modeling. [5th Edition]. 1999. Box 126 MCV, Richmond, VA 23298, Department of Psychiatry.

Neale, M. C. & Cardon, L. R. 1992, *Methodology for genetic studies of twins and families* Kluwer Academic Publishers, Dordrecht.

Neiswanger K, Hill SY, Kaplan BB (1995) Association and linkage studies of the TAQI A1 allele at the dopamine D2 receptor gene in samples of female and male alcoholics. Am J Med Genet B Neuropsychiatr Genet 60:267–271.

Neville MJ, Johnstone EC, Walton RT (2004) Identification and characteriza- tion of ANKK1: a novel kinase gene closely linked to DRD2 on chromo- some band 11q23.1. Hum Mutat 23(6):540–545.

Noble EP (2000a) Addiction and its reward process through polymorphisms of the D2 dopmaine receptor gene: a review. Eur Psychiatry 15:79–89.

Noble EP (2000b) The DRD2 gene in psychiatric and neurological disorders and its phenotypes. Pharmacogenetics 1:309–333.

Noble EP (2003) D2 dopamine receptor gene in psychiatric and neurologic dis- orders and its phenotypes. Am J Med Genet B Neuropsychiatr Genet 116:103–125.

Noble EP, Syndulko K, Fitch RJ, Ritchie T, Bohlman MC, Guth P, Sheridan PJ, Montgomery A, Heinzmann C, Sparkes RS (1994) D2 dopamine recep- tor TawI A alleles in medically ill alcoholic and nonalcoholic patients. Alco- hol Alcohol 29:729–744.

Noble EP, Zhang X, Ritchie TL, Sparkes RS (2000) Haplotypes at the DRD2 locus and severe alcoholism. Am J Med Genet B Neuropsychiatr Genet 96:622–631.

Nyholt DR (2004) A simple correction for multiple testing for single-nucleo- tide polymorphisms in linkage disequilibrium with each other. Am J Hum Genet 74:765–769.
Parsian A, Cloninger CR, Zhang Z-H (2000) Functional variant in the DRD2 receptor promoter region and subtypes of alcoholism. Am J Med Genet B Neuropsychiatr Genet 96:407–411.

O'Connell, J. R., & Weeks, D. E. (1998). PedCheck: A program for identification of genotype incompatibilities in linkage analysis. American Journal of Human Genetics, 63(1), 259-266.

Parsian A, Todd RD, Devor EJ, O'Malley KL, Suarex BK, Reich T, Clonin- ger CR (1991) Alcoholism and alleles of the human D2 dopamine receptor locus. Arch Gen Psychiatry 48:655–663.

Ponce G, Jimenez-Arriero MA, Rubio G, Hoenicka J, Ampuero I, Ramos JA, Palomo T (2003) The A1 allele of the DRD2 gene (TaqI A polymor- phisms) is associated with antisocial personality in a sample of alcohol- dependent patients. Eur Psychiatry 18:356–360.

Reich T (1996) A genomic survey of alcohol dependence and related pheno- types: results from the Collaborative Study on the Genetics of Alcoholism (COGA). Alcohol Clin Exp Res 20(Suppl 8):133A–137A.

Rose, R. J. 1998, "A developmental behavior-genetic perspective on alcoholism risk", *Alcohol Health Res.World*, vol. 22, no. 2, pp. 131-143.

Sander T, Harms H, Podschus J, Finckh U, Nickel B, Rolfs A, Rommelspa- cher H, Schmidt LG (1995) Dopamine D1, D2, and D3 receptor genes in alcohol dependence. Psychiatr Genet 5:171–176.

Sander T, Ladehoff M, Samochowiec J, Finckh U, Rommelspacher H, Sch- midt LG (1999) Lack of an allelic association between polymorphisms of the dopamine D2 receptor gene and alcohol dependence in the German population. Alcohol Clin Exp Res 23:578–581.

SAS Institute (2008). SAS OnlineDoc Version 9.2. SAS Institute, Inc. : Cary, NC.

Schuckit MA, Danko GP, Smith TL, Hesselbrock V, Kramer J, Bucholz KK (2003) A 5-year prospective evaluation of DSM-IV alcohol dependence with and without a physiological component. Alcohol Clin Exp Res 27:818–825.

Schwab S, Soyka M, Niederecker M, SAckenheil M, Scherer J, Wilderauer DB (1991) Allelic association of human dopamine D2-receptor DNA poly- morphism ruled out in 45 alcoholics. Am J Hum Genet 49(Suppl):203.

Schumann G, Coin LJ, Lourdusamy A, Charoen P, Berger KH, Stacey D, Desrivières S, Aliev FA, Khan AA, Amin N, Aulchenko YS, Bakalkin G, Bakker SJ, Balkau B, Beulens

JW, Bilbao A, de Boer RA, Beury D, Bots ML, Breetvelt EJ, Cauchi S, Cavalcanti-Proença C, Chambers JC, Clarke TK, Dahmen N, de Geus EJ, Dick D, Ducci F, Easton A, Edenberg HJ, Esko T, Fernández-Medarde A, Foroud T, Freimer NB, Girault JA, Grobbee DE, Guarrera S, Gudbjartsson DF, Hartikainen AL, Heath AC, Hesselbrock V, Hofman A, Hottenga JJ, Isohanni MK, Kaprio J, Khaw KT, Kuehnel B, Laitinen J, Lobbens S, Luan J, Mangino M, Maroteaux M, Matullo G, McCarthy MI, Mueller C, Navis G, Numans ME, Núñez A, Nyholt DR, Onland-Moret CN, Oostra BA, O'Reilly PF, Palkovits M, Penninx BW, Polidoro S, Pouta A, Prokopenko I, Ricceri F, Santos E, Smit JH, Soranzo N, Song K, Sovio U, Stumvoll M, Surakk I, Thorgeirsson TE, Thorsteinsdottir U, Troakes C, Tyrfingsson T, Tönjes A, Uiterwaal CS, Uitterlinden AG, van der Harst P, van der Schouw YT, Staehlin O, Vogelzangs N, Vollenweider P, Waeber G, Wareham NJ, Waterworth DM, Whitfield JB, Wichmann EH, Willemsen G, Witteman JC, Yuan X, Zhai G, Zhao JH, Zhang W, Martin NG, Metspalu A, Doering A, Scott J, Spector TD, Loos RJ, Boomsma DI, Mooser V, Peltonen L, Stefansson K, van Duijn CM, Vineis P, Sommer WH, Kooner JS, Spanagel R, Heberlein UA, Jarvelin MR, Elliott P (2011) Genome-wide association and genetic functional studies identify autism susceptibility candidate 2 gene (AUTS2) in the regulation of alcohol consumption. Proc Natl Acad Sci U S A. 2011 Apr 26;108(17):7119-24. Erratum in: Proc Natl Acad Sci U S A. 2011 May 31;108(22):9316.

Service, S., DeYoung, J., Karayiorgou, M., Roos, J. L., Pretorious, H., Bedoya, G., et al., (2006) Suarez BK, Parsian A, Hampe CL, Todd RD, Reich T, Cloninger CR (1994) Linkage disequilibria at the D2 dopmaine receptor locus (DRD2) in alco- holics and controls. Genomics 19:12–20.

Thanos PK, Taintor NB, Rivera SN, Umegaki H, Ilkari H, Roth G, Ingram DK, Hitzemann R, Fowler JS, Gatley SJ, Wang G-J, Volkow ND (2004) DRD2 gene transfer into the nucleus accumbens core of the alcohol prefer- ring and nonpreferring rats attenuates alcohol drinking. Alcohol Clin Exp Res 28:720–728.

Turner E, Ewing J, Shilling P, Smith TL, Irwin M, Schuckit M, Kelsoe JR (1992) Lack of association between an RFLP near the D2 dopamine recep- tor gene and severe alcoholism. Biol Psychiatry 31:285–290.

Volkow ND, Fowler JS, Wang G-J, Swanson JM (2004) Dopamine in drug abuse and addiction: results from imaging studies and treatment implica- tions. Mol Psychiatry 9:557–569.

Wang KS, Liu X, Aragam N, Jian X, Mullersman JE, Liu Y, Pan Y., Waldman ID, Robinson BF, Rhee SH (1999) Family-based association analysis of alcohol dependence in the COGA sample and replication in the Australian twin-family study. J Neural Transm. 118(9):1293-9.

Waldman ID, Robinson BF, Rhee SH (1999) A logistic regression extension of the transmission disequilibrium test for continuous traits: application to linkage disequilibrium between alcoholism and the candidate genes DRD2 and ADH3. Genet Epidemiol 17(Suppl 1):S379–S384.

White, H. R. & Labouvie, E. W. 1989, "Toward the assessment of adolescent problem drinking", *Journal of Studies on Alcohol*, vol. 50, pp. 30-37.

Wise RA, Rompre P-P (1989) Brain dopamine and reward. Annu Rev Psychol 40:191–225.

Chapter 3

A Genome Wide Association Study of Alcohol Dependence Symptoms in the Population Based Finnish Twin Cohort, *FinnTwin12*

Abstract

Background: In 2009, the first genome wide association study (GWAS) on alcohol dependence was published. Since then, several alcohol dependence GWAS have been reported without producing robust, replicable genetic association signals, with a notable few exceptions.

Methods: In the present study, we conducted a genome wide association study of DSM-IV Alcohol Dependence symptoms (AD sx) in the population-based Finnish twin study, *Finntwin12*. GWAS data was available on ~1,069 individuals (406 MZs; 614 DZs) who were genotyped on the Illumina 670K Single Nucleotide Polymorphism (SNP) Custom Array. Primary GWAS analyses of AD sx presented in this study included SNP-based analyses (PLINK), gene-based analyses (VEGAS) and gene enrichment analyses of gene-based results (ToppFun). In addition, we also analyzed two genetic factor scores that emerged from the multivariate twin analyses of five measures of alcohol consumption and problems conducted in this sample (Mx). In an effort to capture the most robust associations, comparisons between AD sx genetic association results and the genetic factors were carried out on both the SNP and gene level.

Results: GWAS analyses of AD sx indicated that no individual SNP met criteria for the genome wide significance threshold. However many SNPs were approaching this threshold, including several SNPs located on 4p16.3 in *docking protein 7(DOK7)*. Additionally, we ran gene-based analyses that produced a number of top gene results detailed in this manuscript, including *gamma-aminobutyric acid receptor subunit gamma-1 (GABRG1)* and *DOK7*. Gene enrichment analyses suggested that genes with ion-channel activity were overrepresented in the AD sx gene-based results. Comparisons between genetic association results from AD sx and the genetic factors implicated different variants on both the SNP level (correlations between AD sx SNP based results and genetic factor scores range from 0.10-0.50) and gene level (correlations between AD sx gene-based results and genetic factor scores range from 0.06-0.25).

Conclusions: We provide modest evidence of association between AD sx and several novel genetic variants (both SNPs and genes) that approach genome wide significance, including *DOK7*, which was implicated in both SNP and gene-based analyses. In addition, gene-based results implicated a previously reported genetic association between *GABRG1* and alcohol dependence. Discordance between genetic association results from AD sx and the genetic factors underscores the difficultly in replicating genetic effects and in differentiating real findings from spurious ones. Convergence in results across phenotypes, methods, and samples may provide us the most robust genetic association signals.

Introduction

In 2009, the first genome wide association study (GWAS) on alcohol dependence (AD) was published (Treutlein et al., 2009). This study included 487 German male inpatients with alcohol dependence as defined by the DSM-IV and an age at onset younger than 28 years, and 1,358 population-based control individuals. This study also included a follow-up sample of 1,024 German male inpatients and 996 age-matched male controls. This initial GWAS implicated two novel intergenic single nucleotide polymorphisms (SNPs) that reached stringent genome wide significance thresholds required to correct for multiple testing (rs7590720, rs1344694). Since then, several alcohol dependence GWAS have been reported and are detailed in table 11. From 2010-2011, six large GWA studies were published (Lind et al., 2010, Bierut et al., 2010, Edenberg et al., 2010, Kendler et al., 2011, Heath et al., 2011, Wang et al, 2011), none of which reported genome wide significant findings. Thus far, two very large alcohol dependence GWAS have been published in 2012 (Zuo et al., 2012, Frank et al., 2012), both of which have reported genome wide significant findings. Earlier this year, Zuo and colleagues combined the Study of Addiction Genetics and Environment (SAGE) data and Australian family study of alcohol use disorder (OZ-ALC) with the goal of discovering novel risk loci for alcohol dependence. The authors reported that variants within *KIAA0040* and the *PHF3-PTP4A1* gene complex might harbor a causal variant for AD (Zuo et al., 2012). Frank and colleagues (Frank et al., 2012) conducted an AD

GWAS on 1,333 German (inpatient) cases and 2,168 German controls and reported genome-wide significant support for the role of the ADH gene cluster (*ADH1B/ADH1C*). In addition to these AD GWAS reports, several studies have conducted association with alcohol-related phenotypes, such as alcohol consumption. Many studies have suggested that use of a quantitative measure could improve power to detect variants of small effect (Agrawal et al., 2009). In 2010, Joslyn and colleagues conducted a GWAS on level of response to alcohol in 367 individuals and reported no genome wide significant findings. However in 2011, two large studies conducted GWAS on alcohol consumption (Baik et al., 2011, Schumann et al., 2011) and reported genome wide significant findings. Baik and colleagues reported genome wide significant signals in (or near) *C12orf51, CCDC63*, and *MYL2* that were successfully replicated in a sample of Korean male drinkers; rs2074356, located in *C12orf51*, was in high linkage disequilibrium with SNPs in *ALDH2*, but other SNPs were not (Baik et al., 2011). The largest alcohol related GWAS to date examined alcohol consumption in 12 population-based samples of European ancestry, comprising 26,316 individuals, with replication genotyping in an additional 21,185 individuals. SNP rs6943555 in autism susceptibility candidate 2 gene (*AUTS2*) was associated with alcohol consumption at a genome-wide significant level (Schumann et al., 2011). Most recently, Agrawal and colleagues conducted a GWAS on alcohol craving in 3,976 individuals and reported no genome wide significant findings.

In reviewing the current state of alcohol dependence GWAS findings, six of the sixteen studies reviewed in table 11 report genome-wide significant findings. At this point, evidence that the genome-wide significant variants implicated in these studies replicate in an independent sample is limited. However, there is some suggestion from this literature that larger sample sizes and quantitative measures of alcohol use may increase the likelihood (via an increase in statistical power) of identifying genome wide significant findings.

For these reasons, conducting GWAS on quantitative measures of alcohol consumption has gained popularity. Consideration of the genetic relationship between alcohol consumption and alcohol dependence is prudent. Twin studies indicate that the genetic correlation between measures of regular alcohol consumption and dependence is strong (Grant et al., 2009; Kendler et al., 2010), however there is also evidence that there are genetic risk factors unique to alcohol problems (Dick et al. 2011). Thus, different measures of alcohol use and problems may be mediated by different genetic factors. This has implications for gene identification studies in that there are valid reasons why true genetic findings may not replicate across studies that have assessed different aspects of alcohol use and/or dependence.

We have previously extended these twin studies to examine the relationship between measures of alcohol use/problems and candidate gene, *DRD2* in *Finntwin16*, another cohort of the Finnish Twin Studies (Meyers et al., 2012 under review). The multivariate twin analyses of the seven measures of alcohol use and problems generated two genetic factors of interest; a general alcohol consumption and problems factor score which represents shared genetic variance across alcohol measures, and an alcohol problems genetic factor score which loads onto the two indices of problematic drinking (Michigan Alcohol Screen Test (Selzer et al., 1971) and Rutgers Alcohol Problems Index (White HR, Labouvie, 1989)). The results provided modest evidence for the association between *DRD2* and alcohol outcomes, including frequency of drinking and drinking problems. More importantly, the results indicated that one may increase power to detect genetic association by modeling the genetic architecture of the trait/disorder. This is in part achieved by reducing the number of phenotypes for analysis.

In the present study, we conducted a genome wide association study (GWAS) on DSM-IV Alcohol Dependence symptoms (AD sx) within *Finntwin12*, an independent cohort from the population based Finnish Twin Studies. In this study, we present GWAS analyses of AD sx including individual SNP-based association and gene-based association. Among the top genes associated with AD sx, we conducted gene enrichment analyses in which we tested for the overrepresentation of a particular gene function within the set. In addition, we conducted GWAS on two genetic factor scores that emerged from the multivariate twin analyses of five measures of alcohol consumption and problems conducted in this sample. In an effort to capture the most robust associations for alcohol use/problems in this sample, we compared genetic association results from AD sx with genetic association results from the genetic factors on both the SNP and gene level.

Table 11. Summary of Published Genome-Wide Association Studies on Alcohol Dependence and Consumption

Study	Alcohol Phenotype	Sample	GWAS design	Genetic Variants Implicated in Study	Genome-Wide Sig?
Treutlein et al., 2009	DSM-IV AD (age at onset younger than 28 years)	487 German male inpatient cases and 1,358 population-based controls; Follow-up study: 1,024 German male inpatient cases and 996 controls.	Case/ Control	**rs7590720, rs1344694**, PECR, PPP2R2B	Yes
Lind et al., 2010	DSM-IV AD	1,224 Australian cases and 1,162 controls	Case/ Control	CTBP2, KRT3, TJP1	No
Lind et al., 2010	DSM-IV AD/ND	599 cases and 488 controls	Case/ Control	**rs7530302, rs1784300, rs12882384** (located in KIAA1409), CTBP2, MYOM1, ORIL6, MALTI, ARHGAP10, ENPP6, PRAGMI, MTR	Yes**
Bierut et al., 2010	DSM-IV AD	1,897 cases and 1,932 controls.	Case/ Control	GABRA2, PNOX2, CC2D2B, SHBP5, GRM5	No
Edenberg et al., 2010	DSM-IV AD	847 cases; 552 controls	Case/ Control	SLC22A18, PHLDA2, NAP1L4, SNORA54, CARS,OSBPL5 CPE, DNASE2B, SLC10A2, ARL6IP5, ID4, GATA4, SYNE1, ADCY3, BBX	No
Joslyn et al., 2010	Level of Response to Alcohol	367 individuals	Quantitative		No
Kalsi et al., 2010	DSM-IV AD symptoms	562 cases	Quantitative		No
Baik et al., 2011	Average daily alcohol consumption	1721 Korean males from a population-based cohort. Replication sample: 1113 males	Quantitative	**C12orf51, CCDC63, MYL2, OAS3, CUX2, RPH3A**	Yes
Kendler et al., 2011	Alcohol Dependence Factor Score	3,169 individuals from the population-based Molecular Genetics of Schizophrenia (MGS2) control sample.	Quantitative	KCNMA1, AKAP9, PIGG, CEACAM6, KCNQ5, SLC35B4, MGLL, ADH1C, NFKB1, ANKK1 ADH5, POMC, CHRM2	No
Schumann et al., 2011	Alcohol Consumption	26,316 individuals, with replication genotyping in an additional 21,185 individuals.	Quantitative	AUTS2	Yes
Heath et al., 2011	alcohol dependence, dependence factor score, and heaviness of drinking factor score,	2062 Australian cases and 3393 controls	Case/ Control	TMEM108, ANKS1A	No
Wang et al., 2011	DSM-IV AD	1283 EA cases and 1416 EA controls	Case/ Control	ALK, CASC4, and SEMA5A,KIAA0040,THSD7B, NRD1, PKNOX2	No
Zuo et al., 2012	DSM-IV AD	1409 EA cases with AD, 1518 EA controls	Case/ Control	**KIAA0040**, TNN, TNR	Yes**
Frank et al., 2012	DSM-IV AD	1333 German male in-patient cases and 2168 controls	Case/ Control	**rs1789891**, which is located between the ADH1B and ADH1C genes	Yes
Edwards et al., 2012	DSM-IV AD/MDD	467 EA cases and 407 EA controls	Case/ Control	CDH13, CSMD2, GRID1, and HTR1B	No
Agrawal et al., 2012	Alcohol Craving	3976 individuals	Quantitative	ITGAD	No

Note: **Genome-wide significant finding are bolded** ** Replicated Genome-wide significant finding

Methods

Sample

FinnTwin12 is a longitudinal population based developmental twin study that followed five consecutive birth cohorts of twins born 1983-1987 identified through Finland's central population registry (n = 5600 twins). The study was initially designed to examine genetic and environmental influences on health-related behaviors. Questionnaire assessments of both twins and their parents were collected at baseline, before the twins reached age 12, with follow-up of all twins at ages 14, 17.5 and 22. At the age 22 follow up, GWAS data was collected on a subset (n=1,069; 406 MZs and 614 DZs) of the sample. In all, 1,347 questionnaires were returned at age 22 out of 4,236 of those already participating in earlier questionnaires. Zygosity was initially determined using a well-validated questionnaire completed by both co-twins at the baseline (Kaprio, Pulkkinen, & Rose 2002). Later, DNA from venous blood or saliva samples were used to confirm the zygosity in same-sex pairs with 97% accuracy. Here we focus on the age 22 assessments, as we were interested in examining genetic risk factors for young adult drinking problems and related behavior.

Measures

DSM-IV AD Symptoms (AD sx) were assessed for lifetime in the interviews based on seven DSM-IV criteria (American Psychological Association, 1994). Scores ranged from (0) No Symptoms endorsed to (7) All Seven AD symptoms endorsed (mean=1.09, SD=1.37). AD sx scores were highly skewed, with over 70% of the sample endorsing one symptom or fewer. 180 individuals (16.84% of the sample) endorsed three or more alcohol dependence criteria. Only individuals who had evidence of alcohol exposure were included in twin analyses so that genetic influences on the decision to initiate alcohol are not confounded with genetic influences on alcohol consumption or problems. 34 individuals (3.2% of the genotyped sample) indicated that they had never tried alcohol.

Twin Modeling: Genetic Factor Scores

Measures

Parallel to current practice in gene identification efforts for alcohol dependence, only individuals who had evidence of alcohol exposure were included in twin analyses so that genetic and environmental influences on the decision to initiate alcohol are not confounded with genetic and environmental influences on alcohol consumption or problems. All measures were coded so that an increased score indicated more frequent drinking or more drinking problems. *Frequency of Drinking* (Frequency) was assessed by the following question: “How many days per week do you drink alcohol?” Response options included: 0-7 and were recoded into five categories based on a quintile split of the data: (0) 0, (1) 1, (2) 2, (3) 3 and (4) 4-7 days. *Drinking Quantity* (Quantity) was assessed with the following question: “On those days when you drink, how many drinks did you usually have in a day?” Responses ranged from 1-29 and were collapsed into the following categories based on a quintile split of the data: (0) 1-3, (1) 4-6, (2) 7-8, (3) 9-12 and (4) 13+ drinks. *Frequency of Intoxication* (Intoxication) was assessed with the following question: “How often did you use alcohol to get drunk?” Response options included: 0-7 and were recoded into five categories based on a quintile split of the data:

(0) 0, (1) 1, (2) 2, (3) 3 and (4) 4-7 days. *Maximum Drinks/ 24 hr. period* (Max Drinks) was assessed with the following question: "What is the largest number of drinks you had on any single day?" Responses ranged from 1-54 and were collapsed into the following categories based on a quintile split of the data: (0) 1-9, (1) 10-12, (2) 13-17, (3) 18-23 and (4) 24+ drinks.

Twin Model

All details of the twin modeling have been detailed in previous publications (Dick et al., 2011). Briefly, a multivariate Cholesky model was used to estimate genetic and environmental influences across the measures of consumption/problem drinking (Neale and Cardon, 1992). Alternative models, including variations on the independent (Akaike's Information Criterion (AIC): 7019.077) and common pathway models (AIC: 7156.380), were tested for fit comparison (detailed in supplemental table 20); preliminary model fitting suggested that the Cholesky model provided the best fit to the data (AIC: 4495.392). Analyses were conducted using the seven measures of alcohol consumption and problems. The Cholesky model allows us to evaluate (1) the magnitude of genetic and environmental influences on each phenotype and (2) the extent to which these influences contribute to the covariation between the phenotypes. The full model calculated variance components separately by sex. Additional models were tested to evaluate goodness-of-fit in which estimates of the variance components were constrained to be equal across sex. Estimates were obtained from observed twin data using maximum likelihood estimation in the software program Mx (Neale et al., 1999). Model fit was evaluated by (AIC), and the probability (p) value associated with the χ^2 statistic. Lower AIC values indicate an optimal balance between explanatory power and parsimony. Additionally, nonsignificant χ^2 values ($\underline{p}$ >. 05) indicate a good fit. We compared nested alternative models by the change in chi-square between models, which is used to evaluate the significance of dropping parameters. A significant change in χ^2 ($\underline{p} < .05$) for the difference in degrees of freedom of the models indicates that the model with fewer degrees of freedom should be adopted, because the gain in degrees of freedom of the alternate model caused a significant decrease in fit. Missing data were handled by reading raw data into Mx and fitting to the observed and unobserved data vectors using full information maximum likelihood estimation.

Genetic Factor Scores

The latent genetic factor structure from the best fitting model was used to create individual genetic factor scores for each subject. Using the statistical software package Mx (Neale et al, 1999), individual scores were generated for each subject, weighted by the loadings implicated by the genetic architecture from the best fitting twin model. This genetic factor score is similar to a phenotypic factor score in that it encompasses all shared variance across various measures. It differs in that it incorporates genetic information gained from twin data, therefore partitioning this shared variance into shared genetic variance across various measures. Thus, if an individual has an increased score on the specific alcohol measures that are loaded on by the latent genetic factor (e.g., frequency and quantity of drinking) they will also have an increased score on the genetic factor score (e.g., Figure 6 genetic factor A1, which loads onto frequency and quantity of drinking).

Genome Wide Association Analyses

AD sx

Once data was cleaned for quality control, GWAS data was available on ~1,069 individuals (406 MZs; 614 DZs) who were genotyped on the Illumina 670K Custom Array. An analysis of the population structure of the sample indicated a single ethnicity factor; thus all individuals were included in association analyses. Using the statistical package Plink (Purcell et al., 2007), regression analyses were run treating the phenotype as a quantitative trait and accounting for the twin structure of the data using the Qfam (quantitative trait, family data) command. Because the qfam procedure can specify only one type of familial relationship, both individuals from each DZ pair and one individual from each MZ twin pair was included in the analyses, reducing the sample size from 1,069 to 872 individuals (6 of these individuals were excluded as they had not been exposed to alcohol). GWAS of AD sx included both SNP-based and gene-based analyses. In the SNP-based analyses, each marker was run separately; thus to account for the multiple testing a threshold of 8.89x10E-8 (Bonferoni correction= 0.05/535,613 markers analyzed) was required to meet genome wide significance. In the gene-based analyses, each gene was run separately in Versatile Gene Based Association Study (VEGAS (Liu et al., 2010)). For gene-based tests of association, VEGAS applies a gene-wise correction based on the number of independent signals in each gene. Permutation testing was conducted on both SNP based and gene based analyses that provided corrected (empirical) p-values. Once gene-based tests of association were performed, we conducted a gene enrichment analyses on the top (empirical p-value<0.01) genes associated with AD sx using the Topp Gene Suite tool, Topp Fun

(Chen et al., 2009). Topp Fun empirically tests whether a particular gene function is overrepresented, or enriched, within a set of genes.

Genetic Factors

Parallel SNP-based and gene-based genome-wide association analyses were conducted on the genetic factors. In an effort to capture the most robust associations for alcohol use/problems in this sample, we compared genetic association results from AD sx with genetic association results from the genetic factors on both the SNP and gene level. SNP level results were compared by examining the correlation between the log of the p-values associated with each SNP using the statistical software SAS. Gene level results were compared by examining the concordance between top gene (empirical p-value <0.01) sets for each phenotype using Gene Weaver: a web based system for the integration of functional genomics experiments. (Baker et al., 2012).

Results

AD sx Genome Wide Association Study

SNP Based Analyses

GWAS analyses of AD sx indicated no individual SNP that met criteria for the genome wide significance threshold (8.89x10E-8), however many SNPs were approaching this threshold and are detailed in table 12 below. Of the 535,613 SNPs analyzed, 101 SNPs had an FDR (BH) less than 10%. The most significant SNP result was the association between AD sx and rs10022329 (p-value= 6.02E-07), which resides in Docking protein 7 (*DOK7*).

Table 12. Variants from PLINK's SNP based analyses of AD sx (empirical pvalue<5.98E-05)

Chr	Located in Gene	SNP	BP Location	Prior evidence of associated with:	p-value
1	intergenic	rs9662365	98234031		4.69E-05
1	intergenic	rs1505551	99321856		5.98E-05
1	intergenic	rs10157998	112591162		2.89E-05
2	intergenic	rs13013813	211705048		3.09E-05
3	HMGB1	rs2122369	22507664		3.56E-06
3	HMGB1	rs1947238	22519902		2.17E-06
3	intergenic	rs2713001	111483418		4.96E-05
3	UPK1B/TSPAN20	rs6797796	120387620		3.71E-06
4	DOK7	rs10022329	3437317	congenital myasthenic syndromes (Muller et al., 2007)	6.02E-07
4	Dok7	rs7680504	3468951	congenital myasthenic syndromes (Muller et al., 2007)	5.71E-05
4	intergenic	rs16988673	32829156		5.94E-05
4	intergenic	rs1497499	55412719		5.65E-05
4	intergenic	rs1874647	55416226		5.65E-05
4	intergenic	rs6824301	81368193		2.00E-05
4	intergenic	rs11934116	81369133		2.20E-05
4	intergenic	rs2033613	142315874		2.32E-05
7	intergenic	rs1888349	138314672		5.75E-05
9	BNC2	rs10810585	16661045	ovarian cancer (Goode et al., 2010)	4.90E-05
9	intergenic	rs7042753*	87291493		5.02E-05
11	FXYD6	rs564989*	117214964	schizophrenia (Ito et al., 2008)	1.68E-05
11	FXYD6	rs6589624	117222507	schizophrenia (Ito et al., 2008)	5.38E-05
11	FXYD6	rs531855*	117231637	schizophrenia (Ito et al., 2008)	5.38E-05
14	ZBTB7A (3' UTR)	rs1542313	64069791		1.69E-05
14	C14orf50	rs3742604	64095995		3.07E-05
14	C14orf50 (non-synonymous)	rs6573560	64101287		2.30E-05
19	upstream	rs12461092	19180484		1.74E-05
19	intergenic	rs7246529	22856963		2.53E-06
19	intergenic	rs12460438	22867619		2.53E-06
20	TSHZ2	rs6022360	51313268	breast and prostate cancer (Yamamoto et al., 2011)	8.70E-06
22	TCN2	rs740234	29338745		3.68E-05

Note: Boxes indicate that SNPs are in high Linkage Disequilibrium (r2>.8); *Nominally significant association (p<0.01) with AD sx in the Collaborative Study on the Genetics of Alcoholism GWAS (Edenberg et al., 2010)

Gene Based Analyses

Additionally, we ran gene-based analyses that produced a number of genes associated with AD sx (detailed in table 13). Amongst the top genes (empirical p-value<0.001) associated with AD sx was gamma-aminobutyric acid receptor subunit gamma-1 (*GABRG1*). Also associated with AD sx in this sample was heat shock protein (*HSPA2). DOK7* was both implicated in the SNP based and gene-based GWAS analyses of AD sx. In further examination of the gene-based analyses of AD sx, we tested whether a particular gene function was overrepresented, or enriched, in this set of highly associated genes for AD sx (empirical p-value<0.001). Gene enrichment analyses of this gene set indicated that no particular function was significantly overrepresented. When the threshold for significance of top gene-based results was relaxed (empirical p-value <0.01), the gene set is significantly enriched for ion channel activity genes. Associated AD sx genes that involve aspects of ion channel activity include *ACCN1, KCNMB1, KCTD3, KCNH1, P2RX1, ITPR2, FXYD6, BEST1, KCNIP1, CACNA1C, BSND, TRPC7, TRPA1, GRID1, GRIN2B, FXYD2*.

Table 13. Genes from VEGAS gene-based analyses of AD sx (corrected p-value<0.001)

Chr	Gene	Corrected p-value	Previous Literature
1	TOR3A	4.31E-04	
1	FAM20B	7.81E-04	
3	LRRN1	8.81E-04	Autism (Davis et al., 2009),
3	C3orf54	9.21E-04	
3	LOC389118	9.27E-04	
3	IHPK1	9.36E-04	Type II Diabetes (Kamimura et al., 2004), Insulin sensitivity (Chakraborty et al., 2010)
4	LRPAP1	1.16E-04	Alzheimer disease (Sanchez et al., 2001), Gallstone disease (Dixit et al., 2006), Degenerative dementia (Pandey et al., 2008)

4	HGFAC	3.20E-05	
4	DOK7	5.00E-05	Congenital myasthenic syndromes (Muller et al., 2007)
4	PRDM8	5.20E-04	Diastolic blood pressure (Newton-Chech, 2009)
4	GABRG1	9.87E-04	Alcohol dependence (Edenberg et al., 2004; Covault et al., 2008; Enoch et al., 2009; Ray et al., 2009; Ittiwut et al., 2012; Wang et al., 2012), Autism (Ma et al., 2005; Kakinuma et al., 2008)
5	PLEKHG4B	1.30E-03	
8	TRPA1	4.09E-04	
11	COPB1	1.37E-03	
11	SCGB2A2	7.86E-04	Breast Cancer (Al-Joudi et al., 2011)
11	SCGB1D4	7.91E-04	Rhinosinusitis (Lu et al., 2011)
11	SCGB1D2	8.13E-04	Breast Cancer (Carter et al., 2002)
13	SLC46A3	8.34E-04	Fatty liver disease (Chalasani et al., 2010)
14	C14orf181	1.30E-04	Type I Diabetes (Reddy et al., 2011)
14	ZFP36L1	2.04E-04	Celiac disease (Dubois et al., 2010), Crohn’s disease (Franke et al., 2010), Multiple sclerosis (Sawcer et al., 2011)
14	LTBP2	2.94E-04	Height (Lango et al., 2010), bone mineral density variation (Cheung et al., 2008), Glacoma (Krumbiegel et al., 2009; Rao et al., 2012),
14	C14orf50	3.20E-05	
14	EXOC5	5.58E-04	Polycystic kidney disease (Fogelgren et al., 2011)
14	HSPA2	6.00E-06	Sensitive to chronic ethanol treatment in mice (Bowers et al., 2009)
14	ZBTB25	9.57E-04	
14	ZBTB1	9.70E-05	Lymphoid development (Siggs et al., 2012)
17	MSL-1	1.33E-03	
17	THRA	1.49E-03	Thyroid cancer (Rasmussen, 2001)
17	ATP2A3	5.42E-04	
17	NR1D1	6.67E-04	Mood disorders and sleep disturbances (Partonen, 2012)
19	ZNF492	2.65E-04	
20	NAT5	1.40E-03	
21	TTC3	6.48E-04	Eye color (Liu et al., 2010),

Genetic Factor Scores

Twin Modeling Results

Multivariate twin analyses produced five latent genetic factors. We focus on two genetic factors of interest: a first genetic factor (Figure 6. A1), which accounts for the genetic variation shared across five measures of alcohol consumption and problems (drinking frequency, drinking quantity, intoxication frequency, maximum drinks/24 hr period, and DSM-IV AD symptoms) and a second genetic factor (Figure 6. A5), that loads exclusively onto DSM-IV AD symptoms. Throughout this manuscript, we will refer to A1 (figure 6) as the *consumption and problems genetic factor* and we will refer to A5 (figure 6) as the *alcohol dependence genetic factor.* The genetic factor scores were significantly related to each other (r^2=0.468) and to AD sx (r^2=0.478 and 0.928 for the *consumption and problems genetic factor* and the *alcohol dependence genetic factor* respectively), with the strongest relationship existing between AD sx and the *alcohol dependence genetic factor*, as would be expected. In addition, the *general consumption and problems genetic factor* was more related to adolescent alcohol consumption, DSM-IV Conduct Disorder symptoms, and age 22 smoking frequency than either AD sx or the *alcohol dependence genetic factor* (Table 14). AD sx were more related to adolescent alcohol consumption, DSM-IV Conduct Disorder symptoms, and age 22 smoking frequency, and DSM-IV Adult Antisocial Behavior Symptoms than the *alcohol dependence genetic factor* (Correlations detailed in table 14). These correlations confirm two assumptions. First, the *general consumption and problems genetic factor* represents the genetic variance captured across five measures of alcohol consumption and problems that is related to frequency of alcohol (and related substance, tobacco) use and the *alcohol dependence genetic factor* represents the genetic variance that is related to AD sx. Second, the *alcohol dependence genetic factor* is somewhat less related to general frequency of alcohol (and related behaviors/disorders) than AD sx, as the variance shared with measures of consumption is (theoretically) removed from this genetic factor.

Figure 6. Genetic Architecture of Measures of Alcohol Consumption and Problems in *FinnTwin12*

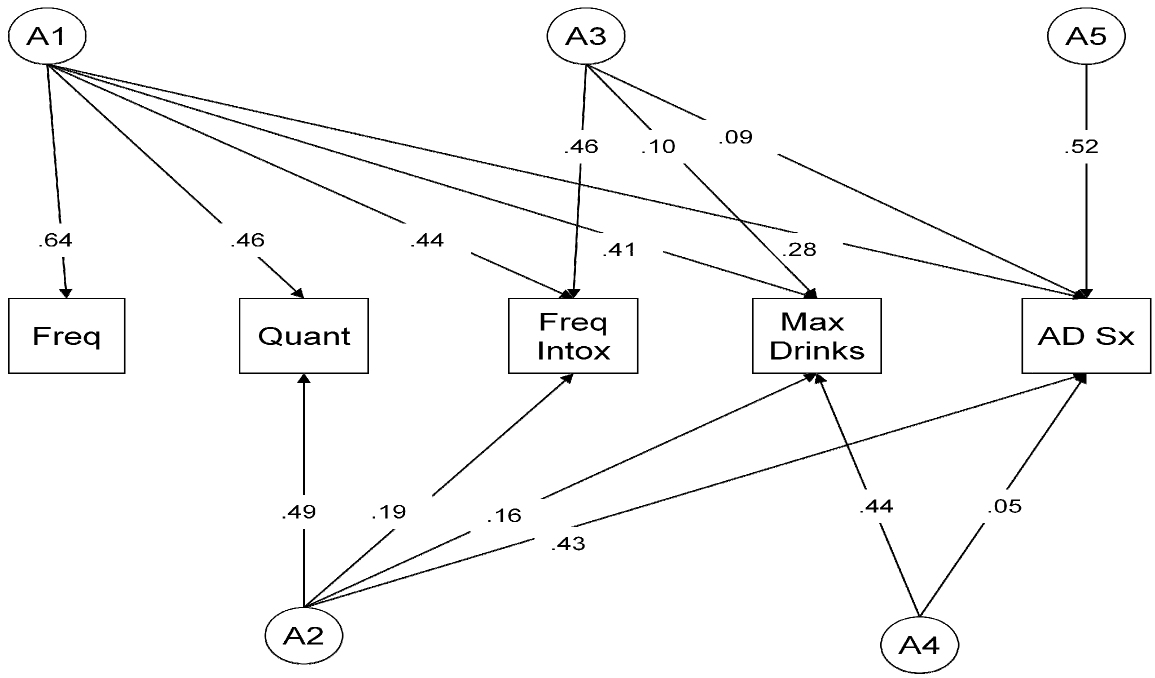


Table 14. Phenotypic Correlations between AD symptoms, the *general consumption and problems genetic factor*, and the *alcohol dependence genetic factor* scores (yielded from twin data), and Related Outcomes

	AD Symptoms	*Consumption and Problems Genetic Factor*	*Alcohol Dependence Genetic Factor*
AD symptoms	1		
Consumption and Problems Genetic Factor	.478**	1	
Alcohol dependence Genetic Factor	.928**	.468**	1
Related Outcomes			
Age 14 Drinking Frequency	.076**	.155**	.060
Age 17 Drinking Frequency	.115**	.437**	.098*
Age 22 Drinking Frequency	.274**	.695**	.200**
Age 22 Smoking Frequency	.198**	.362**	.162**
Conduct Disorder Sx	.260**	.271**	.207**
Antisocial Behavior Sx	.379**	.380**	.348**

**Pearson correlation significant at a p<0.01 *Pearson correlation significant at a p<0.01

Comparing GWAS Results from AD sx and Genetic Factors

SNP Based Analyses

The correlation between the SNP based genetic association results (log of the p-values) for AD sx and *the consumption and problems genetic factor* was 0.103. The correlation between the SNP based genetic association results (log of the p-values) for AD sx and the *alcohol dependence genetic factor was* .514 (table 15).

Under the assumption that SNPs associated with all outcomes may represent the most robust results, we compared associated SNPs across the three alcohol phenotypes. Three individual SNPs, that were significantly associated with AD sx (FDR<10%), were also associated with the genetic factor scores (p-value<0.05). Two of these three SNPs reside in genes: *UPK1B/TSPAN20* and *DOK7* (table 16).

Table 15. Pearson Correlation between SNP level results (log of p-values) for AD sx and the Genetic Factors

P-value Results	AD Symptoms	*Consumption and Problems Genetic Factor*	*Alcohol Dependence Genetic Factor*
AD symptoms	1		
Genetic Factor Scores	---	---	---
Consumption and Problems Genetic Factor	.103**	1	
Alcohol dependence Genetic Factor	.514**	.086**	1

Table 16. Top (FDR (BH) less than 10%) AD sx SNPs Also Associated with Genetic Factor Scores (p-value<0.05)

Chr	Gene	SNP	BP	Corrected p-value		
				Alcohol Dependence Sx	*Consumption and Problems Genetic Factor*	*Alcohol Dependence Genetic Factor*
3	Intergenic	rs9310823	26933421	6.03E-05	9.79E-03	9.80E-03
3	UPK1B/TSPAN20	rs6797796	120387620	3.71E-06	7.20E-02	3.15E-06
4	DOK7	rs10022329	3437317	6.02E-07	2.63E-03	4.49E-05
4	Intergenic	rs13136935	3466738	6.16E-05	4.68E-03	2.55E-04
4	DOK7	rs7680504	3468951	5.71E-05	4.64E-03	3.17E-04
4	Intergenic	rs11934116	81369133	2.20E-05	1.46E-02	1.88E-03
4	Intergenic	rs2033613	142315874	2.32E-05	1.08E-03	3.09E-04
19	Intergenic	rs7246529	22856963	2.53E-06	2.83E-02	2.18E-05
19	Intergenic	rs12460438	22867619	2.53E-06	2.83E-02	2.18E-05

Note: Boxes indicate that SNPs are in high Linkage Disequilibrium (r2>.8); Dashed Box indicates that SNPs are in moderate LD (r2>.5)

Gene Based Analyses

Additionally, we compared associated genes across the three alcohol phenotypes. Below, we have presented venn diagrams depicting the overlap in gene sets, consisting of genes that passed a relaxed gene-based significance threshold (p<0.01), for AD sx and the genetic factors in Figure 7. Below each diagram, we have presented the associated Jaccard coefficient (J) a statistic that assesses the similarity between gene-sets. Results indicate a larger degree of overlap between the AD sx gene-set and the *alcohol dependence genetic factor* gene-set (J= 0.25) then between the AD sx gene-set and *the consumption and problems genetic factor* gene-set (J= 0.07). Of all genes highly associated with AD sx (p-value<0.001), four genes were significantly associated with both genetic factors (p-value<0.05). These include three genes on chromosome 14: *C14orf181* and *ZFP36L1/Brfn1,* and *LTBP2* and one gene on chromosome 13, *SLC46A3*.

Figure 7.Top Gene Results (p-value<0.01) and Overlap for AD symptoms and two genetic factor scores, *General Consumption and Problems* and *Alcohol* Problems.

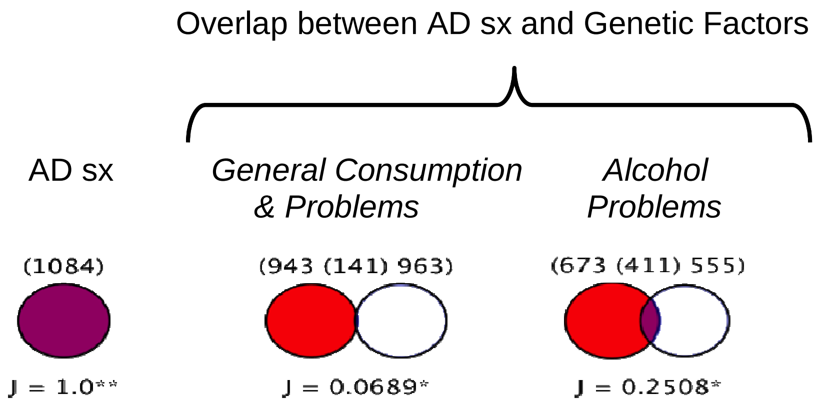


Table 17. Top Genes from VEGAS gene-based analyses of AD sx also associated with the genetic factor scores (corrected p-value <0.05)

Chr	**Gene**	**p-value**	***Consumption and Problems Genetic Factor***	***Alcohol Dependence Genetic Factor***	**Previous Literature**
13	SLC46A3	8.30E-04	1.63E-02	7.90E-04	Fatty liver disease (Chalasani et al., 2010)
14	C14orf181	1.30E-04	3.60E-03	1.50E-03	Type I Diabetes (Reddy et al., 2011)
14	ZFP36L1/ Bfn1	2.00E-04	4.68E-03	2.54E-03	Celiac disease (Dubois et al., 2010), Crohn's disease (Franke et al., 2010), Multiple sclerosis (Sawcer et al., 2011)
14	LTBP2	2.90E-04	3.28E-02	3.30E-03	Height (Lango et al., 2010), bone mineral density variation (Cheung et al., 2008), Glacoma (Krumbiegel et al., 2009; Rao et al., 2012),

Discussion

To date, several Genome Wide Association Studies (GWAS) on AD have been published without producing robust, replicable genetic association signals. In the present study, we conducted a GWAS on DSM-IV Alcohol Dependence symptoms in a Finnish population based sample of twins. No individual SNP met the genome-wide threshold of significance (8.89 $x10^{-8)}$, however many SNPs were approaching this threshold. In addition to analyzing our primary phenotype of interest, AD sx, we analyzed two genetic factor scores that emerged from the multivariate twin analysis. We believe that there are several interesting observations to make regarding the results from this study.

AD GWAS

The most striking novel genetic association result from this study is Docking Protein 7 (*DOK7)* SNP rs10022329. rs10022329 is both the most significant individual SNP result and resides in the most highly associated gene (from the gene-based tests), Docking Protein 7 (*DOK7)*. *DOK7* is essential for neuromuscular synaptogenesis and mutations in this gene are a cause of familial limb-girdle myasthenia autosomal recessive, which is also known as congenital myasthenic syndrome type 1B (Muller et al., 2007). Of the 18 *DOK7* SNPs available on the Illumina Platform, five independent signals are represented. Each of the independent signals in *DOK7* was associated (p-value<0.001) with AD sx. When considered as a set, the association between the 18 *DOK7* SNPs and AD sx was highly significant (empirical p= 7.9E-06). Also of note is SNP based association result rs531855, which resides in domain-containing ion transport regulator 6 (FXYD6). *FXYD6* belongs to the FXYD family of ion transport regulators and has previously been associated with schizophrenia (Choudhury et al., 2007, Ito et al., 2008). The *FXYD6* gene encodes the protein phosphohippolin (Kadowaki et al., 2004), which is highly expressed in regions of the brain likely involved in schizophrenia. *FXYD6* is also moderately associated (nominal p-value<0.01) with alcohol dependence in the COGA study of severely affected alcohol dependence cases (Edenberg et al., 2010).

Gene-based analyses produced several interesting genes associated with AD sx, including gamma-aminobutyric acid receptor subunit gamma-1 (*GABRG1)*. *GABRG1* belongs to the ligand-gated ionic channel family and plays an important role in inhibiting neurotransmission by binding to the benzodiazepine receptor and opening an integral chloride channel. $GABA_A$ receptors have been implicated in biological processes related to the acute and chronic effects of alcohol (Koob et al., 2004; Krystal et al., 2006) *GABRG1* has previously been associated with alcohol dependence in several studies (Edenberg et al., 2004; Covault et al., 2008; Enoch et al., 2009; Ray et al., 2009; Ittiwut et al., 2012; Wang et al., 2012). *GABRG1* has also been previously associated with autism (Ma et al., 2005; Kakinuma et al., 2008). Another gene-based association result of note is heat shock 70k Da protein 2 (*HSPA2)* which has been previously associated with alcoholic pancreatitis in Korean patients (Lee et al., 2007). In cooperation with other chaperones, *HSPA2* stabilizes preexistent proteins against aggregation and mediate the folding of newly translated polypeptides. They bind extended peptide segments during translation and membrane translocation, or following stress-induced damage (Bonnycastle et al., 1994). Gene enrichment analyses indicated that the AD sx gene-set is significantly enriched for ion channel activity genes. Associated AD sx genes that involve aspects of ion channel activity include *ACCN1, KCNMB1, KCTD3, KCNH1, P2RX1, ITPR2, FXYD6, BEST1, KCNIP1, CACNA1C, BSND, TRPC7, TRPA1, GRID1, GRIN2B, FXYD2*. Ion channel activity has previously been linked to alcohol dependence in humans (Lind et al, 2010) and in ethanol responsiveness in model systems (Bettinger et al., 2012). Gene enrichment analyses performed on the top signals from an Australian case/control study of AD (Lind et al., 2010) also indicated that that ion-channel activity genes were overrepresented. A recent study in *caenorhabditis elegans found* that genetic alterations in this gene can modify the phenotype of gain-of-function mutations in the ethanol-inducible ion channel SLO-1 (Bettinger et al., 2012).

Comparing GWAS Results from AD sx and Genetic Factors

While the phenotypic correlations between AD sx and the genetic factors were strong (.478-.978; table 14), the relationship between the genetic association results was significantly weaker, on both the SNP level (r=.103-.514; table 15) and gene level (J=. 06-.25; figure 7). The high phenotypic correlation between the commonly used AD sx and the *alcohol dependence genetic factor* suggests that the AD sx is comparable to the genetic factor implicated by twin modeling. The nominally higher correlations between AD sx and related externalizing outcomes (adolescent alcohol consumption, DSM-IV Conduct Disorder symptoms, age 22 smoking frequency, DSM-IV Adult Antisocial Behavior Symptoms) than with the *alcohol dependence genetic factor* may indicate that use of the genetic factor score is reducing the variance shared between alcohol consumption and problems. However, no substantive advantage of the *alcohol dependence genetic factor* over AD sx is noted. More striking is the discordance between the genetic association results for AD sx and the genetic factors. In this sample, both SNP based and gene-based results suggest that there are a small proportion of genetic variants shared across these three phenotypes, but the majority of variants are unique to each outcome. This discordance between GWAS results from AD sx and the genetic factors underscores the difficultly in replicating genetic effects and in differentiating real findings from spurious ones. The variability in genetic association results for highly correlated phenotypes suggests that convergence in results across phenotypes, methods, and samples may provide us the most robust genetic association signals.

If we operate under the assumption that genetic variants associated with AD sx and the genetic factor scores are the most robust results, there are six independent SNP signals that stand out. Two of these SNPs reside in genes: *UPK1B*/*TSPAN20* and *DOK7*. *UPK1B* encodes the Uroplakin 1B protein, a member of the tetraspanin family. These proteins mediate signal transduction events in the regulation of cell development, activation, growth and motility (Olsburgh et al., 2002). Prior studies suggest a link between *UPK1B* and bladder function (Kalma et al., 2009). The converging evidence of association between SNPs in *DOK7,* AD sx, and both genetic factors, lends further support to this genetic association result. Results from comparisons of the gene-based tests indicated that four genes were associated with AD sx and the genetic factors. Butyrate response factor 1 (*ZFP36L1/Brf1)* is a member of the TIS11 family of early response genes, which are induced by various agonists (Hacker et al., 2010). A 2010 study reported that chronic alcohol administration in mice leads to enhanced expression of Brf1 in the liver (Zhong et al., 2010). Chromosome 14 open reading frame 181

(*C14orf181)* has previously been implicated in Type I Diabetes (Reddy et al., 2011). Latent transforming growth factor beta binding protein 2 (*LTBP2)* belongs to the family of latent transforming growth factor (TGF)-beta binding proteins (LTBP) and may be involved in cell adhesion (Vehviläinen et al., 2003). *LTBP2* has previously been associated with height (Lango et al., 2010), bone mineral density variation (Cheung et al., 2008), and glacoma (Krumbiegel et al., 2009; Rao et al., 2012). The fourth gene associated with all three phenotypes is solute carrier family 46, member 3 (*SLC46A3)*. *SLC46A3* has been implicated in non-alcoholic fatty liver disease (Chalasani et al., 2010). Note that none of these SNPs based or gene-based variants have been previously associated with alcohol dependence.

There are several limitations of this study to consider. Most notable is the small sample size and subsequent lack of power to detect association for alcohol phenotypes at a genome wide threshold. Several studies have demonstrated that very large sample sizes are required to detect the subtle genetic influences thought to be acting on complex behavioral phenotypes such as alcohol use/problems (Risch & Merikangas, 1996). Relatedly, the power of this sample was further diminished by the constraints of the statistical program, qfam that was implemented to conduct the GWAS. This program can only specify one type of familial relationship (in this sample the relationship between dizygotic twins), thus the second MZ twin was not included and the sample size was diminished. In an effort to test the effect this limitation had on the genetic association results, we re-analyzed the data using two different statistical packages in R that specify both MZ and DZ relationships, GENABLE (Aulchenko et al., 2007) and GWAF (Chen et al., 2004). While GWAF uses a kinship matrix to specify the genetic relationship between twins (imposing 1 for MZ's and 0.5 for DZ's), GENABLE empirically establishes the genetic relationship between twins. Future work is necessary to ensure that these methods handle the genetic relationship between MZ twins adequately. In examining the correlation between the log of the p-values associated with each SNP and AD sx from each package, the results were strongly related. The correlation between results from qfam (PLINK) and GWAF was 0.94, the correlation between results from qfam (PLINK) and GENABLE was 0.96. The correlation between results from GWAF and GENABLE was 0.94. We believe that this demonstrates that the power lost from the use of qfam did not substantively affect the genetic association results, however future work should compare specific genetic variants associated with AD sx produced by each program. In addition, sex differences for genetic risk factors were not formally tested for in the context of the GWAS analyses. Because the best fitting twin model implicated that male and female alcohol phenotypes could be collapsed without a significant decrease to model fit, we analyzed males and females together in the genetic association analyses. However, collapsing male and female data may introduce further heterogeneity into the phenotype. Another potential limitation of this study was choice of statistical model. In this manuscript, we chose to use a cholesky decomposition model. Preliminary model fitting suggested that the fit of the cholesky to this data was an improvement on the independent and common pathway models. However, a comparison of GWAS results from alternative genetic factors should be carried out in future studies.

In summary, this study has provided modest evidence of association between AD sx and several novel genetic variants (both SNPs and genes) that approach genome wide significance, including *DOK7*, which was implicated in both SNP and gene-based analyses conducted in a Finnish population based sample. In addition, we have replicated a previously reported genetic association between *GABRG1* and alcohol dependence. Each of the genetic variants presented in this study should be replicated in an independent sample with comparable phenotypic measurement of DSM-IV alcohol dependence symptoms. Finally, discordance between genetic association results from AD sx and the genetic factor scores illustrates the inconsistency of GWAS results for complex psychiatric phenotypes. Harmonization of phenotypes and methods across comparable study designs is likely to result in the most robust genetic association signals.

Supplemental Tables

Table 18. Top SNP Results for the *General Consumption and Problems Genetic Factor*

CHR	GENE	SNP	Lit	P-value
2	DOCK10	rs12469757**	Cancers (Yelo et al., 2008)	1.12E-05
2		rs11688439**		1.16E-05
2	NRXN1	rs10490175	Alcoholism (Yang et al., 2005), Nicotine dependence (Bierut et al., 2007), Sz (Moore et al., 2011), Autism (Hedges et al., 2012)	4.18E-05
3	BCHE	rs7429483	ADHD (Lesch et al., 2009); Az (Atack et al., 1985)	8.77E-07
3		rs1587425		1.18E-05
3		rs1909526		1.37E-05
3		rs11921615		2.04E-05
3	SPATA16	rs506433	Male infertility (Dam et al., 2008)	2.07E-05
4		rs7657618		1.12E-06
4	PDGFC	rs4691381	Speech perception in dyslexia (Roeske et al., 2009)	5.70E-06
4		rs17035181		1.14E-05
4	DOCK10	rs983473	Cancers (Yelo et al., 2008)	1.16E-05
4	DOCK10	rs1816164	Cancers (Yelo et al., 2008)	1.16E-05
4		rs1907091		1.26E-05
4		rs17036640		2.40E-05
4	CYP4V2	rs13146272*	Bietti Crustalline dystrophy (Okialda et al., 2012)	2.48E-05
5		rs1560919		4.57E-06
5	FSTL4	rs10515460	Ischemic stroke (Luke et al., 2008)	1.88E-05
5	FSTL4	rs17166631	Ischemic stroke (Luke et al., 2008)	1.88E-05
7		rs10244707		2.60E-05
8	SLC7A2	rs13270915	Hyperthyroidism	2.37E-06
8	SLC7A2	rs13252649	Hyperthyroidism	3.03E-06
9		rs6476012		1.69E-05
9		rs8181181		2.40E-05
11		rs4075242		4.67E-06
13		rs974288**		7.32E-06
16		rs1437169		2.20E-05
20	PLCB1	rs6056006	Schizophrenia, Depression (Vasco et al., 2012)	8.83E-07
20	PLCB1	rs2295179	Schizophrenia , Depression (Vasco et al., 2012)	4.22E-06
20		rs6056230		2.46E-05

*Nominally significant association (p<0.01) with AD sx in the Collaborative Study on the Genetics of Alcoholism GWAS **Nominally significant association (p<0.01) with max drinks phenotype (maximum drinks in a 24 hr period) in the Collaborative Study on the Genetics of Alcoholism GWAS

Table 19. Top SNP Results for *the Alcohol Problems Genetic Factor*

CHR	GENE	SNP	Lit	P-value
1		rs10046065		2.26E-05
1	ST6GALNAC5	rs12461092	cancers (Oster et al., 2011)	2.67E-05
1	AMPD3	rs7587040	skeletal muscle changes (Fortuin et al., 1998),	2.87E-05
2		rs9355980		3.2E-06
2		rs11629182		1.31E-05
2		rs9459056		1.39E-05
3		rs7587040		3.1E-06
3	HMGB1	rs531855		1.11E-05
3		rs7616907		2.57E-05
3		rs978743		2.57E-05
3		rs2664904		2.63E-05
5	RASGRF1	rs1283924	myopia (Hysi et al., 2010)	3.8E-06
5		rs17027082		0.000023
6	PARK2	rs17214843	parkinson' s disease (Matsumine et al., 1998)	3.5E-06
6		rs1283926		5.7E-06
6		rs10454559		5.7E-06
6		rs995085		7.9E-06
6	SYNJ2	rs12584812		1.49E-05
6	PARK2	rs7756400	parkinson' s disease (Matsumine et al.	2.79E-05
6	PARK2	rs2269340	parkinson' s disease (Matsumine et al.	2.79E-05
10		rs2392038		2.66E-05
11	FXYD6	rs6589624*	schizophrenia (Jiao et al., 2011)	1.12E-05
11	FXYD6	rs7563569	schizophrenia (Jiao et al., 2011)	1.12E-05
13		rs2025641		6.8E-06
13		rs1947238		9.6E-06
13		rs9578135		1.67E-05
13		rs8075075		1.67E-05
14		rs13035719		1.31E-05
17		rs872387		1.76E-05
17		rs12460438		1.76E-05
19		rs7246529		2.18E-05
19		rs7544426*		2.18E-05
19		rs2023053		0.000027

*Nominally significant association (p<0.01) with AD sx in the Collaborative Study on the Genetics of Alcoholism GWAS

Table 20. Alternative Twin Model Fit Statistics

Model	***-2 times LL***	***DF***	***AIC***	***BIC***
(1) Independent Pathway	24832.923	6026	7019.077	-16301.402
(5) Cholesky Decomposition	17677.392	6591	4495.392	-15655.054
(1) Common Pathway	20524.380	6684	7156.380	-14577.171
(2) Independent Pathway	24825.650	6091	6956.350	-16194.384
(3) Cholesky Decomposition	16773.936	6601	4767.936	-10387.851
(3) Common Pathway	21245.490	6558	7211.490	-14997.122

References

1. Bierut LJ, Agrawal A, Bucholz KK, Doheny KF, Laurie C, Pugh E, Fisher S, Fox L, Howells W, Bertelsen S, Hinrichs AL, Almasy L, Breslau N, Culverhouse RC, Dick DM, Edenberg HJ, Foroud T, Grucza RA, Hatsukami D, Hesselbrock V, Johnson EO, Kramer J, Krueger RF, Kuperman S, Lynskey M, Mann K, Neuman RJ, Nöthen MM, Nurnberger JI Jr, Porjesz B, Ridinger M, Saccone NL, Saccone SF, Schuckit MA, Tischfield JA, Wang JC, Rietschel M, Goate AM, Rice JP; Gene, Environment Association Studies Consortium. A genome-wide association study of alcohol dependence. Proc Natl Acad Sci U S A. 2010 Mar 16;107(11):5082-7.

2. Zuo L, Zhang CK, Wang F, Li CS, Zhao H, Lu L, Zhang XY, Lu L, Zhang H, Zhang F, Krystal JH, Luo X. A novel, functional and replicable risk gene region for alcohol dependence identified by genome-wide association study. PLoS One. 2011;6(11):e26726.

3. Wang KS, Liu X, Zhang Q, Wu LY, Zeng M. Genome-wide association study identifies 5q21 and 9p24.1 (KDM4C) loci associated with alcohol withdrawal symptoms. J Neural Transm. 2012 Apr;119(4):425-33.

4. Wang KS, Liu X, Aragam N, Jian X, Mullersman JE, Liu Y, Pan Y. Family-based association analysis of alcohol dependence in the COGA sample and replication in the Australian twin-family study. J Neural Transm. 2011 Sep;118(9):1293-9.

5. Frank J, Cichon S, Treutlein J, Ridinger M, Mattheisen M, Hoffmann P, Herms S, Wodarz N, Soyka M, Zill P, Maier W, Mössner R, Gaebel W, Dahmen N, Scherbaum N, Schmäl C, Steffens M, Lucae S, Ising M, Müller-Myhsok B, Nöthen MM, Mann K, Kiefer F, Rietschel M. Genome-wide significant association between alcohol dependence and a variant in the ADH gene cluster. Addict Biol. 2012 Jan;17(1):171-80.

6. Baik I, Cho NH, Kim SH, Han BG, Shin C. Genome-wide association studies identify genetic loci related to alcohol consumption in Korean men. Am J Clin Nutr. 2011 Apr;93(4):809-16.

7. Schumann G, Coin LJ, Lourdusamy A, Charoen P, Berger KH, Stacey D, Desrivières S, Aliev FA, Khan AA, Amin N, Aulchenko YS, Bakalkin G, Bakker SJ, Balkau B, Beulens JW, Bilbao A, de Boer RA, Beury D, Bots ML, Breetvelt EJ, Cauchi S, Cavalcanti-Proença C, Chambers JC, Clarke TK, Dahmen N, de Geus EJ,

Dick D, Ducci F, Easton A, Edenberg HJ, Esko T, Fernández-Medarde A, Foroud T, Freimer NB, Girault JA, Grobbee DE, Guarrera S, Gudbjartsson DF, Hartikainen AL, Heath AC, Hesselbrock V, Hofman A, Hottenga JJ, Isohanni MK, Kaprio J, Khaw KT, Kuehnel B, Laitinen J, Lobbens S, Luan J, Mangino M, Maroteaux M, Matullo G, McCarthy MI, Mueller C, Navis G, Numans ME, Núñez A, Nyholt DR, Onland-Moret CN, Oostra BA, O'Reilly PF, Palkovits M, Penninx BW, Polidoro S, Pouta A, Prokopenko I, Ricceri F, Santos E, Smit JH, Soranzo N, Song K, Sovio U, Stumvoll M, Surakk I, Thorgeirsson TE, Thorsteinsdottir U, Troakes C, Tyrfingsson T, Tönjes A, Uiterwaal CS, Uitterlinden AG, van der Harst P, van der Schouw YT, Staehlin O, Vogelzangs N, Vollenweider P, Waeber G, Wareham NJ, Waterworth DM, Whitfield JB, Wichmann EH, Willemsen G, Witteman JC, Yuan X, Zhai G, Zhao JH, Zhang W, Martin NG, Metspalu A, Doering A, Scott J, Spector TD, Loos RJ, Boomsma DI, Mooser V, Peltonen L, Stefansson K, van Duijn CM, Vineis P, Sommer WH, Kooner JS, Spanagel R, Heberlein UA, Jarvelin MR, Elliott P. Genome-wide association and genetic functional studies identify autism susceptibility candidate 2 gene (AUTS2) in the regulation of alcohol consumption. Proc Natl Acad Sci U S A. 2011 Apr 26;108(17):7119-24.

8. Grant JD, Agrawal A, Bucholz KK, Madden PA, Pergadia ML, Nelson EC, Lynskey MT, Todd RD, Todorov AA, Hansell NK, Whitfield JB, Martin NG, Heath AC. Alcohol consumption indices of genetic risk for alcohol dependence. Biol Psychiatry. 2009 Oct 15;66(8):795-800.

9. Kendler KS, Myers J, Dick D, Prescott CA. The relationship between genetic influences on alcohol dependence and on patterns of alcohol consumption. Alcohol Clin Exp Res. 2010 Jun;34(6):1058-65.

10. Dick DM, Meyers JL, Rose RJ, Kaprio J, Kendler KS. Measures of current alcohol consumption and problems: two independent twin studies suggest a complex genetic architecture. Alcohol Clin Exp Res. 2011 Dec;35(12):2152-61.

11. Meyers JL, Nyman E, Loukola A, Rose RJ, Kaprio J, Dick DM. The Association between *DRD2/ANKK1* and Genetically Informed Measures of Alcohol Use and Problems. Under review in Addiction Biology

12. Kaprio J, Pulkkinen L, Rose RJ. Genetic and environmental factors in health-related behaviors: studies on Finnish twins and twin families. Twin Res. 2002 Oct;5(5):366-71. Review.

13. American Psychological Association, 1994

14. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ & Sham PC (2007) PLINK: a toolset for whole-genome association and population-based linkage analysis. American Journal of Human Genetics, 81.

15. Liu JZ, McRae AF, Nyholt DR, Medland SE, Wray NR, Brown KM, AMFS Investigators, Hayward NK, Montgomery GW, Visscher PM, Martin NG, MacGregor S. (2010). A Versatile Gene-Based Test for Genome-wide Association Studies. *American Journal of Human Genetics*, 87.

16. Müller JS, Herczegfalvi A, Vilchez JJ, Colomer J, Bachinski LL, Mihaylova V, Santos M, Schara U, Deschauer M, Shevell M, Poulin C, Dias A, Soudo A, Hietala M, Äärimaa T, Krahe R, Karcagi V, Huebner A, Beeson D, Abicht A, Lochmüller H. Phenotypical spectrum of DOK7 mutations in congenital myasthenic syndromes. Brain. 2007 Jun;130(Pt 6):1497-506.

17. Yamamoto M, Iguchi G, Takeno R, Okimura Y, Sano T, Takahashi M, Nishizawa H, Handayaningshi AE, Fukuoka H, Tobita M, Saitoh T, Tojo K, Mokubo A, Morinobu A, Iida K, Kaji H, Seino S, Chihara K, Takahashi Y. Adult combined GH, prolactin, and TSH deficiency associated with circulating PIT-1 antibody in humans. J Clin Invest. 2011 Jan;121(1):113-9. doi: 10.1172/JCI44073. Epub 2010 Dec 1.

18. Ito Y, Nakamura Y, Takahashi N, Saito S, Aleksic B, Iwata N, Inada T, Ozaki N. A genetic association study of the FXYD domain containing ion transport regulator 6 (FXYD6) gene, encoding phosphohippolin, in susceptibility to schizophrenia in a Japanese population. Neurosci Lett. 2008 Jun 13;438(1):70-5.

19. Goode EL, Chenevix-Trench G, Song H, Ramus SJ, Notaridou M, Lawrenson K,

Widschwendter M, Vierkant RA, Larson MC, Kjaer SK, Birrer MJ, Berchuck A, Schildkraut J, Tomlinson I, Kiemeney LA, Cook LS, Gronwald J, Garcia-Closas M, Gore ME, Campbell I, Whittemore AS, Sutphen R, Phelan C, Anton-Culver H, Pearce CL, Lambrechts D, Rossing MA, Chang-Claude J, Moysich KB, Goodman MT, Dörk T, Nevanlinna H, Ness RB, Rafnar T, Hogdall C, Hogdall E, Fridley BL, Cunningham JM, Sieh W, McGuire V, Godwin AK, Cramer DW, Hernandez D, Levine D, Lu K, Iversen ES, Palmieri RT, Houlston R, van Altena AM, Aben KK, Massuger LF, Brooks-Wilson A, Kelemen LE, Le ND, Jakubowska A, Lubinski J, Medrek K, Stafford A, Easton DF, Tyrer J, Bolton KL, Harrington P, Eccles D, Chen A, Molina AN, Davila BN, Arango H, Tsai YY, Chen Z, Risch HA, McLaughlin J, Narod SA, Ziogas A, Brewster W, Gentry-Maharaj A, Menon U, Wu AH, Stram DO, Pike MC; Wellcome Trust Case-Control Consortium, Beesley J, Webb PM; Australian Cancer Study (Ovarian Cancer); Australian Ovarian Cancer Study Group; Ovarian Cancer Association Consortium (OCAC), Chen X, Ekici AB, Thiel FC, Beckmann MW, Yang H, Wentzensen N, Lissowska J, Fasching PA, Despierre E, Amant F, Vergote I, Doherty J, Hein R, Wang Gohrke S, Lurie G, Carney ME, Thompson PJ, Runnebaum I, Hillemanns P, Dürst M, Antonenkova N, Bogdanova N, Leminen A, Butzow R, Heikkinen T, Stefansson K, Sulem P, Besenbacher S, Sellers TA, Gayther SA, Pharoah PD; Ovarian Cancer Association Consortium (OCAC). A genome-wide association study identifies susceptibility loci for ovarian cancer at 2q31 and 8q24. Nat Genet. 2010 Oct;42(10):874-9.

20. Bowers BJ, Radcliffe RA, Smith AM, Miyamoto-Ditmon J, Wehner JM. Microarray

analysis identifies cerebellar genes sensitive to chronic ethanol treatment in PKCgamma mice. Alcohol. 2006 Aug;40(1):19-33.

21. Siggs OM, Li X, Xia Y, Beutler B. ZBTB1 is a determinant of lymphoid development. J Exp Med. 2012 Jan 16;209(1):19-27.

22. Pandey P, Pradhan S, Mittal B. LRP-associated protein gene (LRPAP1) and susceptibility to degenerative dementia. Genes Brain Behav. 2008 Nov;7(8):943-50.

23. Sánchez L, Alvarez V, González P, González I, Alvarez R, Coto E. Variation in the LRP-associated protein gene (LRPAP1) is associated with late-onset Alzheimer disease. Am J Med Genet. 2001 Jan 8;105(1):76-8.

24. Dixit M, Choudhuri G, Keshri LJ, Mittal B. Association of low density lipoprotein receptor related protein-associated protein (LRPAP1) gene insertion/deletion polymorphism with gallstone disease. J Gastroenterol Hepatol. 2006 May;21(5):847-9.

25. Reddy MV, Wang H, Liu S, Bode B, Reed JC, Steed RD, Anderson SW, Steed L, Hopkins D, She JX. Association between type 1 diabetes and GWAS SNPs in the southeast US Caucasian population. Genes Immun. 2011 Apr;12(3):208-12.

26. Chen MH, Yang Q. GWAF: an R package for genome-wide association analyses

with family data. Bioinformatics. 2010 Feb 15;26(4):580-1.

27. Aulchenko YS, Struchalin MV, van Duijn CM. ProbABEL package for genome-wide association analysis of imputed data. BMC Bioinformatics. 2010 Mar 16;11:134.

28. Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science. 1996 Sep 13;273(5281):1516-7.

29. Kadowaki K, Sugimoto K, Yamaguchi F, Song T, Watanabe Y, Singh K, Tokuda M.Phosphohippolin expression in the rat central nervous system. Brain Res Mol Brain Res. 2004 Jun 18;125(1-2):105-12.

30. Koob GF (2004). A role for GABA mechanisms in the motivational effects of alcohol. *Biochem Pharmacol* **68**: 1515–1525.

31. Krystal JH, Staley J, Mason G, Petrakis IL, Kaufman J, Harris RA *et al* (2006). Gamma-aminobutyric acid type A receptors and alcoholism: intoxication, dependence, vulnerability, and treatment. *Arch Gen Psychiatry* **63**: 957–968.

32. Bonnycastle LL, Yu CE, Hunt CR, Trask BJ, Clancy KP, Weber JL, Patterson D, Schellenberg GD. 1994. Cloning, sequencing, and mapping of the human chromosome 14 heat shock protein gene (HSPA2). *Genomics* **23**: 85–93.

33. Olsburgh J, Weeks R, Selby P, Southgate J. Human uroplakin lb gene structure and

promoter analysis. Biochim Biophys Acta. 2002 Jun 7;1576(1-2):163-70.

34. Hacker C, Valchanova R, Adams S, Munz B. ZFP36L1 is regulated by growth factors and cytokines in keratinocytes and influences their VEGF production. Growth Factors. 2010 Jun;28(3):178-90.

35. Vehviläinen P, Hyytiäinen M, Keski-Oja J. Latent transforming growth factor-beta-binding protein 2 is an adhesion protein for melanoma cells. J Biol Chem. 2003 Jul 4;278(27):24705-13.

Chapter 4

Environmental Moderation of Alcohol Use in Adolescence: Common and/or Unique Influences

*****This chapter is adapted from the following manuscript:***
Meyers JL, Latendresse SJ, Pulkkinen L, Korhonen T, Rose D, Kaprio J, Dick DM. Environmental Moderation of Alcohol Use in Adolescence: Common and/or Unique Influences. Under review in Alcohol Clin Exp Res.

Abstract

Background: There is an emerging literature documenting how specific environmental factors moderate the importance of genetic effects on substance use and related behaviors. In previous Finnish twin studies, we have found genetic influence on adolescent substance use to be enhanced in environments characterized by *lower levels of* parental monitoring and *higher levels of* deviant peer behavior. It remains unclear whether these findings reflect a shared process, whereby both factors are reflecting general environmental risk that creates a social opportunity for adolescents to express genetic dispositions to problematic behavior, or whether there are unique contributions of these respective environmental factors.

Methods: In this study, we follow-up on our previous findings (parental knowledge and peer deviance), and test another potential environment of importance, frequency of family dinner, as a moderator of etiological factors influencing frequency of alcohol use at ages 14 and 17. Our dataset included 4,236 Finnish twins followed longitudinally. We compared moderation effects at the level of shared variance, encompassing what is common across these three variables, to the residual sources of variance specific to

each variable. We use the longitudinal study design to explore the relationship between these environmental moderators and behavioral outcomes across the span of adolescence.

Results: All three environmental variables played a moderating role on the importance of genetic and environmental influences on adolescent alcohol use, both jointly, through common variance, and uniquely, through residual sources specific to each.

Conclusions: There are both common and unique moderation effects associated with family and peer factors. The moderating effects associated with the common variance may conceptually map onto an overarching, shared mechanism of social opportunity/control. However, there is also important and distinct information captured in the variance unique to each individual environmental moderator. The moderating effects associated with familial context (parental knowledge residual and frequency of family dinner residual) were more robust in early adolescence, whereas the moderating effects unique to the peer deviance residual persist throughout adolescence.

Introduction

Alcohol use and alcohol-related disorders are known to be under considerable genetic influence (Goldman 1993; Kendler et al. 1995; Tsuang et al. 2001). However, there is growing recognition that static measures of heritability may mask important changes in the relevance of genetic influences as a function of the environment. Many specific environments have been demonstrated to moderate the magnitude of genetic influences on individual variation in alcohol use. The earliest illustration of genetic moderation in alcohol use research demonstrated that within an adult population-based sample, genetic influences on alcohol use were greater among unmarried women, whereas having a marriage-like relationship reduced the impact of genetic influences on drinking (Heath et al. 1989). In 1999, Koopmans et al. demonstrated that religiosity moderates genetic influences on alcohol use among adult females, with genetic factors playing a larger role among those without a religious upbringing (Koopmans et al. 1999).

In addition to these studies that examined moderation of genetic influences in adult samples, adolescent specific gene-environment interactions (GxE) have also been a burgeoning area of study. As adolescent phenotypes have been shown to be powerful indicators of risk for adult alcohol problems, adolescent alcohol use and related behavior problems are relevant in understanding the genetic epidemiology of emerging alcohol problems. Further, there is accumulating evidence that adolescent behaviors may be particularly susceptible to environmental moderation of genetic effects since most adolescents are not yet autonomous individuals and are highly influenced by their home environment, family and peer group. In 2001, Rose and colleagues observed in

Finnish adolescent twins, that genetic factors had more influence on frequency of alcohol use in urban than in rural settings from age 16 to 18.5 years, whereas common environmental factors accounted for more variation in alcohol use frequency in rural areas (Rose et al. 2001a). Following up on these findings, Dick and colleagues (Dick et al. 2001) found that specific neighborhood characteristics (ie. higher percentage of young adults, migration and regional alcohol sales) also moderated the genetic influence on alcohol use frequency in late adolescence (age 18). In 2009 (Dick et al. 2009), Dick et al. examined the moderating effects of socioregional factors on alcohol use and behavior problems in younger twins (age 14). Their results were in line with the original study of older adolescents, indicating that the genetic effects on adolescent behavior problems were greater in urban settings and in neighborhoods characterized by more slightly older adolescents and increased social mobility, whereas, common environmental influences played a larger role in rural settings. Their results suggest that communities characterized by older adolescent role models and greater social mobility allow for increased expression of genetic dispositions that contribute to individual differences in adolescent behavior problems. Conversely, communities with fewer older peers and more social structure create opportunities in which common environmental effects, within families and within communities, assume greater importance. The authors hypothesized that higher rates of migration reflected reduced neighborhood cohesion, stability, and monitoring, thus creating more opportunity for individual expression of genetic predispositions. In 2003, Cleveland and Wiebe found that in adolescent males, genetic influences on drinking were potentiated by exposure to parental drinking; again, this may suggest a more opportunistic drinking environment for expression of genetic predispositions toward alcohol use (Cleveland & Wiebe 2003). In 2006, Dick et al. reported a moderating effect of parental monitoring on the genetic and environmental influences on adolescent smoking (age 14) in Finnish twins (Dick et al. 2006). Genetic influences were enhanced in environments with lower parental monitoring and reduced in environments with higher parental monitoring. These analyses suggest that when adolescents receive little parental monitoring, it creates an environment that allows for greater opportunity to express genetic predispositions and conversely when adolescents receive more monitoring, the environment attenuates the opportunity for genetic expression. Additionally, peer alcohol use was found to moderate the genetic and environmental influences on adolescent drinking at age 17 within the Finnish twin sample (Dick et al. 2007a): among adolescents with a larger number of peers who used alcohol, there was greater expression of genetic predispositions. Finally, an interdependent sibling relationship is an important modifier of drinking habits, and it appears to reduce the impact of inherited liabilities on alcohol-related behavior especially in adolescence (Penninkilampi-Kerola et al. 2005).

More recently, these moderation effects have been extended into the molecular literature. In 2009, Chen and colleagues extended these findings when they reported that genetic risk for nicotine dependence associated with *CHRNA3* SNP *rs16969968* was modified by level of parental monitoring (Chen et al. 2009). In 2009, Dick et al. reported that the association of *GABRA2* with externalizing trajectories across development (ages 12-22) diminished with high levels of parental monitoring (Dick et al. 2009) and more recently reported an interaction in which the association between several SNPs in *CHRM2* and externalizing behavior was stronger in environments with lower parental monitoring (Dick et al. 2011). In 2010, Johnson et al. reported that peer smoking had a substantially lower effect on nicotine dependence among those with the high-risk AA genotype at the functional SNP rs16969968 (*CHRNA5*) than among those with lower-risk genotypes. Converging evidence from twin studies and molecular genetic studies provide additional support for these GxE effects, as recently reviewed by Young-Wolff et al, Clinical Psychology Review (2011).

Previously, we observed (Dick et al. 2007b) that the diverse interactions observed in the alcohol literature appear to converge on a common mechanism, namely that of social control versus opportunity. The various environments that have been found to exacerbate genetic effects all appear to allow greater opportunity to express individual predispositions (absence of a marital partner, presence of deviant or substance using peers, lower parental monitoring, less religiosity, reduced community monitoring/more alcohol availability, independence from co-twin), whereas environments that provide greater social constraints allow less opportunity for genetic predispositions to play a role; in these cases the environmental factors are more important in individual's drinking patterns. This raises question as to whether there is anything specific about the moderation effects associated with different environmental moderators, or whether moderation is concentrated at the level of common variance shared across the theoretically different environmental dimensions. The present study used data from a sample of Finnish twins to examine common versus unique moderating effects associated with three environmental variables, parental knowledge, peer deviance, and frequency of family dinner, on the genetic and environmental influences on alcohol use at ages 14 and 17. This study used a longitudinal sample to explore the developmental relationship between these environmental moderators and frequency of alcohol use at ages 14 and 17. Parental monitoring and peer deviance were selected for further study based on our previous evidence of moderating effects associated with these outcomes in the Finnish twin samples (Dick et al. 2007a, Dick et al. 2007b). In addition, we added frequency of family dinner. Previous studies suggest that more frequent family meals may reduce problem behaviors by providing structure, stability, and improving family communications (Sen 2010). For these reasons, we hypothesized that frequent family dinner has potential to operate as a social control in a similar fashion to *high* parental monitoring and *low* peer deviance. While parental knowledge and peer deviance have previously been shown to moderate adolescent substance use, to our knowledge, frequency of family dinner has not yet been studied in this context. In this paper, we expand on previous work by testing whether the genetic moderation observed operates at the level of the shared and/or unique variance of these environmental moderators. We test for moderation associated with a general latent factor that encompasses the common variance between these three variables, as well as for moderation associated with three individual factors consisting of the residual variance specific to each environment.

Methods

Sample

The *FinnTwin12* has been described in previous chapters (chapter 3). Briefly, the study was designed to examine genetic and environmental influences on health-related behaviors. Questionnaire assessments of both twins and their parents were collected at baseline, late in the year before the twins reached age 12 (87% participation rate), with follow-up of all twins at age 14 (response rate 88%), and again at age 17 years (92.2%). In all, 4,236 questionnaires were returned at age 17 out of the 4,594 already participating in earlier questionnaires. For the current study, each environmental moderator was measured at age 14, and the outcome variables (frequency of alcohol use) were measured at ages 14 and 17.

Measures

Frequency of Drinking

At age 14, the questionnaire item asked the individual how frequently they drank alcohol and included four response options: (0) never, I don't drink alcohol, (1) less than once a month, (2) about 1 to 2 times a month, and (4) once a week or more. At age 17.5, the item included nine response options: (0) I don't drink alcohol, (2) once a year or less, (3) 2-4 times per year, (4) about once every two months, (5) about once a month, (6) a couple of times a month, (7) once a week, (8) a couple of times a week, (9) daily. The latter response options were collapsed into four categories to parallel the age 14 data; (0) never (1) weekly (3) monthly (4) daily. Non-drinkers were excluded from all analyses. The four categories from each of the two drinking variables were transformed into a continuous numeric scale so that they became semi-continuous variables; individuals who reported they never drank were given a value of 0, individuals who reported they drank less than once a month were given a value of .33, individuals who reported using alcohol about 1 or 2 times per month were given a value of .50, and individuals who reported using alcohol once per week or more were given a value of 1. Age 14 drinking frequency was available on 5,656 same-sex twin individuals (1,395 MZ twin pairs, 1,433 DZ twin pairs). Age 17.5 drinking frequency was available on 4,732 same-sex twin individuals (1,168 MZ pairs, 1,198 DZ pairs).

Parental Knowledge (Knowledge)

Knowledge was assessed with four questions included in the twins' questionnaire administered at age 14. The questions, created by Chassin and colleagues (Chassin et al. 1993), asked the adolescents to report on the degree to which their parents (1) know about their daily plans (2) know of their interests, activities, and whereabouts (3) know how they spend their money, and (4) know where and with whom they are outside of the home. Responses were made on a 4-point scale ranging from 1 (*almost always*) to 4 (*rarely or never*). A sum score based on the tallying of these items was created on 4,542 adolescents. We note that we have previously referred to this measure as "parental monitoring" in Finnish Twin Study publications, however, this variable likely reflects both solicited information and spontaneous information provided by the child and therefore we will refer to this measure as parental knowledge (Kerr & Stattin 2000).

Peer Deviance (Peers)

At age 14, the adolescents were asked the four following questions regarding their friends' behavior: (1) Do any of your friends /acquaintances drink? (2) Do any of your friends/acquaintances smoke? (3) Do any of your friends/acquaintances use drugs? (4) Do any of your friends/acquaintances get into trouble at school? For each of these questions, the response options included: (1) None, (2) One, (3) 2–5, (4) More than five. The term 'friends /acquaintances' rather than 'friends' was used here, because the illegal nature of underage alcohol use and illicit drug use was considered and we assumed that an adolescent would be more willing to report illegal behavior if it was not narrowly pinned to his or her own circle of friends (Rimpelä et al. 2006). A sum score based on the tallying of these items was created on 4,542 adolescents.

Frequency of Family Dinner (Dinner)

Frequency of family dinner consisted of two items assessed at age 14: (1) frequency of dinner together on weekdays and (2) frequency of dinner together on weekends. Response options ranged from 1 (always) to 4 (never). Family dinner was defined as having dinner with at least one parent/guardian. A sum score based on the tallying of these items was created on 4,542 adolescents.

Statistical Analyses

Data Reduction

Prior to analysis, each moderator variable was re-coded so that higher scores on each factor reflected higher risk to the adolescent (*less* parental knowledge, *more* peer deviance and *less* frequent family dinner). Using Mplus version 6.1 **(Muthen & Muthen 2006),** a second-order confirmatory factor analysis was used to differentiate a second-order *common environmental factor*, reflecting the shared variance across the three distinct environments, from three residual first-order factors reflecting the variance uniquely attributable to individual environments ($\chi^2_{(21df)}$ = 253.072, $p \leq .0001$, CFI = .99). This higher-order factor structure yielded an improvement in fit over a model in which all indicators loaded onto a single environmental factor ($\chi^2_{(21df)}$ = 4795.811, $p \leq .0001$, CFI = .77). The common factor accounted for 63% of the variance in parental knowledge, 24% of the variance in frequency of dinner with family, and 25% of the variance in peer deviance, leaving residual variances of 37%, 76%, and 75%, respectively, in the three unique environmental factors. These percentages, which equate to the squared path coefficients in Figure 8, suggest that the common factor is somewhat more indicative of parental knowledge than it is of frequency of dinner with family and peer deviance. These four environmental factors (one common and three unique, residuals) were used as moderators of the genetic and environmental sources of variability in adolescent drinking in subsequent analyses.

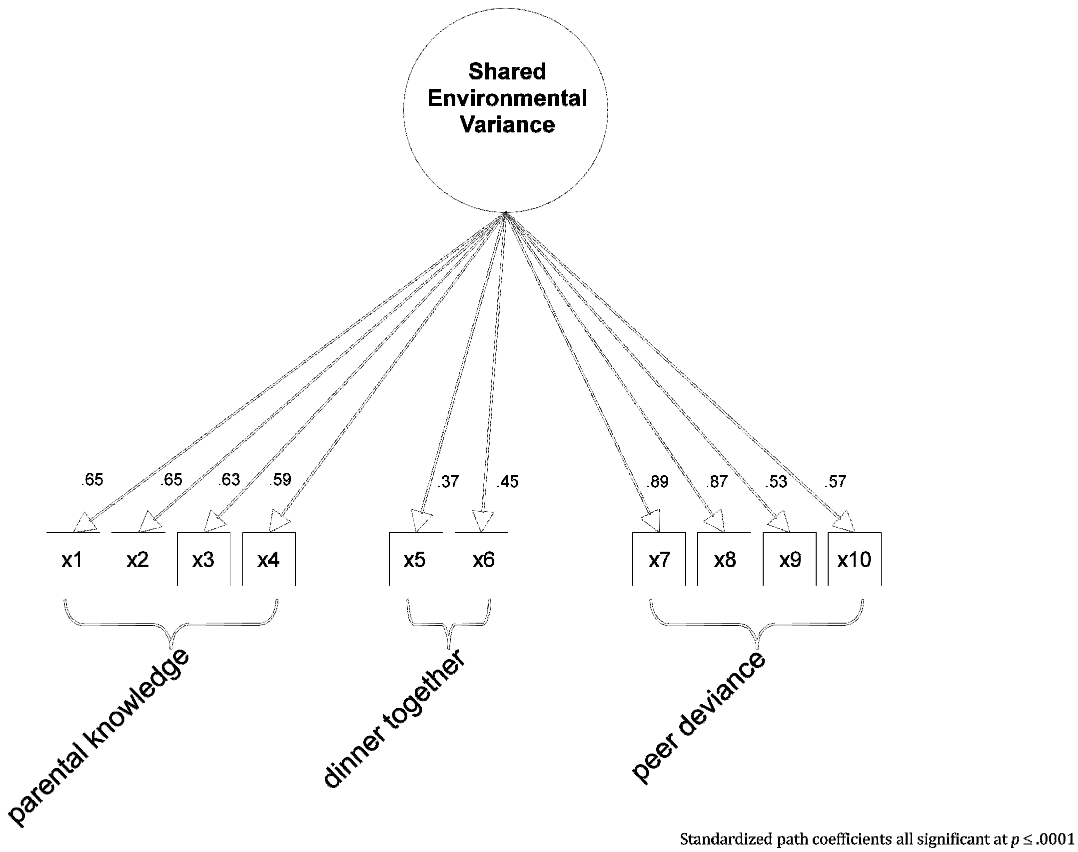


Figure 8. Second Order Confirmatory Factor Analysis of the three Environments: Parental Knowledge, Family Dinner, and Peer Deviance

Statistical Model

Comparisons of the similarity of monozygotic (MZ) and dizygotic (DZ) twin pairs yield information about the degree of influence that can be attributed to genetic and environmental factors for a particular outcome (Plomin et al. 2001). The basic genetically informative twin model partitions variance in a behavior into additive genetic influences (A), dominant genetic influences (D), common/shared environmental influences or (C), and unique environmental influences (E). Genetic influences correlate 1.0 between monozygotic (MZ) twins, who share all of their genetic variation identical-by-descent, and 0.5 between dizygotic (DZ) twins, who share, on average, 50% of their segregating genetic variation, as do ordinary siblings. Shared environmental effects, as defined in biometrical twin modeling, refer to all environmental influences that make siblings more similar to one another. By definition, these influences correlate 1.0 between both MZ and DZ twins. Unique environmental influences are uncorrelated between co-twins and have the effect of decreasing the covariance between siblings. As dominant genetic influences (D) and shared environmental influences (C) cannot be simultaneously modeled in twin-only data, we modeled shared environmental influences (C) because the DZ twin correlation exceeded ½ of the MZ twin correlation for each of the present study's outcomes. Moderation models were fit to test whether the variance components for each of the phenotypes differed as a function of shared and unique environmental factors. Figure 9 shows a classic twin model (for clarity, including only 1 twin in the pair) that has been modified to include a moderation component (Purcell 2002). The standard paths a, c, and e, indicating the magnitude of effect of additive genetic influences, shared environmental influences, and unique environmental influences, now each include a β term, which indicates the significance of a potential moderator variable M on each of these genetic and environmental influences. The value of M changes from subject to subject, taking on the value of the measured variable for that subject (i.e., parental knowledge, peer deviance and family dinner in our models).

In the moderation model, the additive genetic value is a linear function of the moderator M, represented by the equation $a + \beta_X M$, where β_X is an unknown parameter to be estimated from the data, representing the magnitude of the moderating effect. If β_X is significantly different from zero, there is evidence for a moderating effect. A similar logic follows for the β_Y and β_Z pathways, which represent the extent to which a specific moderator variable alters the importance of shared and unique environmental influences, respectively. In other words, the moderation model allows us to test whether the importance of additive genetic effects (a), shared environmental effects (c), and unique environmental effects (e) are changing as a function of the measured variable. The pathway $l + \beta_M M$ models main effects of the moderator variable on the outcome.

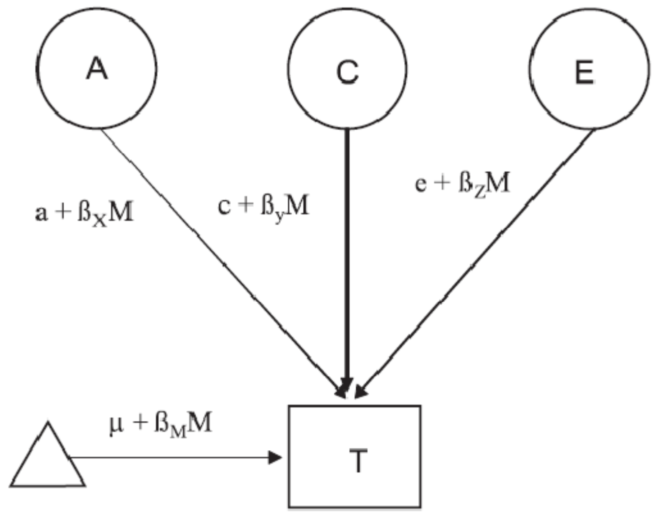
There is some evidence of genetic influence on each of the previously studied environmental moderators, parental knowledge and peer deviance (Kendler et al, 2007; Latendresse et al., 2010). For each of the presumed environmental moderators, heritability estimates were 0.27, 0.35, 0.15 for parental knowledge, peer deviance and family dinner respectively. However, previous analyses in this sample have suggested that even for those environments showing some small degree of genetic influence, the correlation with drinking frequency in early adolescence was largely environmentally mediated (Latendresse et al., 2010). Further, any covariance between the moderator and the outcome (and accordingly, any gene-environment correlation) is incorporated into the means model.

All modeling was conducted using the raw data option in Mx (Neale 2000). Mx is a structural equation modeling program developed specifically for the use of twin and family data. The significance of each of the parameters in the model can be tested by dropping a parameter and evaluating the change in 2 log likelihood between the initial model and the nested submodel. This difference is evaluated using a chi-square distribution. A significant change in fit between the models ($p < 0.05$) for the difference in degrees of freedom indicates that dropping the parameter caused a significant decrease in the model fit, indicating that this dropped pathway significantly contributes to the outcome trait and should be retained in the model.

Model-fitting proceeded in a series of steps. First, we tested the significance of the main effect of the moderator separately on drinking frequency at age 14 and 17. Next, we tested the significance of total moderation effects by dropping all moderating effects of the environment on the genetic, shared and unique environmental influences on drinking frequency simultaneously (3 df test, β_X, β_Y, and β_Z dropped). When this test was significant, we conducted further testing to determine what specific variance components showed significant moderation by sequentially dropping and testing the significance of each of the moderating effects one by one (moderation of (1) A, (2) C, and (E)). We followed this series of analyses for each moderator: the common environment, the parental knowledge residual, the peer deviance residual, and the family dinner residual. We fit all models separately for frequency of drinking at age 14 and frequency of drinking at age 17.

Figure 9. Moderation model

The latent variable A, represented in a circle, indicates additive genetic influences on the trait (T) of interest. C represents common (shared) environmental influences on a trait, and latent E represents unique environmental influences, which are uncorrelated between the twins. The triangle indicates the mean/thresholds for T and is necessary when modeling raw data. The standard paths a, c, and e, indicating the magnitude of effect of each latent variable on the trait, each include a b term, which indicates the significance of a measured moderator variable M on each of these genetic and environmental influences.



Preliminary power analyses suggested that there was low power to discriminate sex effects, because of the large sample sizes necessary to simultaneously model moderation and sex effects with ordinal outcomes. Accordingly, female and male twins were collapsed by zygosity in modeling, though thresholds for variables were allowed to differ between the sexes when indicated by the data reflecting sex differences in prevalences of alcohol use.

Results

Descriptive Statistics

For age 14 alcohol use frequency, 64.9% of the sample reported that they had never used alcohol, 20.4% reported drinking less often than once a month, 12.1% reported using alcohol about 1 or 2 times per month, and 2.6% reported using alcohol once per week or more. For alcohol use frequency at age 17, 11.9% of the sample reported that they had never used alcohol, 22.3% reported drinking less often than once a month, 41.5% reported using alcohol about 1 or 2 times per month, and 24.3% reported using alcohol once per week or more. Scores for parental knowledge ranged from 4 to 16

(M=6.5, SD=2.14). Peer deviance scores ranged from 4 to 16 (M =7.47, SD=3.18). Scores for frequency of family dinner ranged from 2 to 8 (M =4.82, SD=1.35). The common environmental factor, parental knowledge, peer deviance, and family dinner were each positively and significantly correlated with each other (all correlations are detailed in Table 21).

Table 21. Pearson Correlations of Environmental Moderators and Behavioral Outcomes

Correlations	Common Environment	Parental Knowledge Residual	Peer Deviance Residual	Freq. of Family Dinner Residual	Age 14 Alcohol Use	Age 17 Alcohol Use
Common Environment	1					
Parental Knowledge Residual	.693*	1				
Peer Deviance Residual	.243*	.306*	1			
Frequency of Family Dinner Residual	.258*	.239*	.107*	1		
Age 14 Alcohol Use	.444*	.163*	.387*	.069*	1	
Age 17 Alcohol Use	.272*	.117*	.208*	.041*	.323*	1

*Significant at a p<0.0001

Moderation Models

The results from each of the models, testing for moderation effects associated with the shared variance and with the residual variances of parental knowledge, peer deviance and family dinner, respectively, are displayed in Table 22 and graphically in Figure 10 according to moderator and outcome. There was a significant main effect of all environmental variables (shared variance, knowledge residual, peers residual, dinner residual) on alcohol use frequency at age 14. Dropping moderation effects on additive genetic, shared environmental and unique environmental factors significantly reduced model fit for the shared variance factor, knowledge residual and peers residual. Figure 10 depicts the direction of these effects. For the dinner residual, only dropping the shared environmental moderation effects significantly reduced model fit.

At age 17, there was a significant main effect of the shared variance and residual peer deviance on alcohol use frequency. There was no main effect of residual parental knowledge or residual family dinner on age 17 alcohol use frequency. Simultaneously dropping additive genetic, shared environmental and unique environmental moderation effects significantly reduced model fit only for peer deviance. Figure 10 depicts the direction of these effects. For the shared variance factor, dropping additive genetic and unique environmental moderation effects significantly reduced model fit. Although there was not a significant main effect of the parental knowledge residual on alcohol use frequency at age 17, modest genetic moderation and borderline significant shared and unique environmental moderation ($p<0.10$) was observed. There were no statistically significant moderating effects of genetic or environmental influences on frequency of drinking at age 17 associated with the frequency of family dinner residual.

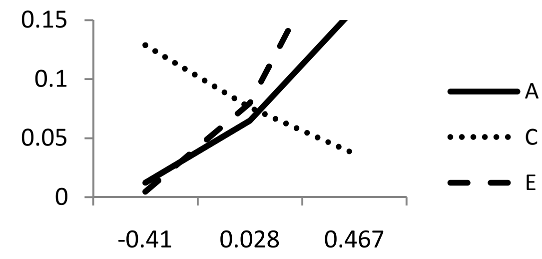
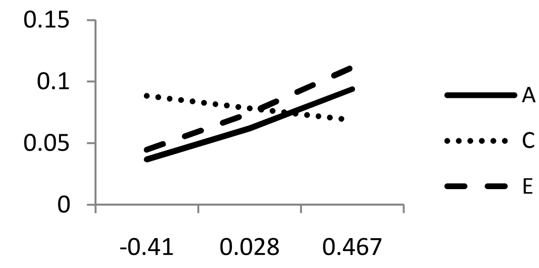
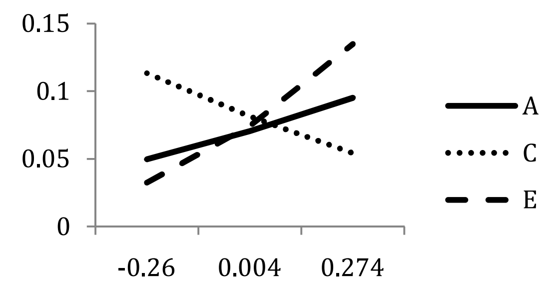
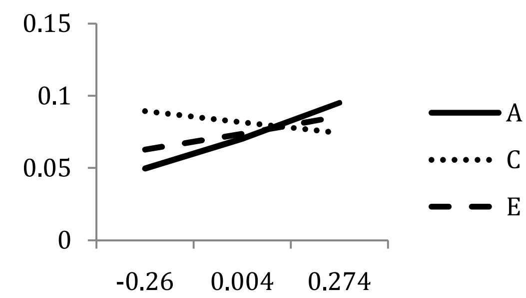
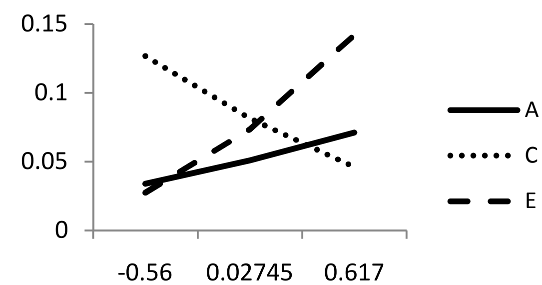
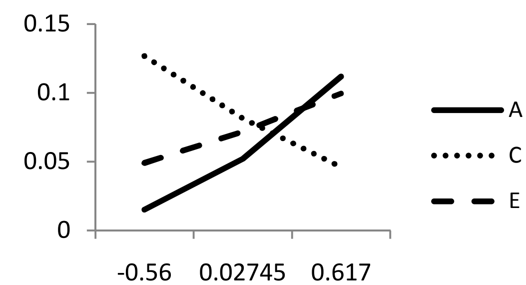
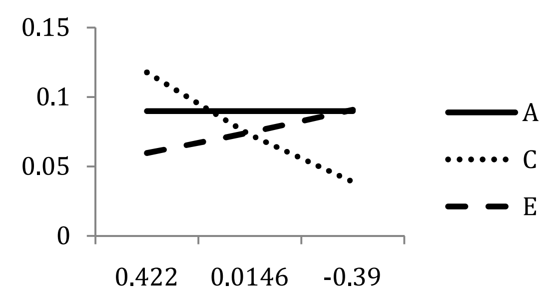
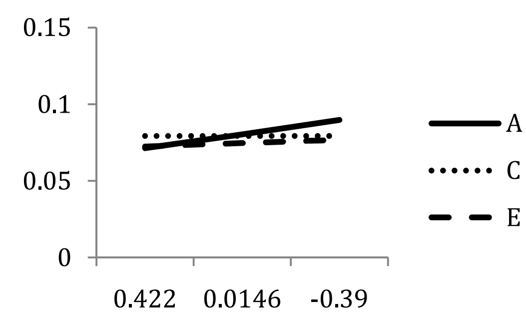
Table 22. Results from each of the models, testing for moderation effects associated with the common environmental risk variance and the residual variance associated with parental knowledge, peer deviance and frequency of family dinner

	Outcome	
	Age 14	Age 17
Drop Moderator	**Freq Drinking**	**Freq Drinking**
	X^2(p-value)	*X^2(p-value)*
Common Environment		
Main effect	**408.33(<0.001)**	**19.19(<0.001)**
Genetic Moderation	**93.51(<0.001)**	**6.13(0.01)**
Shared Env. Moderation	**23.82(<0.001)**	3.17(0.08)
Unique Env. Moderation	**501.11(<0.001)**	**6.82(0.01)**
Parental Knowledge Residual		
Main effect	**44.44(<0.001)**	0.35(0.56)
Genetic Moderation	**7.96(0.01)**	**3.94(0.05)**
Shared Env. Moderation	**10.29(<0.001)**	2.90(0.09)
Unique Env. Moderation	**54.12(<0.001)**	3.31(0.07)
Peer Deviance Residual		
Main effect	**297.89(<0.001)**	**24.35(<0.001)**
Genetic Moderation	**4.34(0.04)**	**14.20(<0.001)**
Shared Env. Moderation	**9.64(<0.001)**	**12.68(<0.001)**
Unique Env. Moderation	**27.42(<0.001)**	**3.93(0.05)**
Frequency of Family Dinner Residual		
Main effect	**7.65(0.01)**	1.36(0.24)
Genetic Moderation	0.85(0.36)	0.29(0.59)
Shared Env. Moderation	**12.14(<0.001)**	0.01(0.92)
Unique Env. Moderation	2.03(0.15)	1.98(0.16)

*All tests df=1

**All significant results(p-value<0.05) are bolded

Figure 10. Results from each of the models, testing for moderation effects associated with the general environmental factor, parental knowledge, peer deviance and frequency of family dinner on frequency of drinking at ages 14 and 17. Note: The dichotomous depiction of the environmental moderators is used only for illustration; a semi-continuous variable was used in the models

Environment	**Outcome**	
	Age 14 Drinking Frequency	Age 17 Drinking Frequency
	Social Control ***Social Opportunity***	***Social Control*** ***Social Opportunity***
General Environment		
Parental Knowledge		
Peer Deviance		
Family Dinner		

Discussion

Previous studies have demonstrated that a variety of environmental variables moderate the relative magnitude of genetic effects on substance use, dependence, and related disorders. These include environments that span a number of different domains, including parental factors (Dick et. al. 2007b, Chassin et al.1993), peer influences (Dick et al. 2007a), neighborhood influences (Rose et al. 2001a, Dick et al. 2009), romantic relationships (Heath et al. 1989) and religious influences (Koopmans et al.1999). A common element of many of the detected effects is that genetic influences are enhanced in conditions that allow opportunity to express predispositions and diminished in environments that could be perceived as more constraining. The various environments that have been found to exacerbate genetic effects all appear to allow greater opportunity to express individual predispositions (absence of a marital partner, presence of deviant or substance using peers, lower parental monitoring, less religiosity, reduced community monitoring/more alcohol availability), whereas environments that provide greater social constraints allow less opportunity for genetic predispositions to play a role; in these cases environmental factors are more important in individual's drinking patterns.

These effects map onto Shanahan and Hofer's proposed Social Context as Social Control mechanism of GxE, whereby social controls (such as parental monitoring and involvement, positive peer influences and lack of access to illegal substances) may attenuate the genetic predisposition to adolescent substance use. This mechanism is one of the four potential GxE mechanisms offered in a Shanahan and Hofer's 2005 review, which also delineated *contextual triggering, social context as compensation, and* *social context as enhancement* (Shanahan & Hofer 2005). The first proposed mechanism, *contextual triggering,* refers to a detrimental environment combining with a genetic predisposition to produce the negative outcome. *Social context as compensation* refers to an enriched setting that prevents the expression of a genetic predisposition to a negative outcome. Lastly, *social context as enhancement* refers to the ability the environment has to accentuate genetic predispositions for positive outcomes. The examples of the specific environments that moderate the genetic predisposition to adolescent substance use, parental knowledge, peer alcohol and drug use, neighborhood characteristics, all appear to fall under the "social context as social control" mechanism, and suggest that this mechanism is particularly relevant in alcohol use.

This conceptually shared mechanism begs the question as to whether there is anything uniquely important about each of the individual environments, or whether they are all simply reflective of a shared environmental factor. The present study sought to address this question. We examined the specificity of the moderating effects of three environmental variables, two of which previously have been demonstrated to moderate adolescent substance use (parental knowledge and peer deviance), and one new variable: family dinner. Our results suggested that while there is evidence of genetic moderation by the shared variance across these environmental moderators, there is also important information unique to parental knowledge, peer deviance and frequency of family dinner. All three of these environmental variables play a moderating role in adolescent alcohol use, both jointly, through shared sources of variance, and uniquely, through residual sources specific to each. Further, while all three environments may operate via a shared mechanism, each individual environment is important in its own right; they are not merely operating as proxies of one another or a shared risk environment.

The longitudinal study design we employed allowed us to study the relationships between environmental moderators and alcohol use across adolescent development. Each of the moderators we examined predicted frequency of alcohol use at age 14. However, by age 17 only peer deviance remained significant. As the environments were measured early in adolescence (age 14), it seems reasonable that they were less developmentally relevant by age 17. The moderating effect of the shared factor, parental knowledge unique variance, and family dinner unique variance on genetic influences decreased across time. While the moderating effects of the shared variance factor and variance unique to parental monitoring remained statistically significant by age 17, the moderating effects of the family dinner residual diminished entirely. Alternatively, the moderating effects of the unique variance associated with peer deviance on genetic influences increased, having reached its greatest significance at age 17. We believe that the specific familial contexts (parental knowledge and frequency of family dinner) appeared to be more relevant in early adolescence when individuals have less autonomy, while specific peer influences persisted because individuals are actively engaged in selecting their social networks throughout adolescence.

Note that these developmental effects should be interpreted with caution as drinking frequency at age 14 and age 17 are likely reflecting somewhat different developmental phenomena. Twin studies suggest that age 14 drinking is more closely linked to adolescent externalizing behavior whereas age 17 drinking is more closely related to young adult drinking patterns (Kendler et al., 2011). As the current sample size lacked the power required to simultaneously model drinking initiation and regular drinking frequency, we ran parallel analyses for age 14 externalizing behavior as measured by the Multidimensional Peer Nomination Inventory (Pulkinen et al., 1998). In this sample, age 14 behavior problems and drinking frequency were signficantly correlated (r2=0.256) and the continuous measure of behavior problems lacked the skewed nature of the drinking frequency measure (65% of the sample reported never drinking at age 14). The results from these analyses were virtually identical to the age 14 drinking frequency results, yet showed more robust moderation effects for both genetic and environmental influences. We believe this provides additional support for the age 14 drinking frequency results as well as for the shared genetic relationship between behavior problems and alcohol use in adolescence.

There are several additional limitations of this study to consider. One is the inability to examine sex effects due to a lack of power to simultaneously model moderation and sex effects. While we modeled different means and variances for males and females, the present analyses do not formally test for sex differences. Another consideration is the factor loadings for the shared factor. This factor is more representative of parental knowledge and somewhat less so of peer deviance and family dinner, though we believe the structure of the shared variance is interesting in its own right. However, it is important to keep in mind that the “shared variance factor” is most strongly influenced by parental knowledge. Also note that parental knowledge was assessed at age 14, that is, 3 years before the study of drinking behavior at age 17. The positive parent-child relationship in mid-childhood (indicated by a high parental knowledge score) may potentially have longstanding significance; it may help the parents to cope with adolescent processes and limit the number of deviant peers the child engages with. Thus, while we treat these environments as different variables, they are not necessarily entirely independent of each other. Although measures of socioregional and neighborhood factors which previously showed moderation of adolescent alcohol use were available in this sample, they were not included in these analyses. Preliminary analyses indicated a weak relationship between socioregional demography measures and parental knowledge (r=0.11), peer deviance (r=0.07), and family dinner (r=0.09). As such, including these socioregional and neighborhood measures in the common factor analysis provided a poor fit to the data. We believe that this provides additional evidence that there are effects unique to specific environments, even though the mechanism of influence may be similar.

Currently, large-scale efforts to identify specific genetic risk factors for alcohol use are underway. As researchers continue to refine molecular genetic methods, it is important to use all available information on the epidemiology of alcohol use to inform these methods. This study adds to a literature that provides evidence of environmental moderation of the genetic influence on alcohol use. That is, genetic influences on alcohol use will diminish or strengthen given environmental circumstance. Our findings suggest that there are moderating effects of the shared environmental variance on adolescent alcohol use, as well as information captured by the variance unique to each individual environment. The shared environmental variance may conceptually map onto an overarching mechanism of social opportunity/social control. In addition, our findings indicate that it is important to carefully consider the most influential environments for a given age group in a given sample, as the relevance of particular environments on the genetic influences on alcohol use tend to shift across development. We believe that these twin studies have important implications for gene-finding studies in that ignoring the effects of the environment on genetic risk for alcohol use may lead to missed opportunities in identifying key risk factors for alcohol use.

References

1. Chassin, L., Pillow, D., Curran, P., Molina, B., & Barrera, M. (1993) Relation of parental alcoholism to early adolescent substance use: A test of three mediating mechanisms. *Journal of Abnormal Psychology*, 102:3-19.

2. Cleveland, H. H. & Wiebe, R. P. (2003) The moderation of genetic and shared-environmental influences on adolescent drinking by levels of parental drinking. *Journal of Studies on Alcohol*, 64:182-194.

3. Dick, D. M., Bernard, M., Aliev, F., Viken, R., Pulkkinen, L., Kaprio, J., & Rose, R. J. (2009) The Role of Socio-Regional Factors in Moderating Genetic Influences on Early Adolescent Behavior Problems and Alcohol Use. *Alcoholism: Clinical and Experimental Research*, 33 (10):1739-48.

4. Dick, D. M., Pagan, J. L., Holliday, C., Viken, R., Pulkkinen, L., Kaprio, J., & Rose, R. J. (2007a) Gender Differences in Friends' Influences on Adolescent Drinking: A Genetic Epidemiological Study. Alcoholism: Clinical and Experimental Research, 31: 2012-19.

5. Dick, D. M., Pagan, J. L., Viken, R., Purcell, S., Kaprio, J., Pulkkinen, L., & Rose, R. J. (2007b) Changing Environmental Influences on Substance Use Across Development. Twin Research and Human Genetics, 10: 315-326.

6. Dick, D. M., Purcell, S., Viken, R. J., Kaprio, J., Pulkkinen, L., & Rose, R. J. (2007) Parental Monitoring Moderates the Importance of Genetic and Environmental Influences on Adolescent Smoking, *Journal of Abnormal Psychology*, 116(1):213-8.

7. Dick, D. M., Rose, R. J., Viken, R. J., Kaprio, J., & Koskenvuo, M. (2001) Exploring gene-environment interactions: Socioregional moderation of alcohol use. *Journal of Abnormal Psychology*, 110: 625-632.

8. Goldman, D. (1993) Recent developments in alcoholism: genetic transmission. *Recent Dev.Alcohol*, 11: 231-248.

9. Heath, A. C., Jardine, R., & Martin, N. G. (1989) Interactive effects of genotype and social environment on alcohol consumption in female twins. *Journal of Studies on Alcohol*, 50:38-48.

10. Kaprio, J., Pulkkinen, L., & Rose, R. J. (2002) Genetic and environmental factors in health-related behaviors: studies on Finnish twins and twin families. *Twin Res*, 5(5):366-71.

11. Kendler, K. S., Walters, E. E., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1995) The structure of the genetic and environmental risk factors for six major psychiatric disorders in women: Phobia, generalized anxiety disorder, panic disorder, bulimia, major depression, and alcoholism. *Archives of General Psychiatry*, 52: 374-383.

12. Kerr, M. & Stattin, H. (2000) What parents know, how they know it, and several forms of adolescent adjustment: Further support for a reinterpretation of monitoring. *Developmental Psychology*, 36: 366-380.

13. Koopmans, J. R., Slutske, W. S., van Baal, G. C., & Boomsma, D. I. (1999) The influence of religion on alcohol use initiation: evidence for genotype X environment interaction. *Behav Genet*, 29(6):445-53.

14. Muthen, L. K. & Muthen, B. O. Mplus Users Guide (2006) Los Angeles, CA.

15. Neale, M. C. (2000) The use of Mx for association and linkage analyses. *GeneScreen*, 1;107-111.

16. Plomin, R., DeFries, J. C., McClearn, G. E., & McGuffin, P. (2001) *Behavioral Genetics* Worth, London.

17. Pulkkinen, L., Kaprio, J., & Rose, R. J. (1999) Peers, teachers and parents as assessors of the behavioral and emotional problems of twins and their adjustment: the Multidimensional Peer Nomination Inventory. *Twin Res*, 2 (4):274-85.

18. Purcell, S. (2002) Variance components models for gene-environment interaction in twin analysis. *Twin Res*, 5(6):554-71.

19. Rose, R. J., Dick, D. M., Viken, R. J., & Kaprio, J. (2001a) Gene-environment interaction in patterns of adolescent drinking: Regional residency moderates

longitudinal influences on alcohol use. *Alcoholism: Clinical and Experimental Research*, 25:637-643.

20. Rose, R. J., Dick, D. M., Viken, R. J., Pulkkinen, L., & Kaprio, J. (2001b) Drinking or abstaining at age 14? A genetic epidemiological study. *Alcohol Clin Exp Res*, 25(11):1594-604.

21. Rutter, M. (2010) Gene-environment interplay. *Depress.Anxiety,* 27(1):1-4.

22. Sen, B. (2010) The relationship between frequency of family dinner and adolescent problem behaviors after adjusting for other family characteristics. *J.Adolesc.,* 33(1): 187-196.

23. Shanahan, M. J. & Hofer, S. M. (2005) Social context in gene-environment interactions: retrospect and prospect. *J.Gerontol.B Psychol.Sci.Soc.Sci.*, 60(1): 65-76.

24. Tsuang, M. T., Bar, J. L., Harley, R. M., & Lyon, M. J. (2001) The Harvard twin study of substance abuse: What we have learned. *The Harvard Review of Psychiatry*, 9: 267-279.

Chapter 5

Life Events Moderate Genetic and Environmental Influences on Adolescent Externalizing Disorders

Abstract

Background: The well documented association between life events and adolescent alcohol use has led researchers to examine this candidate environment as a moderator of genetic influences on alcohol-related outcomes. A recent twin study found that as the number of stressful life events increased, additive genetic influences on adolescent externalizing disorders also increased (Hicks et al., 2009). The goal of the present study is to examine life events, one important environmental context related to adolescent alcohol use, as a moderator of genetic influences on adolescent alcohol use using two complementary methods: twin modeling and genetic risk scores.

Methods: We first used twin data from the *Finntwin12* to examine the moderation effects of life events at age 14 on concurrent alcohol use (age 14) and later use at age 17. We then used available GWAS data on these same twins to create genetic risk sum scores (GRSS; an index of aggregate genetic risk for frequent adolescent alcohol use) and examined whether life events in early adolescence moderated this measured genetic risk.

Results: Our twin study found that in conditions of more life events, both additive genetics, shared and unique environment play a more important role; conversely, in the

conditions of less life events, latent additive genetic factors, shared and unique environmental factors were attenuated. This effect was significant at age 14 only. The GRSS created for the twins significantly predicted frequency of use at both ages 14 and 17; however, the interaction between the GRSS and age 14 life events was only significant at age 14.

Conclusions: Testing for environmental moderation at the level of aggregate molecular genetic risk allows us to parallel the established latent gene-environment interaction effects reported from twin studies. This method also allows us to begin to more systematically characterize the specific environments that are critical for moderating the importance of a genetic predisposition, and the ages and developmental stages at which these gene environment interactions operate.

Introduction

Alcohol use is a normative part of adolescent life (Johnston et al., 2006), but frequent or heavy adolescent alcohol use is associated with a host of problems at both personal and societal levels (Gaffney et al., 1998; Jelalian et al., 2000) and may develop into pervasive adulthood disorders (Schulenberg and Maggs, 2002; Brown et al., 2008). Both genetic and environmental factors contribute to alcohol use in adolescence (Rose et al., 2001; Maes et al., 1999); furthermore we also know that genetic and environmental risk and protective factors often do not exist independent of one another. Examining the interactive effects of factors associated with adolescent alcohol use is important for understanding the contexts in which risk is amplified or attenuated. The goal of the present study is to examine life events, one important environmental context related to adolescent alcohol use, as a moderator of genetic influences on adolescent alcohol use using two complementary genetically-informed methods: twin modeling (in which we test for moderation of latent, unmeasured genetic influences) and genetic risk sum scores (in which we test for moderation of the effect of aggregated measured genotypes).

Disruptive or stressful life events are related to adolescent alcohol use in both human and non-human animals. For example, female rats exposed to prenatal restraint stress tended to consume higher amounts of ethanol in adolescence (van Waes et al., 2011). Similarly, adolescent rhesus monkeys exposed to a prenatal noise stressor show an increasing alcohol preference across a five-week period in adolescence (Schneider et al., 2002) and rhesus monkeys with a history of stressful rearing experiences (peer rearing) consume more alcohol in adolescence compared to rhesus monkeys raised by their mothers (Higley et al., 1991). In the human literature, fifteen-year-olds who experienced more life events in the three years prior were more likely to have had more than five drinks in a row (four for females) and to have consumed a greater maximum amount of alcohol per occasion relative to those who had not experienced multiple events (Blomeyer et al., 2008). Nineteen-year-olds who experienced more life events in the four years prior reported more binge drinking days and greater number of drinks in the past 45 days relative to those who experienced fewer life events (Laucht et al., 2009). A cross-sectional study of high school juniors likewise found positive associations between stressful life events and concurrent alcohol use and alcohol problems (Windle and Windle, 1996). Convergent findings from the developmental trauma literature indicate that adverse life experiences (e.g., maltreatment) and concurrent life events are associated with clinical alcohol use disorders in adolescents (Clark et al., 1997).

The robust association between life events and adolescent alcohol use has led a number of research groups to examine this candidate environment as a moderator of genetic influences on alcohol-related outcomes. For example, a recent twin study found that as the number of stressful life events increased, additive genetic influences on adolescent externalizing disorders (as measured with a composite of self- or mother-reported symptoms of antisocial behavior, alcohol, nicotine, and illicit drug dependence, and teacher-reported externalizing behaviors) also increased (Hicks et al., 2009). The widely-studied serotonin transporter (5-HTTLPR) polymorphism was also found to interact with past-year life events to predict first-year college students' drinking (Covault et al., 2007). Those homozygous for the short allele drank more frequently and more heavily if they had also experienced multiple stressful life events. Similarly, adolescent carriers of the short allele with a history of maltreatment report an earlier age of alcohol use onset (Kaufman et al., 2007), although in at least one case those homozygous for the long allele appeared to be at greater risk if they had experienced greater early psychosocial adversity or adolescent life events (Laucht et al., 2009). Similar findings emerge from the animal literature, where female rhesus monkey carriers of the long/short allele of the orthologous rh5-HTTLPR genotype exposed to stressful peer-rearing early in life consumed more alcohol as adolescents compared to peer-reared carriers of the long/long allele (Barr et al., 2004). Variation in corticotropin releasing hormone receptor 1 (CRHR1), a gene implicated in stress responsivity, also interacts with stressful life events to predict earlier age of onset of first drink (Schmid et al., 2010) and heavy adolescent drinking for those homozygous for the C allele of rs1876831 (Blomeyer et al., 2008).

Despite these advances, several gaps exist in our understanding of how stressful life events come together with genetic risk to predict adolescent alcohol use. First, although one adolescent twin study indicates that genetic influence on broadband externalizing disorders (including symptoms of alcohol dependence) increases as levels of life stress increase (Hicks et al., 2009), to our knowledge no study has examined whether this effect holds for adolescent alcohol use in particular. Relatedly, whether the moderating effect of adolescent life events is sustained over time or is limited to cross-sectional effects has not yet been examined. Addressing this question is important for understanding the long-term consequences of stressful life events during this period of rapid developmental change, which some have suggested may be a sensitive period for downstream cognitive, behavioral, and emotional problems (Steinberg, 2005). Second, although evidence suggests that heritability for externalizing behavior increases under conditions of greater life stress (Hicks et al., 2009), examining the nature of this interaction effect using a measured genetic risk approach that goes beyond single-gene studies represents an important next step in this area.

The goal of the present research is to address these gaps in the literature by bringing together two complementary methods to examine stressful life events as a moderator of genetic influence on adolescent alcohol use. First, we use data from a genetically informative, population-based sample of monozygotic and dizygotic twins to examine whether life events in early adolescence moderate genetic and environmental risk for alcohol frequency concurrently (age 14) and over time (age 17). Next, we use genome-wide association data available on a subset of participants from the twin sample to create genetic risk sum scores and examine whether and how life events in early adolescence moderate measured aggregate genetic risk to predict alcohol frequency in early and later adolescence.

Methods

Sample

The *FinnTwin12* has been described in previous chapters (chapters 3 and 4). This chapter uses data on drinking frequency from the age 14 and 17.5 assessments since adolescence was hypothesized to be a time when gene-environment interactions would be particularly salient. The genotypic data used data collected at the age 22 follow up.

DNA was collected on a subset of the twins from the epidemiological sample that has been more intensively studied. There were 1,069 individuals with genetic data, including 406 monozygotic (MZ) twin individuals and 614 dizygotic (DZ) twin individuals.

Measures

Frequency of Drinking

At age 14, the questionnaire item asked the individual how frequently they drank alcohol and included four response options: (0) never, I don't drink alcohol, (1) less than once a month, (2) about 1 to 2 times a month, and (4) once a week or more. 64.9% of the sample reported that they had never used alcohol, 20.4% reported drinking less often than once a month, 12.1% reported using alcohol about 1 or 2 times per month, and 2.6% reported using alcohol once per week or more. Parallel response options were created using the age 17.5 data. At age 17.5, 11.9% of the sample reported that they had never used alcohol, 22.3% reported drinking less often than once a month, 41.5% reported using alcohol about 1 or 2 times per month, and 24.3% reported using alcohol once per week or more. The four categories from each of the two drinking variables were transformed into a quasi-continuous numeric scale by creating a scaled ratio for each ordinal value for modeling. Age 14 drinking frequency was available on 5,656 same-sex twin individuals (1,395 MZ twin pairs, 1,433 DZ twin pairs). Age 17.5 drinking frequency was available on 4,732 same-sex twin individuals (1,168 MZ pairs, 1,198 DZ pairs).

Life Events

At age 14, the adolescents were asked if any of the following fifteen life events had happened to them and their family in the past two years. The items included: (1) moved to a new neighborhood or town with your family, (2) a close friend moved away, (3) changed schools, (4) you have experienced a serious illness or accident, (5) someone close to you has been seriously ill or hurt, (6) someone close to you has died, (7) your parents have had serious conflicts, (8) your mother or father has moved out of the house/parents divorced, (9) a new mate of your mother or father has moved in, (10) your sister or brother has moved away from home, (11) a close teacher/coach has changed, (12) a close friendship has ended, (13) mother or father has been unemployed, (14) mother has started working after being home a long time, (15) a new sibling has been born. A sum score was computed for each individual such that higher scores indicated more life events. Life events scores ranged from 0 to 13 (M=2.8, SD=1.61). A z-score of the standardized stressful life events score was used in analyses. We note that previously this life event scale has been referred to as ***stressful life events*** on account of the disruptive nature of events listed above (including such events as the death of a parent). However, the relationship between life events scores and the adolescent's report of stress level induced by these events was moderate (r=0.44), indicating that either some individuals did not perceive these events as stressful, or lacked the insight to describe them as so. Because this scale also includes normative life events (including such events as the birth of a new child), we will refer to this scale as ***life events*.**

Twin Modeling

Comparisons of the similarity of MZ and DZ twin pairs yield information about the degree of influence that can be attributed to genetic and environmental factors for a particular outcome (Plomin et al. 2001). The basic genetically informative twin model partitions variance in a behavior into additive genetic influences (A), dominant genetic influences (D), common environmental influences or (C), and unique environmental influences (E). Genetic influences correlate 1.0 between monozygotic (MZ) twins, who share all of their genes identical-by-descent, and 0.5 between dizygotic (DZ) twins, who share, on average, 50% of their segregating genes, as do ordinary siblings. Common environmental effects, as defined in biometrical twin modeling, refer to all environmental influences that make siblings more similar to one another. By definition, these influences correlate 1.0 between both MZ and DZ twins. Unique environmental influences are uncorrelated between co-twins and have the effect of decreasing the covariance between siblings. As dominant genetic influences (D) and common environmental influences (C) cannot be simultaneously modeled in twin-only data, we modeled common environmental influences (C) because the DZ twin correlation exceeded ½ of the MZ twin correlation for each of the present study's outcomes.

Moderation models were fit to test whether the variance components for each of the phenotypes differed as a function of common and unique environmental factors. Chapter 4, figure 2 shows a classic twin model (for only 1 twin in the pair) that has been modified to include a moderation component (Purcell 2002). The standard paths a, c, and e, indicating the magnitude of effect of additive genetic influences, common environmental influences, and unique environmental influences, now each include a β term, which indicates the significance of a potential moderator variable M on each of these genetic and environmental influences. The value of M changes from subject to subject, taking on the value of the measured variable for that subject (i.e., life events in our models). In the moderation model, the additive genetic value is a linear function of the moderator M, represented by the equation $a + \beta XM$, where βX is an unknown parameter to be estimated from the data, representing the magnitude of the moderating effect. If βX is significantly different from zero, there is evidence for a moderating effect. A similar logic follows for the βY and βZ pathways, which represent the extent to which a specific moderator variable alters the importance of common and unique environmental influences, respectively. In other words, the moderation model allows us to test whether the importance of additive genetic effects (a), common environmental effects (c), and unique environmental effects (e) change as a function of the measured variable. The pathway $I + \beta MM$ models main effects of the moderator variable on the outcome. Also included in this pathway are any gene–environment correlation effects between the moderator variable and outcome. Thus, any covariance between the moderator and the outcome is incorporated into the means model. All modeling was conducted using the raw data option in Mx (Neale 2000). Mx is a structural equation modeling program developed specifically for the use of twin and family data. The significance of each of the parameters in the model can be tested by dropping a parameter and evaluating the change in 2 log likelihood between the initial model and the nested submodel. This difference is evaluated using a chi-square distribution. A significant change in fit between the models ($p < 0.05$) for the difference in degrees of freedom indicates that dropping the parameter caused a significant decrease in fit of the model, indicating that pathway significantly contributes to the outcome trait and should be retained in the model. Inherent in twin modeling are crucial assumptions that must be met when interpreting parameter estimates, the most notable being the “equal environments assumption”, which presumes that twins who are reared together receive equal treatment and essentially have the same “shared” environment for the trait of interest. An additional assumption requires that no differences may exist in the means and variances of variables as a function of zygosity; means and variances must be equivalent for MZ and DZ twins.

Model-fitting proceeded in a series of steps. We first tested the significance of the main effect of the moderator. We then tested the significance of moderation effects by dropping all moderation (3 df test, βX, βY, and βZ dropped). When this test was significant, we conducted further testing to determine what specific variance components showed significant moderation by sequentially dropping and testing the significance of each of the moderating effects one by one. We fit models separately for the moderator (life events) with each the two drinking frequency outcomes: frequency of drinking at age 14 and 17.

Preliminary power analyses suggested that there was low power to discriminate sex effects because of the large sample sizes necessary for adequate power to detect moderating effects with ordinal outcomes. Accordingly, female and male twins were collapsed by zygosity in modeling, though thresholds for variables were allowed to differ between the sexes when indicated by the data.

Genetic Association Analyses

All twins with DNA were genotyped using the Illumina 670K custom chip at the Welcome Trust Sanger Centre. SNPs were excluded if the minor allele frequency was less than 1%; further SNPs were excluded if significant ($P<10^{-4}$) deviation from Hardy–Weinberg equilibrium was observed. The data were checked for minor allele frequency (>1%) and had a genotyping success rate per SNP and per individual (>95%). To guard against the possibility that any pairs of individuals were unexpectedly related, a MDS plot (using a pairwise-IBS matrix) with only one member of each known family was created. After the pedigrees were confirmed to be correct, we reapplied the basic filters (MAF, genotyping success, HWE) to the data. Genotypes for altogether 535,613 polymorphic markers were available for analysis. An additive model was assumed, and, because of the semi-continuous outcome variable, linear regression was used.

In order to create polygenic risk scores for each individual, we first ran genome wide association analyses using frequency of drinking at age 14 and 17.5 (separately) as the outcomes. For the initial GWAS analyses, linear regression was performed on frequency of drinking using PLINK v1.07 for all autosomes. Additionally, the family structure of the data was accounted for using a permutation procedure performed in PLINK (qfam) that randomly shuffles the degree of relatedness among all individuals. Because the qfam procedure can specify only one type of familial relationship, both individuals from each DZ pair and one individual from each MZ twin pair was included in the analyses, reducing the sample size from 1,069 to 872 individuals (1 MZ twin and both DZ twins). Each of the top SNP level results were used to create a weighted genetic risk score for each individual. SNPs with a nominal *p*- value less than 0.01 were included in the genetic risk score.

Once genetic risk scores were computed for each individual, we used linear regression to test if the moderators (life events) interacted with genetic risk for drinking to predict greater frequency of drinking at age 14 and/or age 17.5. The first model included the main effect of the life events score, the second model included both main effects of the genetic risk score and the life events score as well as the interaction term. Sex was collapsed to parallel the twin analyses, and used as a covariate in all analyses. Principal components analyses of the population structure performed in Eigenstrat [26] indicated a single dimension of ancestry. As there was no evidence of ethnic stratification within this sample all individuals were included in the genetic analyses.

Results

Descriptive Statistics

For age 14 drinking frequency, 64.9% of the sample reported that they had never used alcohol, 20.4% reported drinking less often than once a month, 12.1% reported using alcohol about 1 or 2 times per month, and 2.6% reported using alcohol once per week or more. After adjusting for the familial clustering in the data, girls are slightly more likely to report alcohol use than boys at this early age [$F(2.77, 3965.03) = 3.39$, $p = 0.02$], as has been discussed previously in this sample (Rose et al. 2001b). For drinking frequency at age 17, 11.9% of the sample reported that they had never used alcohol, 22.3% reported drinking less often than once a month, 41.5% reported using alcohol about 1 or 2 times per month, and 24.3% reported using alcohol once per week or more. Life events scores (standardized) ranged from 0 to 13 (M=0, SD=1).

Greater life events scores were positively and significantly correlated with drinking frequency at age 14 and 17. Life events were not significantly associated with either genetic risk for drinking frequency at age 14 or 17; thus there is no evidence of a gene-environment correlation in this data. As expected the genetic risk scores were significantly associated with drinking frequency variables from which they were derived, both at age 14 and 17 (all correlations are detailed in Table 23).

Table 23. Correlations of Number of Life Events, Genetic Risk Scores and Frequency of Drinking

Correlations	Life Events Score	Genetic Risk Score Age 14	Genetic Risk Score Age 17	Age 14 Drinking Frequency	Age 17 Drinking Frequency
Life Events Score	1				
Genetic Risk Score Age 14	.059	1			
Genetic Risk Score Age 17	.063	.129*	1		
Age 14 Drinking Frequency	.121*	.543*	.266*	1	
Age 17 Drinking Frequency	.060*	.150*	.871*	.323*	1

*Significant at a $p<0.01$

Twin Analyses

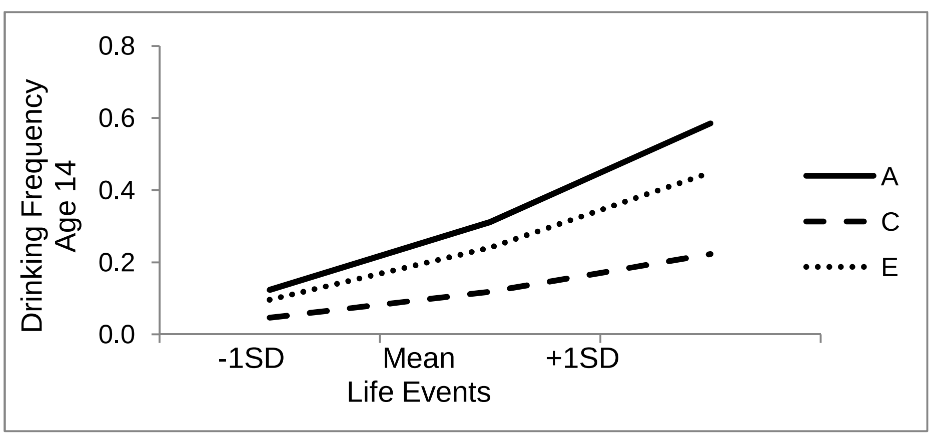
There was a significant main effect of the number of life events on age 14 drinking frequency. A greater number of life events was associated with more frequent drinking. Conversely, fewer life events were associated with less frequent drinking. Similar effects were observed with age 17 drinking frequency; a greater number of life events were associated with more frequent alcohol use at age 17. The results from each of the models, testing for moderation effects associated with number of life events are displayed in Table 24 and graphically in Figure 11. The results for each of the outcomes based on these model fits are detailed below. Dropping additive genetic, shared environmental and unique environmental moderation effects of the number of life events significantly reduced model fit for age 14 drinking frequency. The importance of both additive genetics, shared and unique environment change as a function of the number of life events; figure 12 depicts the direction of these effects. Under conditions of more life events, both additive genetics, shared and unique environment play a more important role; conversely, in the conditions of fewer life events, genetics, shared and unique environment are attenuated. Dropping additive genetic and unique environmental moderating effects of the number of life events did not significantly reduce the fit of the model for drinking frequency at age 17; only dropping the moderating effect of life events on the shared environmental influences on age 17 drinking frequency significantly reduced model fit.

Table 24. Model Fitting Results from Twin Analyses

Freq Drinking 14			Δ Fit		
	Δ X2 units	Probability	Δ DF	AIC	LogLike
Full Model	--	--	2928	275.808	6131.808
Main Effect of Life Events	**488.96**	**<0.001**	1	486.963	6753.292
Additive Genetic Moderation	**12.03**	**<0.001**	1	10.027	6143.834
Shared Environment Moderation	**9.01**	**<0.001**	1	7.005	6140.813
Unique Environment Moderation	**13.26**	**<0.001**	1	11.257	6145.065

Freq Drinking 17			Δ Fit		
	Δ X2 units	Probability	Δ DF	AIC	LogLike
Full Model	--	--	2928	3107.976	8963.976
Main Effect of Life Events	**249.79**	**<0.001**	1	247.794	9301.954
Additive Genetics Moderation	2.274	0.132	1	0.274	8966.25
Shared Environment Moderation	**28.53**	**<0.001**	1	26.525	9002.501
Unique Environment Moderation	1.069	0.0792	1	-1.931	8964.046

Figure 11. Depiction of Twin Moderation Models: The moderating effects of life events on the genetic and environmental influences on drinking frequency at age 14.



Genetic Association Analyses

No genome-wide significant associations were observed; the best association observed for age 14 drinking frequency was rs10101663 (an intergenic SNP downstream of the adenylate cyclase 8 gene [*ADCY3*]) on chromosome 8 with a p-value of 1.2x 10^{-7}, and for age 17 drinking frequency was rs2367979 (an intergenic SNP downstream of the G protein-coupled receptor 158 gene [GPR158] and upstream of the myosin IIIA gene [MYO3A]) on chromosome 10 with a p-value of $5.7x10^{-7}$. Based on the small sample size, these results are not unexpected as we know that the sample is underpowered to detect SNPs of small effect at the genome-wide significance level. This is part of the rationale for focusing on the polygenic scores, which can give an overall index of risk even absent the power to detect individual signals[18]. 1,397 SNPs showed nominal association at p<0.01 for drinking at age 14 and 1,307 SNPs showed nominal association with drinking at age 17.5.

Life events significantly predicted concurrent drinking frequency (age 14) and later drinking frequency (age 17); greater life events were associated with more frequent drinking at both age 14 and 17. Our results indicated that the interaction between greater number of life events and greater genetic risk for age 14 drinking frequency predicts greater frequency of drinking at age 14 but not at age 17. All results are detailed in table 25 and the direction of effect is depicted in figure 12.

Table 25. Genetic Risk Score x Life Events effects on Drinking Frequency

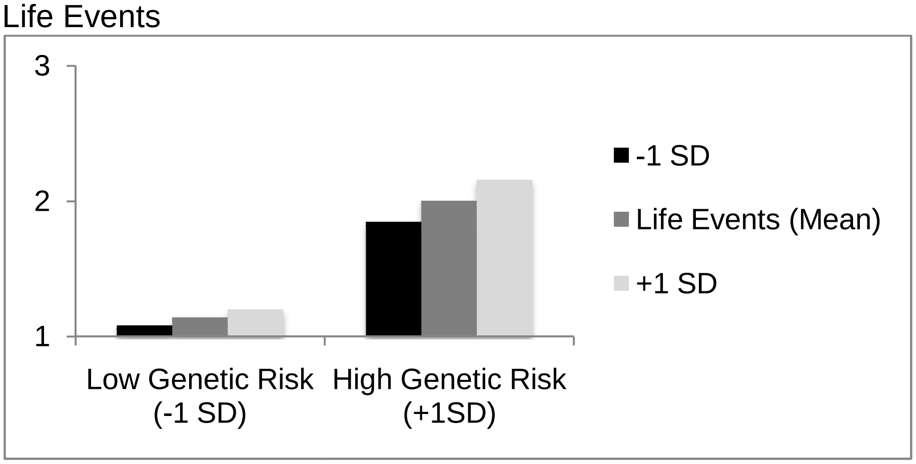
Model	Predictors included in model	Outcome: Drinking Frequency					
		Age 14			Age 17.5		
		ß	*p-value*	$R^2\Delta$	ß	*p-value*	$R^2\Delta$
I	Sex, Genetic Risk Score	0.54	7.9×10^{-6}	0.295	0.87	4.5×10^{-4}	0.801
II	Life Events	0.12	2.3×10^{-10}	0.016	0.06	3.4×10^{-4}	0.004
V	Sex, Genetic Risk Score, Life Events, Genetic Risk Score x Life Events	1.99	0.054	0.003	0.09	0.606	0.000

Note: In the main effects models (models I-IV), the $R^2\Delta$ refers to the proportion of variance accounted for by the listed variable (GRSS, life events). In the model that test for interaction effects (model III), the $R^2\Delta$ refers to the proportion of variance accounted for by the listed interaction term, after accounting for the main effects listed.

*p<0.05

Figure 12. Depiction of the interaction between the genetic sum score for age 14 drinking frequency and life events as a predictor of age 14 drinking frequency. Note: In the following figures, high genetic risk refers to +/-1 SD from the mean.

a. Life Events



The genetic risk score correlated with age 14 drinking frequency at r=0.54 and with age 17.5 drinking frequency at r=0.87. The magnitude of these correlations clearly reflects the fact that these scores consist of a number of false positives that capitalize on chance properties in the sample. To this end, we examined the association between this same age 14 drinking frequency derived genetic risk score and an external outcome we know to be genetically related to the drinking frequency, behavior problems as measured by the Multidimensional Peer Nomination Inventory (Pulkkinen, Kaprio, & Rose, 1999). The phenotypic association between age 14 drinking frequency derived GRSS and age 14 behavior problems is r=0.26. The age 14 drinking frequency derived GRSS correlated with age 14 behavior problems at r=0.23. This association suggests that the GRSS harbors some real risk variants and a number of false positives that will likely diminish the predictive ability of this GRSS in an independent sample. To check that our findings were not purely driven by chance results, we simulated a null distribution of GWAS results by random shuffling of the phenotypes. We created polygene scores using the same parameters as before based on these null simulations; accordingly, these polygene scores will entirely reflect capitalization on chance. We tested for interaction between the null polygene scores and each of the moderators. We repeated this process 100 times. The mean correlation between the null genetic risk scores and outcome was 0.24 (SD=0.01) at age 14 and 0.26 (SD=0.01) at age 17.5, reflecting the degree to which the genetic risk scores can be attributed purely to random chance. However, the interaction between the null GRSS and life events with the simulated null genetic risk scores was not significant, suggesting that the significant interactions detected in our data are not due purely to statistical artifacts purely associated with false positive findings encompassed in the calculation of the genetic risk scores.

In an effort to further reduce the noise included in the sum scores, we recalculated the sum scores using only those SNPs yielding $p<0.01$ at both age 14 and 17. 416 (15.4%) of SNPs were overlapping between the two risk scores. 29.8% of the SNPs showing $p<0.01$ at age 14 were also significant at this level at age 17, and 31.8% of the SNPs showing $p<0.01$ at age 17 were also significant at this level at age 14. When we weighted this subset of SNPs using the age 17 weights (chosen based on the higher heritability at that age) and recalculated the GxE results, all life events showed highly significant interaction effects ($p<0.001$).

Discussion

The goal of the present study was to use two complementary methods to examine the role of early adolescent life events as a moderator of genetic influences on the frequency of adolescent alcohol use, both concurrently and three years later. Consistent with past research showing that life events are positively associated with increased heritability for broadband externalizing disorders (Hicks et al., 2009), the findings from the present study indicate that stressful life events amplified the additive genetic effects to predict concurrent drinking frequency. Also in line with a previous report (Hicks et al., 2009), higher numbers of life events also increased shared and non-shared environmental effects. Note that under conditions of less life events, overall variance in adolescent drinking is diminished as compared with the overall variance in drinking under conditions of more life events. Moreover, under conditions of less life events, there is little variation in drinking frequency for both individuals with and without genetic risk. Variation in drinking frequency is maximized in conditions of more life events. This dual moderation of genetic and environmental effects in predicting frequent early adolescent alcohol use lend further support to the principle that life circumstances marked by unpredictability or change may allow for greater expression of genetic predispositions (Hicks et al., 2009; Shanahan & Hofer, 2005) and/or may render individuals susceptible to the influence of environments related to adolescent alcohol use, such as deviant peers, problematic relationships with parents, or other idiosyncratic experiences such as trauma.

Turning to the question of the legacy of life events, the second twin model indicated a positive main effect for early adolescent life events to predict age 17 drinking frequency. However, there was no evidence that life events moderated genetic influences. This suggests that the moderating effect of early adolescent life events on genetic influences for alcohol use is time-limited. In contrast, early adolescent life events moderated later shared environmental effects, such that their influence increased under greater numbers of life events. Thus, experiencing a greater number of life events in early adolescence may sensitize individuals to the effects of shared environmental risks later in adolescence. In the past, research has focused on the role of prenatal or neonatal life stress in sensitizing individuals to effects of later alcohol use risk factors (Clarke et al., 2011; Higley et al., 1991; Schneider et al., 2002; van Waes et al., 2011). The present results highlight the need to examine early adolescent life stress as a moderator of later environments to predict alcohol use. This is consistent with evidence that adolescence is a period of significant biopsychosocial reorganization (Graber and Brooks-Gunn, 1996; Cicchetti and Rogosch, 2002), and suggests that stressful experiences during this period may have long-lasting consequences.

In the second part of our study, we built upon the findings from the twin models in using genetic risk sum scores (Yang et al., 2010; The International Schizophrenia Consortium, 2009). On a zero-order level, the intercorrelations between genetic risk sum scores and alcohol frequency reveal several interesting effects. Genetic risk sum scores at ages 14 and 17 were only modestly inter-correlated, indicating that different sets of genes are related to frequency of alcohol use for these two ages. Life events were not significantly associated with genetic risk scores at either age, suggesting that frequent alcohol use and life events may not have a shared genetic liability in this age group and reducing concern that the moderation effects would be driven by gene-environment correlation.

As anticipated, given the twin model results, age 14 stressful life events moderated genetic risk sum scores to predict age 14 drinking frequency, but not age 17 drinking frequency. Further, the interaction findings were not significant using a simulated null polygenic risk score. This suggests that although polygenic risk scores are known to encompass both real and false positive effects, the findings are not entirely driven by chance effects encompassed in the creation of polygenic scores in any given sample. An effort to further reduce the noise in the genetic risk score by including only those SNPs included in the calculation of the score at both ages 14 and 17 further increased the significance of the interaction terms. Those who experienced a greater number of life events between ages 13-14 and who were at higher genetic risk drank most often at age 14. Meanwhile, those at low genetic risk drank least often at age 14, even in the context of high life events. Furthermore, the pattern observed here reaffirms the principle that genetic risk can take on a different meaning depending on one's environment.

Our results should be interpreted in the context of the several limitations. One limitation of the study is that we did not verify the predictive power of our polygenic risk scores in an independent sample. To this end, while the genetic risk score accounted for a substantial proportion of the variance in drinking frequency (as expected being that this was the phenotype it was derived from), the proportion of variance in behavior problems (a phenotype genetically related to adolescent drinking frequency) accounted for by the genetic risk score dropped to 1%. In our previous analyses using similarly constructed GWAS risk scores in a sample of similar size, we found that 56% of the variance in alcohol dependence symptoms was accounted for in the discovery sample, whereas only 1% was accounted for in the replication sample (Yan et al., in preparation), consistent with previous analyses of this sort showing the small overall percentage of variance accounted for even by sum scores. Another limitation is that we used a threshold of all snps with $p<0.01$ in the creation of the polygenic risk scores, which is somewhat arbitrary. There are of course several ways to create aggregate risk scores (Evans et al, 2009). Previous studies have shown that risk prediction increases up to a certain point, but then decreases as more false positives are included, overshadowing the real effects that are encompassed [24]. Posthoc analyses of our data suggested that the interaction effects became less significant as the p-value threshold for inclusion of SNPs in the polygenic score became less stringent.

In addition to these limitations, life events were measured only at age 14, and so we are unable to determine the relative influence of early versus later adolescent life events. Although our sample is population based, it is racially homogenous and generalizability to other populations may be limited. Lastly, our measure of life events taps primarily normative stressors. Extreme stressors (e.g., developmental trauma or natural disasters) may moderate genetic risk in a different way, or may swamp genetic risk entirely.

In conclusion, this study brings together latent and measured genetic approaches to better understand how genetic predispositions interact with stressful life events to predict alcohol use frequency across adolescence. We provide new evidence that higher levels of stressful life events increases genetic risk for frequent alcohol use in early adolescence, that some of the genes associated with frequent alcohol use differ between early and later adolescence, and that higher life events amplify the association between high genetic risk and early adolescent alcohol frequency. These findings highlight the benefits of using multiple methods to elucidate the presence and mechanisms of gene-environment interactions in order to better understand the etiology of adolescent alcohol use.

References

Barr CS, Newman TK, Lindell S, et al. (2004) Interaction between serotonin transporter gene variation and rearing condition in alcohol preference and consumption in female primates. *Arch Gen Psychiatry* 61: 1146-1152.

Blomeyer D, Treutlein J, Esser G, et al. (2008) Interaction between CRHR1 gene and stressful life events predicts adolescent heavy alcohol use. *Biological Psychiatry* 63: 146-151.

Brown SA, McGue M, Maggs J, et al. (2008) A developmental perspective on alcohol and youths 16 to 20 years of age. *Pediatrics* 121: S290-S310.

Cicchetti D and Rogosch FA. (2002) A developmental psychopathology perspective on adolescence. *Journal of Consulting and Clinical Psychology* 70: 6-20.

Clark DB, Lesnick L and Hegedus AM. (1997) Traumas and other adverse life events in adolescents with alcohol abuse and dependence. *Journal of American Academy of Child and Adolescent Psychiatry* 36: 1744-1751.

Clarke TK, Laucht M, Ridinger M, et al. (2011) KCNJ6 is associated with adult alcohol dependence and involved in gene x early life stress interactions in adolescent alcohol drinking. *Neuropsychopharmacology* 36: 1142-1148.

Covault J, Tennen H, Armeli S, et al. (2007) Interactive effects of the serotonin transporter 5-HTTLPR polymorphism and stressful life events on college student drinking and drug use. *Biological Psychiatry* 61: 609-616.

Gaffney LR, Thorpe K, Young R, et al. (1998) Social skills, expectancies, and drinking in adolescents. *Addictive Behaviors* 23: 587-599.

Graber JA and Brooks-Gunn J. (1996) Transitions and turning points: Navigating the passage from childhood through adolescence. *Developmental Psychology* 32: 768-776.

Hicks BM, South SC, DiRago AC, et al. (2009) Environmental adversity and increasing genetic risk for externalizing disorders. *Archives of General Psychiatry* 66: 640-648.

Higley JD, Hasert MF, Suomi SJ, et al. (1991) Nonhuman primate model of alcohol abuse: effects of early experience, personality, and stress on alcohol consumption. *Proc Natl Acad Sci U S A* 88: 7261-7265.

Jelalian E, Alday S, Spirito A, et al. (2000) Adolescent motor vehicle crashes: The relationship between behavioral factors and self-reported injury. *Journal of Adolescent Health* 27: 84-93.

Johnston LD, O'Malley PM, Bachman JG, et al. (2006) Monitoring the Future: National survey results on drug use, 1975–2005. Bethesda, MD: National Institute on Drug Abuse.

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

Kendler KS, Gardner C and Dick DM. (2011) Predicting alcohol consumption in adolescence from alcohol-specific and general externalizing genetic risk factors, key environmental exposures and their interaction *Psychological Medicine* 41: 1507-1516.

Laucht M, Treutlein J, Schmid B, et al. (2009) Impact of psychosocial adversity on alcohol intake in young adults: moderation by the LL genotype of the serotonin transporter polymorphism. *Biol Psychiatry* 66: 102-109.

Maes HH, Woodard CE, Murrelle L, et al. (1999) Tobacco, alcohol and drug use in eight- to sixteen-year-old twins: The Virginia Twin Study of Adolescent Behavioral Development. *Journal of Studies on Alcohol* 6: 293-305.

Rose RJ, Dick DM, Viken RJ, et al. (2001) Gene-environment interaction in patterns of adolescent drinking: Regional residency moderates longitudinal influences in alcohol use. *Alcoholism: Clinical and Experimental Research* 25: 637-643.

Schmid B, Blomeyer D, Treutlein J, et al. (2010) Interacting effects of CRHR1 gene and stressful life events on drinking initiation and progression among 19-year-olds. *Int J Neuropsychopharmacol* 13: 703-714.

Schneider ML, Moore CF, Kraemer GW, et al. (2002) The impact of prenatal stress, fetal alcohol exposure, or both on development: perspectives from a primate model. *Psychoneuroendocrinology* 27: 285-298.

Schulenberg JE and Maggs JL. (2002) A developmental perspective on alcohol use and heavy drinking during adolescenec and the transition to young adulthood. *Journal of Studies on Alcohol* Supplement No. 14: 54-70.

Steinberg L. (2005) Cognitive and affective development in adolescence. *Trends in cognitive sciences* 9: 69-74.

The International Schizophrenia Consortium. (2009) Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 460.

van Beek JHDAK, K. S., de Moor MHM, Geels LM, et al. (2011) Stable genetic effects on symptoms of alcohol abuse and dependence from adolescence into early adulthood. *Behavior Genetics*.

van Waes V, Darnaudéry M, Marrocco J, et al. (2011) Impact of early life stress on alcohol consumption and on the short- and long-term responses to alcohol in adolescent female rats. *Behavioural Brain Research* 221: 43-49.

Windle M and Windle RC. (1996) Coping strategies, drinking motives, and stressful life events among middle adolescents: Associations with emotional and behavioral problems and with academic functioning. *Journal of Abnormal Psychology* 105: 551-560.

Yang J, Benyamin B, McEvoy BP, et al. (2010) Common SNPs explain a large proportion of variability for human height. *Nature Genetics* 42: 565-571.

Chapter 6

The Interaction between Adolescent Parental Knowledge and Genetic Risk for Alcohol Dependence Predicts Adult Alcohol Dependence

Abstract

Background: Previous studies demonstrate that parental knowledge moderates latent genetic influences on adolescent externalizing behavior and alcohol use (Dick et al., 2007, Latendresse et al., 2010) as well as specific genetic predispositions, such as *CHRM2*, to predict adolescent externalizing behavior (Dick et al., 2009). Little is known however, about the longitudinal effects of the parental knowledge in moderating genetic risk for alcohol problems from adolescence into adulthood.

Methods: This study examines whether parental knowledge in adolescence continues to moderate genetic influences on alcohol use in young adulthood. We approached this question using data from a longitudinal, population based twin sample, *Finntwin12* (Kaprio et al., 1999). We first conducted twin analyses to examine whether parental knowledge (measured at age 14) moderated genetic and environmental influences on alcohol dependence symptoms at age 22. We then created genetic risk sum scores (Yang et al., 2009) using GWAS data available on the twins (scores were comprised of all SNPs associated at $p<0.01$ with DSM-IV Alcohol Dependence symptoms). Next, we examined the interaction between this aggregate measure of risk genes and parental knowledge, and its effect on age 22 alcohol dependence symptoms.

Results: The twin analyses indicated that parental knowledge significantly moderates the genetic influences on alcohol dependence symptoms at age 22 (χ^2=10.31, p<0.0001). The genotypic analyses indicated that the interaction between genetic risk sum scores and parental knowledge significantly predicted alcohol dependence symptoms at age 22 (β=0.308, p<0.001).

Conclusion: Converging evidence from two analytic methods suggests that parental knowledge in adolescence has an enduring moderating influence on genetic predispositions to alcohol use disorders in young adulthood. Parental knowledge may be an important proxy for some stable aspect of the individual's environment from adolescence into early adulthood, or may scaffold the adolescent's burgeoning behavioral regulation skills. There is a need for future research to elucidate the depth and limitations of the lasting effects of this aspect of adolescent parenting throughout development.

Introduction

Low levels of parental knowledge, or the degree to which a parent is aware of his/her child's whereabouts and actions, are associated with externalizing problems in adolescence, including more frequent adolescent drug and alcohol use (Marshal et al., 2000; Johnstone et al., 1994; Windle et al., 2000; Leventhal et al., 2000; Barnes et al., 1992; Steinberg et al.1994; Chilcoat et al, 1996). In addition, twin studies indicate that parental knowledge moderates latent genetic influences on adolescent externalizing behavior (Dick et al. 2007) and frequency of alcohol use (Meyers et al., 2012 *under review*) throughout adolescence. Moreover, several studies that implement measured genotypic data also find that the interaction between specific genetic variants (e.g., *CHRM2, GABRA2)* and parental knowledge predict adolescent externalizing behavior (Dick et al., 2011) and risk trajectories (Dick et al., 2009). In a recent study, Kendler and colleagues reported a significant interaction between parental monitoring and genetic risk for externalizing behavior and alcohol use disorders as a predictor of alcohol use frequency from ages 12-14 (Kendler et al., 2011). These analyses all suggest that when adolescents report that their parents know little about their whereabouts, associations, and behavior (i.e., less parental knowledge), it creates an environment that allows for greater opportunity to express genetic predispositions for risky alcohol use behavior. These results are in line with previous findings from the Finnish Twin Studies, which indicate that in less stable neighborhoods, where there was presumably less community monitoring, genetic influences on alcohol use frequency become more important (Rose et al., 2001; Dick et al., 2001; Dick et al., 2009).

These cross-sectional and short-term longitudinal effects, whereby parental monitoring moderates genetic influences on adolescent externalizing-spectrum behavior (including alcohol use), beg the question of whether these effects are implicated in the development of young adult alcohol problems. Is parental knowledge tapping into an adolescent-limited phenomenon whereby low levels of parental knowledge in adolescence contribute to greater genetic risk for concurrent alcohol use, and these effects diminish once the adolescent is out of this environment (ie. moves out of the home)? Or, does adolescent parental knowledge continue to impact the individual's behavior into young adulthood? From the perspective that high levels of parental knowledge provide youth an appropriate balance of opportunities to explore their own autonomy while also maintaining one's connection to parents (Pettit et al., 2001), one would expect to observe such enduring effects.

The present study examines whether adolescent parental knowledge continues to moderate genetic influences on alcohol use once the adolescent enters young adulthood. We approached this question using two different methods in a population based twin sample, *Finntwin12*. We first conducted twin analyses that examined the moderating effects of parental knowledge (measured in adolescence) on the latent genetic and environmental contributions to alcohol dependence in young adulthood. We then attempted to address the same research question using measured genotypic data available on the sample. We created genetic risk sum scores (GRSS) using genome wide association study (GWAS) data with the ultimate goal of distinguishing whether the interaction between parental knowledge (measured in adolescence) and genetic risk for alcohol dependence is adolescent limited or has persisting effects on adult alcohol dependence. Instead of testing individual loci sequentially, genetic risk sum scores (GRSS) can be constructed and tested to summarize the total number of risk alleles (Yang et al., 2009; Aulchenko et al. 2009). Using these two complementary methods, we examine whether adolescent parental knowledge moderates aggregate genetic risk on young adult alcohol dependence symptoms.

Methods

Sample

Finntwin12 has been described previously (chapters 3, 4 and 5). For the present study, parental knowledge was measured at age 14 and DSM-IV Alcohol Dependence criteria were measured at age 22, when genotypic data was collected on a subset of individuals.

Assessment

Parental Knowledge (Knowledge) was assessed with four questions included in the twins' questionnaire administered at age 14. The questions, created by Chassin and colleagues (Chassin et al. 1993), asked the adolescents to report on the degree to which their parents (1) know about their daily plans (2) know of their interests, activities, and whereabouts (3) know how they spend their money, and (4) know where and with whom they are outside of the home. Responses were made on a 4-point scale ranging from 1 (*rarely or never*) to 4 (*almost always*), so that *greater* scores indicate *more* parental knowledge. A sum score based on the tallying of these items was created on 4,542 adolescents. We note that we have previously referred to this measure as “parental monitoring” in Finnish Twin Study publications, however, this variable likely reflects both solicited information and spontaneous information provided by the child and therefore we will refer to this measure as parental knowledge (Kerr & Stattin 2000). Scores for parental knowledge ranged from 1 to 16 (M=6.5, SD=2.14).

DSM-IV Alcohol Dependence Symptoms (ADSX) were derived from the Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA; Bucholz et al., 1994) interviews administered at age 22. The SSAGA indexed lifetime prevalence of the seven DSM-IV criteria for Alcohol Dependence (American Psychiatric Association, 1994), including (1) tolerance, (2) withdrawal, (3) drinking in amounts or timeframes larger than intended (4) unsuccessful efforts to cut down on use (5) spending a large amount of time obtaining, using, or recovering from alcohol, (6) important activities reduced because of use, (7) use despite physiological or psychological consequences. Scores for ADSX ranged from 0 to 7 (M=1.75, SD=1.45).

Twin Modeling

Comparisons of the similarity of MZ and DZ twin pairs yield information about the degree of influence that can be attributed to genetic and environmental factors for a particular outcome (Plomin et al. 2001). The basic genetically informative twin model partitions variance in a behavior into additive genetic influences (A), dominant genetic influences (D), common environmental influences or (C), and unique environmental influences (E). Genetic influences correlate 1.0 between monozygotic (MZ) twins, who share all of their genes identical-by-descent, and 0.5 between dizygotic (DZ) twins, who share, on average, 50% of their segregating genes, as do ordinary siblings. Common environmental effects, as defined in biometrical twin modeling, refer to all environmental influences that make siblings more similar to one another. By definition, these influences correlate 1.0 between both MZ and DZ twins. Unique environmental influences are uncorrelated between co-twins and have the effect of decreasing the covariance between siblings. As dominant genetic influences (D) and common environmental influences (C) cannot be simultaneously modeled in twin-only data, we modeled common environmental influences (C) because the DZ twin correlation exceeded ½ of the MZ twin correlation for each of the present study's outcome.

Moderation models were fit to test whether the variance components for alcohol dependence symptom count differed as a function of common and unique environmental factors. Chapter 4, Figure 9 shows a classic twin model (for only 1 twin in the pair) that has been modified to include a moderation component (Purcell 2002). The standard paths a, c, and e, indicating the magnitude of effect of additive genetic influences, common environmental influences, and unique environmental influences, now each include a β term, which indicates the significance of a potential moderator variable M on each of these genetic and environmental influences. The value of M changes from subject to subject, taking on the value of the measured variable for that subject (i.e., parental knowledge in our models). In the moderation model, the additive genetic value is a linear function of the moderator M, represented by the equation a + β XM, where βX is an unknown parameter to be estimated from the data, representing the magnitude of the moderating effect. If βX is significantly different from zero, there is evidence for a moderating effect. A similar logic follows for the βY and βZ pathways, which represent the extent to which a specific moderator variable alters the importance of common and unique environmental influences, respectively. In other words, the moderation model allows us to test whether the importance of additive genetic effects (a), common environmental effects (c), and unique environmental effects (e) change as a function of the measured variable. The pathway I + βMM models main effects of the moderator variable on the outcome. Also included in this pathway are any gene–environment correlation effects between the moderator variable and outcome. There is some evidence of genetic influence on parental knowledge (Kendler et al, 2007; Latendresse et al., 2010). However, previous analyses in this sample have suggested that even with genetic factors accounting for 27% of variance in knowledge, the correlation with alcohol use was largely environmentally mediated (Latendresse et al., 2010). Further, any covariance between the moderator and the outcome (and accordingly, any gene-environment correlation) is incorporated into the means model.

All modeling was conducted using the raw data option in Mx (Neale 2000). Mx is a structural equation-modeling program developed specifically for the use of twin and family data. The significance of each of the parameters in the model can be tested by dropping a parameter and evaluating the change in 2 log likelihood between the initial model and the nested submodel. This difference is evaluated using a chi-square distribution. A significant change in fit between the models ($p < 0.05$) for the difference in degrees of freedom indicates that dropping the parameter caused a significant decrease in fit of the model, indicating that pathway significantly contributes to the outcome trait and should be retained in the model.

Model-fitting proceeded in a series of steps. We first tested the significance of the main effect of the moderator (parental knowledge). We then tested the significance of moderation effects by dropping all moderation (3 df test, βX, βY, and βZ dropped). When this test was significant, we conducted further testing to determine what specific variance components showed significant moderation by sequentially dropping and testing the significance of each of the moderating effects one by one.

Preliminary power analyses suggested that there was low power to discriminate sex effects because of the large sample sizes necessary for adequate power to detect moderating effects with ordinal outcomes. Accordingly, female and male twins were collapsed by zygosity in modeling, though means and variances for ADSX were allowed to differ between the sexes when indicated by the data.

Genetic Association Analyses

To create genetic risk sum scores for each individual, we first ran a genome wide association analysis using the number of alcohol dependence symptoms endorsed at age 22 as the outcome. We then summed the top single nucleotide polymorphism (SNP) results to create a weighted genetic risk score for each individual. For the initial GWAS analysis, a linear regression adjusted for age and sex was performed for ADSX, as a quantitative trait using PLINK v1.07 (Purcell et al. 2007). Additionally, the family structure of the data was accounted for using a permutation procedure (qfam) performed in PLINK that randomly shuffles the degree of relatedness across all individuals. Because the qfam procedure can specify only one type of familial relationship, both individuals from each DZ pair and one individual from each MZ twin pair was included in the analyses, reducing the sample size from 1,069 to 866 individuals. GWAS results from the FT12 analyses of ADSX are described elsewhere (chapter 3). Briefly, no individual single nucleotide polymorphism (SNP) met genome wide criteria for significance in those analyses; however, many SNPs fell just below the threshold. The asymptotic *p* value for the linear regression was calculated and the effect size (beta) was estimated. We then summed the top SNP level results to create a weighted genetic risk score for each individual. All SNPs with nominal *p*- values less than 0.001 were included in the genetic risk sum score. Once genetic risk scores were computed for each individual, we used linear regression to test whether (1) parental knowledge predicted age 22 ADSX, (2) parental knowledge interacted with genetic risk sum scores to predict age 22 ADSX. Sex was used as a covariate in all analyses.

Results

Twin Analyses

Twin analyses indicated that parental knowledge had a significant main effect on ADSX (χ^2=76.92, p<0.001); less parental knowledge was associated with higher ADSX. In addition, parental knowledge significantly moderated the additive genetic, shared, and unique environmental influences on ADSX. As shown in Figure 13, genetic factors had a greater influence on ADSX in early adulthood for individuals who reported low levels of parental knowledge in adolescence. Conversely, shared and unique environmental factors had less of an influence on ADSX in early adulthood for those who reported low levels of parental knowledge in adolescence.

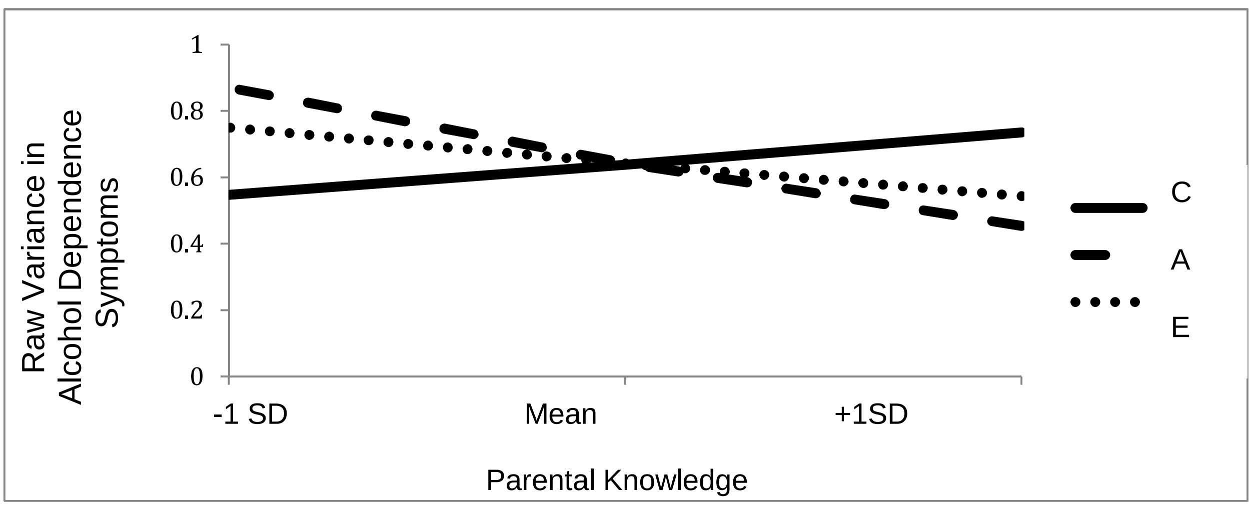
Table 26. Model Fit Statistics from Twin Moderation Models

Alcohol Dependence Symptoms	Model Fit				
	Δ X^2 units	Probability	Δ DF	AIC*	LogLike**
Full Model	--	--	4828	20893.92	30549.92
Main Effect of Parental Knowledge	76.924	<0.0001	1	21228.70	30892.70
Additive Genetic Moderation	10.315	<0.0001	1	20911.23	30569.23
Shared Environment Moderation	22.796	<0.0001	1	20914.71	30572.71
Unique Environment Moderation	50.674	<0.0001	1	20942.59	30600.59

*Akaike's Information Criterion
**-2 times log-likelihood of the data

Figure 13. Latent genetic and environmental influences (raw variance estimates) on alcohol dependence symptom count change as a function of parental knowledge

Note: The parental knowledge scale is coded so that low scores (-1 standard deviation) indicate less parental knowledge and high scores (+1 standard deviation) indicate more parental knowledge.

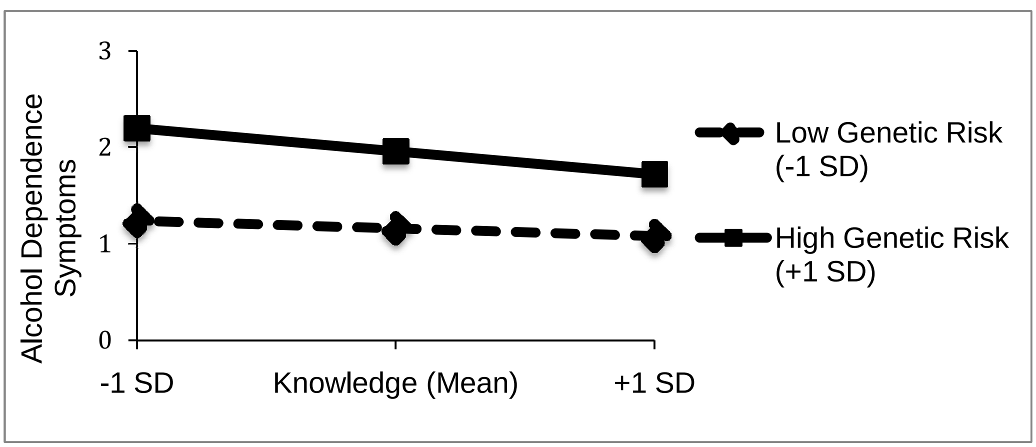


Genetic Risk Sum Scores Analysis

The genetic risk sum scores ranged from -0.065 to 1.27 (M=0.363, SD=0.174) and consisted of 177 SNPs from 85 genes (most significant association results detailed in chapter 3). Genetic risk sum scores were not associated with parental knowledge (r=0.051, p=0.138), suggesting there was no appreciable gene-environment correlation within this sample. The genotypic analyses indicated that parental knowledge moderated GRSS to predict age 22 ADSX (β=0.308, $p<0.001$). As shown in Figure 14, the association between ADSX and adolescent parental knowledge was stronger for those at higher genetic risk for alcohol dependence compared to those who were at lower genetic risk.

Figure 14. Depiction of the interaction between the genetic risk sum scores and parental knowledge as a predictor of age 22 alcohol dependence symptoms

Note: The parental knowledge scale is coded so that low scores (-1 standard deviation) indicate less parental knowledge and high scores (+1 standard deviation) indicate more parental knowledge.



The genetic risk score correlated with ADSX at r=0.69. This estimate is largely inflated as we know the GRSS to consist of some real signal and some false-positives produced by the discovery sample bias; the predictive power of this GRSS would dramatically decrease in an independent (replication) sample (Yang et al., 2011). Seeking validation that this GRSS consisted of some real signal, we examined the relationship between this ADSX GRSS and three external phenotypes that we know share genetic risk with ADSX: smoking frequency, conduct disorder and adult antisocial behavior. The ADSX GRSS correlated with smoking frequency r=0.20 (phenotypic r=0.34), and accounted for 3.7% of the variance in the phenotype. The ADSX GRSS correlated with DSM-IV conduct disorder r=0.264 (phenotypic r=0.269), and accounted for 6.9% of the variance in the phenotype. The ADSX GRSS correlated with DSM-IV adult antisocial behavior r=0.334 (phenotypic r=0.379), and accounted for 11.1% of the variance in the phenotype. These associations suggests that the ADSX GRSS harbors some real risk variants and a number of false positives that will likely diminish the predictive ability of this GRSS in an independent sample. To check that our findings were not purely driven by chance results, we simulated a null distribution of GWAS results by random shuffling of the phenotypes. We created polygene scores using the same parameters as before based on these null simulations; accordingly, these polygene scores will entirely reflect capitalization on chance. We tested for interaction between the null polygene scores and each of the moderators. We repeated this process 100 times. The mean correlation between the null genetic risk scores and outcome was .34 (SD=0.01), reflecting the degree to which the genetic risk scores can be attributed purely to random chance. However, the interaction between the null

GRSS and parental knowledge with the simulated null genetic risk scores was not significant, suggesting that the significant interactions detected in our data are not due purely to statistical artifacts purely associated with false positive findings encompassed in the calculation of the genetic risk score.

Conclusions

A substantial literature has examined the effects of parenting on adolescent alcohol use (Luyckx et al., 2011). Recently, a growing number of studies have examined the interaction between specific aspects of adolescent parenting (and other features of adolescents' social environments) and genetic predispositions to adolescent alcohol use and problems (Enoch, 2012). These gene-environment interaction effects have primarily been explored in the context of cross-sectional and short-term longitudinal studies in adolescence, a period of time when individuals are particularly susceptible to input from their surroundings (Swendsen et al., 2012). In the present study, we extend this literature to examine the enduring effects of one key environmental moderator, adolescent parental knowledge, on adult alcohol dependence symptoms.

In the present study, we provide converging evidence from two analytic methods that the interactive effects observed between parental knowledge in adolescence and genetic predispositions predict alcohol use disorder symptoms in young adulthood. The twin models provide a bird's eye view of this gene-environment interaction and indicate that under conditions of less parental knowledge in adolescence (age 14), latent genetic influences on alcohol dependence symptoms at age 22 are more important than for those who reported greater parental knowledge in adolescence.

The genetic risk sum score data provides further detail in fleshing out this latent model, while still measuring aggregate genetic risk. For individuals who reported less parental knowledge in adolescence, the association between genetic risk factors for alcohol dependence symptoms and alcohol dependence symptoms at age 22 was stronger. In contrast, the association between genetic risk and alcohol dependence symptoms was weaker for those who reported more parental knowledge in adolescence. This parallels the direction of effect reported in previous *Finntwin12* publications that examined these moderation effects in adolescence (Dick et al., 2007, Meyers et al., 2012). Such findings conceptually map onto Shanahan and Hofer's (2005) mechanism of social opportunity versus social control. That is, we have previously hypothesized that lower rates of parental knowledge provide an opportunity for an adolescent to express his/her genetic predisposition for alcohol dependence symptoms, whereas higher parental knowledge may suppress the expression of these same genetic predispositions. The present study extends past work showing that parental knowledge in late middle childhood and early adolescence protects against adolescent alcohol (Dick et al., 2009; Meyers et al., 2012) and substance use (Bohnert et al. 2012) by demonstrating that these effects are carried forward into early adulthood as well.

So the question becomes why parental knowledge measured in adolescence remains relevant in early adulthood. What mediates the relationship between adolescent parental knowledge and symptoms of adult alcohol dependence? Previous studies indicate that adolescent perceptions of various aspects of parenting are positively correlated with measures of warmth and responsiveness and negatively associated with conflict, autocratic parenting, discipline and relational tension (Knofo and Schwartz, 2003; Latendresse et al., 2010). Thus, it may be the case that parental knowledge is a proxy for related dimensions of parenting that are a stable aspect of the individual's environment from adolescence into early adulthood. Alternatively, parental knowledge during this critical developmental period, where adolescents and their parents negotiate autonomy and connectedness (Erikson, 1963; Pettit et al., 2001), may scaffold the adolescent's burgeoning ability to regulate his/her own behavior. Historically, the parenting literature has emphasized the legacy of early child-caregiver experiences for later behavioral regulation (Sroufe et al., 2005), including alcohol use (Englund et al., 2008). The results from the present analyses suggest that specific aspects of later parenting (e.g., parental knowledge in adolescence) may have comparable long-lasting effects.

Our findings should be interpreted in the context of several limitations. First, parental knowledge was only measured once in this sample, and so we are unable to determine the relative influence of early versus later adolescent influences. Second, we did not verify the predictive power of our polygenic risk scores in an independent sample. In previous analyses using similarly constructed GWAS risk scores in a sample of similar size, we found that 56% of the variance in alcohol dependence symptoms was accounted for in the discovery sample, whereas only 1% was accounted for in the replication sample, consistent with previous analyses of this sort showing the small overall percentage of variance accounted for even by sum scores. Another limitation is that we used a threshold of all snps with $p<0.01$ in the creation of the polygenic risk scores, which is somewhat arbitrary. There are of course several ways to create aggregate risk scores (Evans et al, 2009). Previous studies have shown that risk prediction increases up to a certain point, but then decreases as more false positives are included, overshadowing the real effects that are encompassed [24]. Posthoc analyses of our data suggested that the interaction effects became less significant as the p-value threshold for inclusion of SNPs in the polygenic score became less stringent.

In summary, adolescent parental knowledge moderates both latent and measured aggregate genetic predispositions for young adult alcohol dependence symptoms. Our findings suggest that interventions aimed at boosting parental knowledge in adolescence may be one approach to prevent problematic alcohol use in young adulthood. However, future research aimed at elucidating the depth and limitations of the lasting effects of adolescent parenting throughout development is needed.

References

Dick, D. M., Pagan, J. L., Holliday, C., Viken, R., Pulkkinen, L., Kaprio, J., & Rose, R. J. (2007a) Gender Differences in Friends' Influences on Adolescent Drinking: A Genetic Epidemiological Study. Alcoholism: Clinical and Experimental Research, 31: 2012-19.

Dick, D. M., Pagan, J. L., Viken, R., Purcell, S., Kaprio, J., Pulkkinen, L., & Rose, R. J. (2007b) Changing Environmental Influences on Substance Use Across Development. Twin Research and Human Genetics, 10: 315-326.

Dick, D. M., Purcell, S., Viken, R. J., Kaprio, J., Pulkkinen, L., & Rose, R. J. (2007) Parental Monitoring Moderates the Importance of Genetic and Environmental Influences on Adolescent Smoking, *Journal of Abnormal Psychology*, 116(1):213-8.

Latendresse SJ, Rose RJ, Viken RJ, Pulkkinen L, Kaprio J, Dick DM. Examining the etiology of associations between perceived parenting and adolescents' alcohol use: common genetic and/or environmental liabilities? J Stud Alcohol Drugs. 2010 May;71(3):313-25.

Dick, D. M., Bernard, M., Aliev, F., Viken, R., Pulkkinen, L., Kaprio, J., & Rose, R. J. (2009) The Role of Socio-Regional Factors in Moderating Genetic Influences on Early Adolescent Behavior Problems and Alcohol Use. *Alcoholism: Clinical and Experimental Research*, 33 (10):1739-48.
Kaprio et al., 1999?

Kaprio J, Pulkkinen L, Rose RJ. Genetic and environmental factors in health-related behaviors: studies on Finnish twins and twin families. Twin Res.
2002 Oct;5(5):366-71. Review.

International Schizophrenia Consortium, Purcell SM, Wray NR, Stone JL,
Visscher PM, O'Donovan MC, Sullivan PF, Sklar P. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature. 2009 Aug 6;460(7256):748-52.

Marshal MP, Chassin L. Peer influence on adolescent alcohol use: the moderating role of parental support and discipline. *Appl Dev Sci.* 2000;4(2):80-88.

Johnstone BM. Sociodemographic, environmental, and cultural influences on adolescent drinking behavior. In: Zucker R, Boyd G, Howard J, eds. *The Development of Alcohol Problems: Exploring the Biopsychosocial Matrix of Risk.* Rockville, MD: NIH Publication; 1994:169-203.

Windle M. Parental, sibling, and peer influences on adolescent substance use and alcohol problems. *Appl Dev Sci.* 2000;4(2):98-110.

Leventhal T, Brooks-Gunn J. The neighborhoods they live in: the effects of neighborhood residence on child and adolescent outcomes. *Psychol Bull.* 2000;126(2):309-337.

Barnes GM, Farrell MP. Parental support and control as predictors of adolescent drinking, delinquency, and related problem behaviors. *J Marriage Fam.* 1992;54:763-776.

Steinberg L, Fletcher A, Darling N. Parental monitoring and peer influences on adolescent substance use. *Pediatrics.* 1994;93(6 pt 2):1060-1064.

Chilcoat HD, Anthony JC. Impact of parental monitoring on initiation of drug use through late childhood. *J Am Acad Child Adolesc Psychiatry.* 1996;35(1):91-100.

Meyers JL, Latendresse SJ, Rose RJ, Viken RJ, Pulkkinen L, Kaprio J, Dick DM. Environmental Moderation of Alcohol Use in Adolescence: Common and/or Unique Influences. Under review in *Alcoholism: Clinical and Experimental Research***.** 2012

Kendler KS, Gardner C, Dick DM. Predicting alcohol consumption in adolescence from alcohol-specific and general externalizing genetic risk factors, key environmental exposures and their interaction. Psychol Med. 2011 Jul;41(7):1507-16.

Rose RJ, Dick DM, Viken And RJ, Kaprio J. Gene-environment interaction in patterns of adolescent drinking: regional residency moderates longitudinal influences on alcohol use. Alcohol Clin Exp Res. 2001 May;25(5):637-43.

Dick DM, Rose RJ, Viken RJ, Kaprio J, Koskenvuo M. Exploring gene environment interactions: socioregional moderation of alcohol use. J Abnorm Psychol. 2001 Nov;110(4):625-32.

Aulchenko YS, Ripatti S, Lindqvist I, Boomsma D, Heid IM, Pramstaller PP, Penninx BW, Janssens AC, Wilson JF, Spector T, Martin NG, Pedersen NL, Kyvik KO, Kaprio J, Hofman A, Freimer NB, Jarvelin MR, Gyllensten U, Campbell H, Rudan I, Johansson A, Marroni F, Hayward C, Vitart V, Jonasson I, Pattaro C, Wright A, Hastie N, Pichler I, Hicks AA, Falchi M, Willemsen G, Hottenga JJ, de Geus EJ, Montgomery GW, Whitfield J, Magnusson P, Saharinen J, Perola M, Silander K, Isaacs A, Sijbrands EJ, Uitterlinden

AG, Witteman JC, Oostra BA, Elliott P, Ruokonen A, Sabatti C, Gieger C, Meitinger T, Kronenberg F, Döring A, Wichmann HE, Smit JH, McCarthy MI, van Duijn CM, Peltonen L; ENGAGE Consortium. Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. Nat Genet. 2009 Jan;41(1):47-55.

Kaprio, J., Pulkkinen, L., & Rose, R. J. (2002) Genetic and environmental factors in health-related behaviors: studies on Finnish twins and twin families. *Twin Res*, 5(5):366-71.

Kaprio J. Twin studies in Finland 2006. Twin Res Hum Genet. 2006 Dec;9(6):772-7.

Chassin, L., Pillow, D., Curran, P., Molina, B., & Barrera, M. (1993) Relation of parental alcoholism to early adolescent substance use: A test of three mediating mechanisms. *Journal of Abnormal Psychology*, 102:3-19.

Kerr, M. & Stattin, H. (2000) What parents know, how they know it, and several forms of adolescent adjustment: Further support for a reinterpretation of monitoring. *Developmental Psychology*, 36: 366-380.

Rose RJ, Viken RJ, Dick DM, Bates JE, Pulkkinen L, Kaprio J. It does take a village: nonfamilial environments and children's behavior. Psychol Sci. 2003 May;14(3):273-7.

Bucholz KK, Cadoret R, Cloninger CR, Dinwiddie SH, Hesselbrock VM, Nurnberger JI Jr, Reich T, Schmidt I, Schuckit MA. A new, semi-structured psychiatric interview for use in genetic linkage studies: a report on the reliability of the SSAGA. J Stud Alcohol. 1994 Mar;55(2):149-58.

Association, A. P. 1994, *Diagnostic and statustical manual of mental disorders*, Fourth Edition edn, American Psychiatric Association, Washington, DC.

Plomin R, Asbury K, Dunn J. Why are children in the same family so different? Nonshared environment a decade later. Can J Psychiatry. 2001 Apr;46(3):225-33. Review.

Purcell S. Variance components models for gene-environment interaction in twin analysis. Twin Res. 2002 Dec;5(6):554-71.

Kendler KS, Baker JH. Genetic influences on measures of the environment: a systematic review. Psychol Med. 2007 May;37(5):615-26.

Latendresse SJ, Rose RJ, Viken RJ, Pulkkinen L, Kaprio J, Dick DM. Examining the etiology of associations between perceived parenting and adolescents' alcohol use: common genetic and/or environmental liabilities? J Stud Alcohol Drugs. 2010. May;71(3):313-25.

Neale, M. C. (2000) The use of Mx for association and linkage analyses. *GeneScreen*, 1;107-111.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population based linkage analyses. Am J Hum Genet. 2007 Sep;81(3):559-75.

Luyckx K, Tildesley EA, Soenens B, Andrews JA, Hampson SE, Peterson M, Duriez B. Parenting and trajectories of children's maladaptive behaviors: a 12-year prospective community study. J Clin Child Adolesc Psychol. 2011;40(3):468-78.

Enoch MA. The influence of gene-environment interactions on the development of alcoholism and drug dependence. Curr Psychiatry Rep. 2012 Apr;14 (2):150-8.

Swendsen J, Burstein M, Case B, Conway KP, Dierker L, He J, Merikangas KR. Use and Abuse of Alcohol and Illicit Drugs in US Adolescents: Results of the National Comorbidity Survey-Adolescent Supplement. Arch Gen Psychiatry. 2012. Apr; 69 (4): 390-8.

Shanahan, M. J. & Hofer, S. M. (2005) Social context in gene-environment interactions: retrospect and prospect. *J.Gerontol.B Psychol.Sci.Soc.Sci.*, 60(1): 65-76.

Bohnert KM, Anthony JC, Breslau N. Parental monitoring at age 11 and subsequent onset of cannabis use up to age 17: results from a prospective study.
J Stud Alcohol Drugs. 2012 Mar;73(2):173-7.

Knafo A, Schwartz SH. Parenting and adolescents' accuracy in perceiving parental values. Child Dev. 2003 Mar-Apr;74 (2):595-611.

Pettit, GS, Laird RD, Dodge KA, Bates JE, Criss MM. Antecedents and behavior-problem outcomes of parental monitoring and psychological control in early adolescence. Ch Dev, 72(2) 583-598.

Englund MM, Egeland B, Oliva EM, et al. (2008) Childhood and adolescent predictors of heavy drinking and alcohol use disorders in early adulthood: a longitudinal developmental analysis. *Addiction* 103: 23-35.

Sroufe LA, Egeland B, Carlson EA, et al. (2005) *The development of the person: The Minnesota Study of Risk and Adaptation from Birth to Adulthood,* New York: Guilford Press.

Erikson E. (1963) *Childhood and society,* New York: Norton.

Chapter 7

Genetic Influences on Alcohol Consumption Have Diverging Developmental Trajectories

Abstract

Background: Both alcohol-specific genetic factors (Kendler et al. 2003 ; Hicks et al. 2004 ; Macgregor et al. 2009) and non-specific genetic factors related to externalizing behavior influence high alcohol consumption and the risk for developing alcohol use disorders across adolescence into adulthood (Kendler et al. 2003 ; Hicks et al. 2004, 2007). Although there is a substantial literature on genetic influences on externalizing disorders in adolescence (Stallings et al., 2005; Dick et al., 2009; Stephens et al., 2011) and alcohol use disorders in adulthood (Treutlein et al., 2011), little is known about the etiologic role of these two classes of genetic risk on alcohol-related behaviors across development. Recently, Kendler et al. (2010) found that non-specific (general externalizing) genetic factors are important for predicting alcohol use in early and mid-adolescence, but that their influence wanes over time as alcohol-specific genetic factors increase in importance during the transition to adulthood.

Methods: In the present study, we build and expand upon these findings using prospective, longitudinal twin data from the population-based FinnTwin12 study. Our primary goal was to attempt to replicate Kendler et al.'s (2010) findings, examining the impact of alcohol-specific and non-specific (general externalizing) genetic factors on alcohol-related behaviors from early adolescence through early adulthood (ages 12-22). Each twin's genetic risk for alcohol use disorders was indexed by their parents' and co-

twin's alcohol dependence symptom counts. The non-specific genetic risk score for externalizing disorders was a composite measure of parents' and co-twin's self-reported symptom count of Conduct Disorder (CD) and Antisocial Personality Disorder (ASPD), each derived from DSM-IV criteria obtained by the SSAGA.

Results: The regression coefficient for non-specific genetic risk begins quite low at age 12 ($\beta = -0.05$), rising to a peak at age 14 ($\beta = 0.23$), decreasing at age 17 ($\beta = 0.13$), and then falling at age 22 ($\beta = 0.09$). The pattern is somewhat different for alcohol-specific genetic risk, which also starts at a relatively low value at age 12 ($\beta = -0.06$) and then rises slowly from age 12-17 and reaches a peak value at age 22 ($\beta = 0.22$).

Conclusions: In accord with previous findings (Kendler et al., 2010), we found divergent developmental trajectories for specific and non-specific genetic factors on alcohol use. Overall, we found more robust prediction of alcohol outcomes with genetic risk for externalizing behaviors earlier in adolescence (12-14) and a more robust prediction of alcohol outcomes with alcohol-specific genetic risk later in adolescence into young adulthood (17-22). These results suggest that, in early adolescence, genetic influences on alcohol use and problems are largely non-specific and may reflect a more general picture of largely adolescent-limited externalizing behaviors (Moffitt, 1993; Moffitt et al. 2002). However, the alcohol-specific genetic risk factors become more important than non-specific genetic influences in early adulthood (Rose et al., 2003). This shift in genetic influences maps onto the typical developmental timing for the onset of serious alcohol problems (Schuckit et al. 1995).

Introduction

Adolescence is typically the period of the lifespan where alcohol use is initiated and regular patterns of use are established (Swendsen et al., 2012). This period is also characterized by rapid transitions in the degree to which alcohol consumption is attributed to genetic or environmental factors, with environmental factors predominating in early adolescence, and genetic factors increasing in importance over time (Kendler et al., 2008; Viken et al.,1998, Dick et al., 2007). Both alcohol-specific genetic factors (Kendler et al. 2003 ; Hicks et al. 2004 ; Macgregor et al. 2009) and non-specific genetic factors related to externalizing behavior influence high alcohol consumption and the risk for developing alcohol use disorders across adolescence into adulthood (Kendler et al. 2003 ; Hicks et al. 2004, 2007). Although there is a substantial literature on the genetic influences on externalizing disorders in adolescence (Stallings et al., 2005; Dick et al., 2009; Stephens et al., 2011) and alcohol use disorders in adulthood (Treutlein et al., 2011), little is known about the etiologic role of these two classes of genetic risk on alcohol-related behaviors across development.

Kendler and colleagues recently began to address this issue in a male cohort of the Virginia Adult Twin Study of Psychiatric and Substance Use Disorders (Kendler et al., 2010; Kendler & Prescott, 2006). Using retrospective reports of alcohol use across the lifespan, their results indicated that the importance of non-specific genetic factors related to externalizing behavior on maximal alcohol consumption is greatest in early to mid-adolescence, peaking at ages 15–17 years and then declining slowly into adulthood. In contrast, the influence of alcohol-specific genetic factors increases slowly through mid-adulthood.

In the present study, we build and expand upon these findings using prospective, longitudinal twin data from the population-based FinnTwin12 study. Our primary goal was to extend the findings of Kendler et al. (2010), by examining the impact of alcohol-specific and non-specific (general externalizing) genetic factors on alcohol-related behaviors from early adolescence through early adulthood. This study expands on previous work in several ways. First, data for both males and females are available, while the Kendler study (2010) used exclusively males. Second, data on drinking from the VATSPSUD sample were retrospective; in contrast, prospective reports from various stages of development are used in the present study. Finally, although overall rates of drinking frequency and problems are similar in Finland and the United States, drinking culture, and age of legal drinking differ (Helasoja et al. 2004; Bloomfield et al., 2010).

Methods

Sample

FinnTwin12 has been described in previous chapters (3, 4, 5 and 6). Nested in this study lays an intensive assessment of a subsample of 1035 families, comprising about 40% of all twins, mostly selected at random (72.3%, 748 families). A small part of the subsample (27.3%, 287 families) is enriched with families with twins assumed to be at elevated familial risk for alcoholism risk. Details about the sub-sample have been described earlier (Rose et al., 2001). In this subsample, both twins and parents were interviewed using the SSAGA (Semi-Structured Assessment for the Genetics of Alcoholism (Buzholz et al., 1994). The interviews were highly age-standardized; the mean age at interviews was 14.19 years, with 75% of interviews completed by 14 years and 3 months of age and all interviews completed before the age of 15. The final sample consisted of 1,854 interviewed boys (N = 945, 51%) and girls (N = 909, 49%). Due to the longitudinal study design, some variables were available on fewer individuals (exact frequencies for each measure described below). Zygosity was determined using a well-validated questionnaire completed by both co-twins at the baseline (Kaprio, Pulkkinen, & Rose 2002). This was supplemented by parental information and comparisons of school photographs for the 3% of twins whose zygosity could not be determined definitively from information in the questionnaires (Kaprio et al., 2002; Kaprio et al., 2006b).

Assessment

Calculation of Genetic Risk Scores

Each twin had his/her genetic risk for alcohol use disorders indexed by their parents' and co-twin's alcohol dependence symptom counts. Alcohol dependence symptom counts were derived from the Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA; Bucholz et al., 1994) interviews based on the criteria outlined in the Diagnostic and Statistical Manual for Psychiatric Disorders (DSM-IV; American Psychiatric Association, 1994). For each of the twins, the SSAGA assessments were administered when the twins were age 22. For the parents, all SSAGA data were collected when the twins were age 14. The DSM-IV criteria for alcohol dependence consist of seven criteria that include both physiological and psychological symptoms associated with problematic alcohol use. The contribution of each measure (parents' symptom sum score, and co-twin's symptom sum score) to the total alcohol use disorder risk was based on a modified ridit score (Kendler et al, 2010). When data on both parents were available, symptom counts from the most severely affected parent was used in risk score calculation. The correlation between parents' SSAGA symptom counts was 0.32, $p<0.0001$. Scores from monozygotic (MZ) co-twins were weighted twice as strongly as scores from dizygotic (DZ) co-twins or parents. Alcohol-specific genetic risk scores (AD-GR) were computed on 1,854 twins.

The non-specific genetic risk score for externalizing disorders was a composite measure of the parents' and co-twin's self-reported symptom count of Conduct Disorder (CD) and Antisocial Personality Disorder (ASPD), each derived from DSM-IV criteria obtained by the SSAGA. According to the DSM-IV classification system (American Psychological Association, 1994), CD is a repetitive and persistent pattern of behavior in which the basic rights of others or major age-appropriate societal norms or rules are violated. ASPD is an Axis II personality disorder characterized by a pervasive pattern of disregard for, and violation of, the rights of others that begins in childhood or early adolescence and continues into adulthood. Note that an adolescent CD diagnosis is an adult ASPD criteria. Throughout this manuscript, we will describe the adolescent criteria as CD and the adult criteria as antisocial behavior (ASB). Non-specific genetic risk scores related to externalizing disorders (EXT-GR) were computed on 2,029 twins.

Drinking Frequency Measures

Current alcohol consumption was assessed at each of the four time points. At age 12, subjects were asked if they had ever used alcohol when they were not in the presence of an adult (n=2,826 twins). Adolescent alcohol use was assessed at age 14 and 17 by asking the participants to report how frequently they drink alcohol. On the age 14 questionnaire, the item included four response options: (1) Never, I don't drink alcohol; (2) Less often than once a month; (3) About 1 to 2 times a month; and (4) Once a week or more. At age 14, a total of 2,828 twins responded to the item. On the age 17 questionnaire, the item included nine response options: (1) Daily; (2) A couple of times a week; (3) Once a week; (4) A couple of times a month; (5) About once a month; (6) About once every two months; (7) 2-4 times per year; (8); Once a year or less; (9) I don't drink any alcohol. The latter response options were collapsed into four categories to parallel the age 14 data; (1) Never, (2) Yearly, (3) Monthly, and (4) Weekly. At age 17, a total of 2,366 twins responded to the item. At age 22, subjects (n = 2,158) were asked *how many weeks in the last 6 months did you drink alcohol?*.

Statistical Analysis

The original distribution of the alcohol use data was highly skewed, and preliminary analyses indicated that a log transformation was optimal at stabilizing the variance. The residual correlation within twin pairs was substantial and stronger in MZ twin pairs; accordingly, regression models were run as hierarchical linear models using PROC

MIXED and PROC GENMOD in SAS (SAS Institute, 2008), with twin pairs and individuals within twin pairs being treated as separate levels.

Results

Descriptive Statistics

At age 12, 93.2% of the sample responded that they had not ever used alcohol outside the presence of an adult. At age 14, 64.9% of the sample reported that they had never used alcohol, 20.4% reported drinking less often than once a month, 12.1% reported using alcohol about 1 or 2 times per month, and 2.6% reported using alcohol once per week or more. At age 17, 11.9% of the sample reported that they had never used alcohol, 22.3% reported drinking less often than once a month, 41.5% reported using alcohol about 1 or 2 times per month, and 24.3% reported using alcohol once per week or more. At age 22, the subjects reported drinking alcohol an average of 13.83 (SD=8.2) weeks in the last 6 months (range 0-26 weeks).

AD-GR scores were based on parent and co-twin DSM-IV AD symptoms. Consistent with expectations for a population-based sample, the parents of the twins largely fell within sub-threshold ranges of alcohol dependence (AD) symptom counts (range=0-7, M=1.03, SD=1.68), with 6.2% of the parents meeting criteria for an AD diagnosis (3 or more AD criteria endorsed). The twins' AD symptom scores ranged from 0-7 (M=1.09, SD=1.37), with 13.4% of the sample meeting criteria for DSM-IV AD. AD-GR scores ranged from 0-8 (M=1.16, SD=1.41). EXT-GR scores were based on parent and co-twin DSM-IV CD and ASB. The majority of the parents were within sub-threshold ranges of CD symptoms (range=0-7, M=0.65, SD=0.99), with 1.7% of the parents meeting criteria for a CD diagnosis. Twins' CD symptom sum scores ranged from 0-8 (M=1.08, SD= 1.26), with 12.3% meeting CD diagnosis criteria. Most of the parents of twins were within normative sub-threshold ranges of ASB (Range=0-6, M=1.12, SD=1.18), with <1% of the parents meeting criteria for the adult portion of the ASPD diagnosis. Twins' ASB sum scores ranged from 0-6 (M= 0.63, SD= 0.96), with 2.6% meeting the adult ASPD diagnosis criteria. EXT-GR scores ranged from 0-7 (M=0.87, SD=1.27).

Zero-order correlations among focal variables are shown in Table 27. Age 12 drinking initiation significantly predicted drinking frequency at ages 14 and 17, but not at age 22. The relation was strongest between age 12 drinking initiation and age 14 frequency of drinking. Age 14, 17, and 22 drinking frequency were all significantly associated, with the stronger relationships existing between age 14 and 17 drinking frequency and between age 17 and age 22 drinking frequency. AD-GR and EXT-GR were correlated at 0.38.

Table 27. Correlations Between Twins' Alcohol Consumption and Problem Outcomes Across Development

Pearson Correlations Alcohol Consumption	Measure of Alcohol Consumption			
	Age 12	Age 14	Age 17	Age 22
Age 12	1.000	-0.213**	-0.142**	-0.004
Age 14	-0.213**	1.000	0.316**	0.142**
Age 17	-0.142**	0.316**	1.000	0.300**
Age 22	-0.004	0.142**	0.300**	1.000

**Correlation is significant at p<0.01.

Diverging developmental trajectories of alcohol-specific and non-specific genetic risk factors

The relationship between AD-GR, EXT-GR, and the alcohol outcomes over development are depicted in Figure 15a and detailed in Table 28. The regression coefficient for EXT-GR begins quite low at age 12 ($\beta = -0.05$), rising to a peak at age 14 ($\beta = 0.23$), decreasing at age 17 ($\beta = 0.13$), and then falling at age 22 ($\beta = 0.09$). In contrast, AD-GR starts at a relatively low value at age 12 ($\beta = -0.06$) and then rises slowly from age 12-17 and reaches a peak value at age 22 ($\beta = 0.22$).

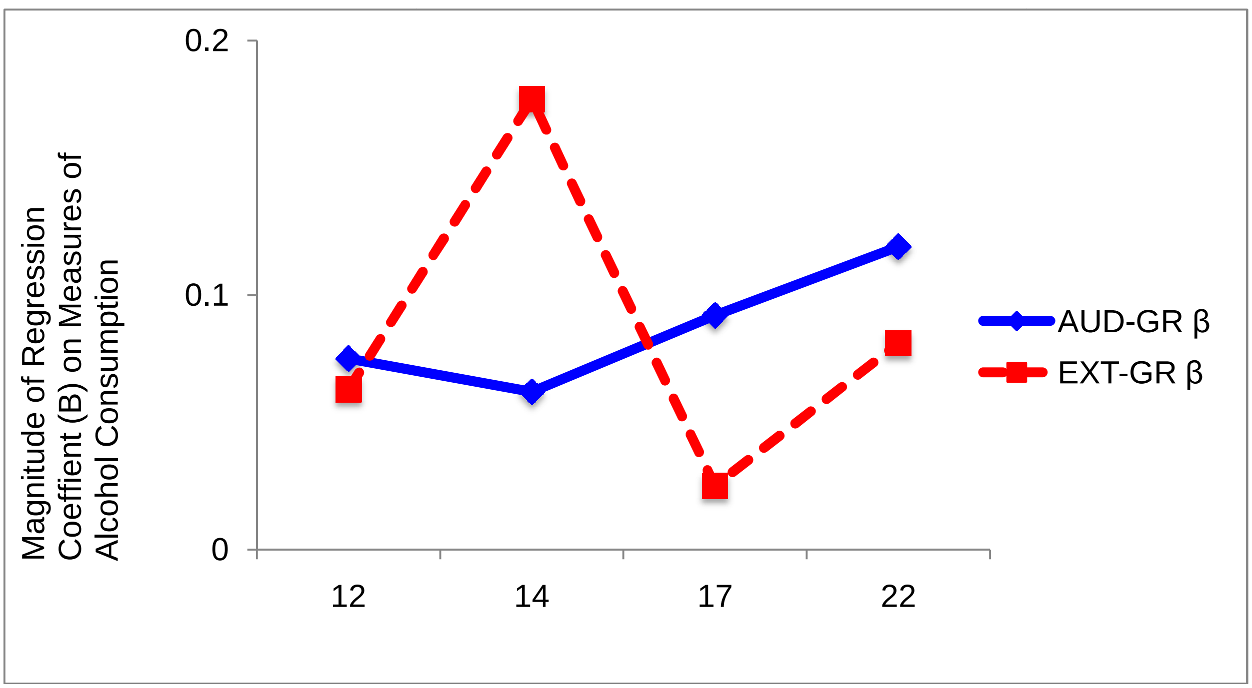
Table 28. Genetic Risk Scores Predicting the Twins' Alcohol Consumption Across Development

	AD-GR		EXT-GR	
		Twin 1		*Twin 1*
Alcohol Consumption	*β*	*p-value*	*β*	*p-value*
Age 12	.064	0.052	.052	0.112
Age 14	.141	0.00004*	.232	<0.0000001*
Age 17	.196	0.0000003*	.132	0.001*
Age 22	.179	0.000003*	.085	0.029*

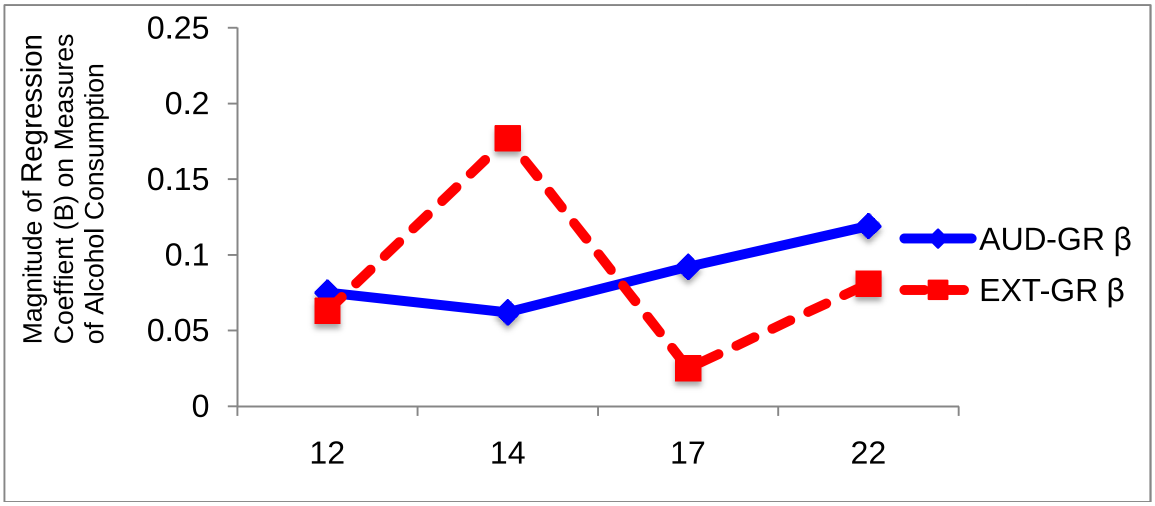
In further examination of the relationship between AD-GR, EXT-GR, and the alcohol outcomes over development, we performed secondary analyses separately by sex. In males, the regression coefficient for EXT-GR begins low at age 12 ($\beta = 0.06$), rising to a peak at age 14 ($\beta = 0.17$), decreasing at age 17 ($\beta = 0.03$), and then rising slightly at age 22 ($\beta = 0.08$). In contrast, AD-GR starts at a moderate value at age 12 ($\beta = 0.08$) and then falls slightly at age 14 ($\beta = 0.06$), rising again at age 17($\beta = 0.92$) and reaches a peak value at age 22 ($\beta = 0.12$). The relationship between AD-GR, EXT-GR, and the alcohol outcomes over development for males is depicted in Figure 15b. In females, the regression coefficient for EXT-GR begins relatively higher at age 12 (β = 0.18), rising to a peak at age 14 (β = 0.28), and then slowly decreasing from age 17 (β = 0.21) through age 22 (β = 0.06). In contrast, AD-GR starts at a relatively high value at age 12 (β =0.12) that continues to increase at age 14 (β = 0.21), and reaches its peak at age 17(β = 0.28) and decreases slightly at age 22 (β = 0.190). The relationship between AD-GR, EXT-GR, and the alcohol outcomes over development for females is depicted in Figure 15c.

Figure 15. Developmental Trajectories of Two Classes of Genetic Risk for Alcohol Consumption

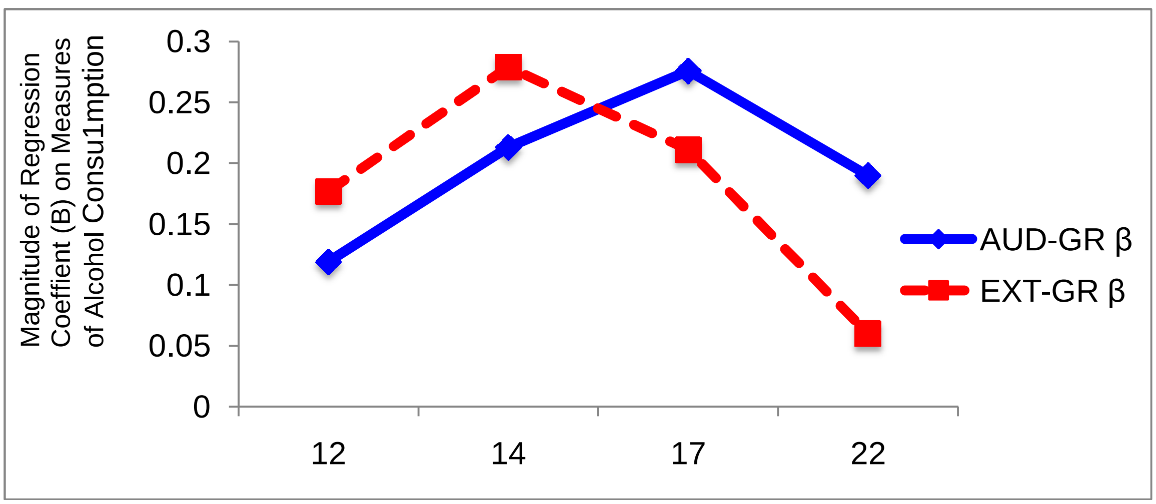
a) Sexes Collapsed



b) Males



c) Females



Discussion

Epidemiological studies have demonstrated that the genetic influences on alcohol-related outcomes have both an alcohol-specific component and a general externalizing component (Kendler et al., 2001; Kendler et al., 2003). Until recently (Kendler et al., 2010), the relative importance of these sets of genetic influences across time remained unexamined. In consideration of twin study findings which indicate that the importance of genetic influences on alcohol use change across adolescence, we sought to examine the relative contribution of each of these aspects of the genetic influence, both alcohol specific influences and general externalizing influences, on alcohol use across adolescence into young adulthood.

Our study used a population based, longitudinal sample of Finnish twins to follow up on findings from a recent study (Kendler et al. 2010) that found that specific and non-specific genetic influences on alcohol consumption have different development trajectories. Supporting evidence from epidemiological twin studies, which suggested that alcohol use and problems are influenced by both alcohol specific genetic risk factors and externalizing genetic risk factors, we found that both alcohol-specific genetic risk and general externalizing genetic risk predict alcohol outcomes from early adolescence to young adulthood. Furthermore, and in accord with previous findings (Kendler et al., 2010), we also found divergent developmental trajectories for specific and non-specific genetic factors on alcohol use. Overall, we found more robust prediction of alcohol outcomes with genetic risk for externalizing behaviors earlier in adolescence (12-14) and a more robust prediction of alcohol outcomes with alcohol-specific genetic risk later in adolescence into young adulthood (17-22). These results suggest that, in early adolescence, genetic influences on alcohol use and problems are largely non-specific and may reflect a more general picture of largely adolescent-limited externalizing behaviors (Moffitt, 1993; Moffitt et al. 2002). However, the alcohol-specific genetic risk factors become more important than non-specific genetic influences in early adulthood (Rose et al., 2003). This shift in genetic influences maps onto the typical developmental timing for the onset of serious alcohol problems (Schuckit et al. 1995).

In further examination of the relationship between AD-GR, EXT-GR, and the alcohol outcomes over development, we performed analyses separately by sex. Overall, the relative influence of AUD-GR and EXT-GR on alcohol consumption across development was maintained; in early adolescence, genetic influences on alcohol use and problems are largely non-specific and later in adolescence and young adulthood, alcohol specific genetic influences on alcohol use are more influential. However, several interesting sex differences in the trajectories of these influences emerged. Most striking is the relatively early influence of AUD-GR on alcohol consumption in females. Twin studies have indicated that drinking frequency is heritable in girls at a younger age than boys (Rose et al., 2001; Maes et al., 1999). The authors pointed to increased alcohol use, pubertal timing, and having a greater number of older friends (that are presumably providing drinking opportunities) as an explanation for these findings. Perhaps this earlier access to alcohol and earlier evidence of heritability in drinking frequency is related to the earlier influence of alcohol specific genetic risk for consumption in early adolescence. Also of note is the relative influence of EXT-GR in late adolescence and early adulthood. In females, risk for alcohol consumption at age 22 is largely influenced by AUD-GR, with EXT-GR playing a very small role. In males, both AUD-GR and EXT-GR appear to substantively influence age 22 alcohol consumption. Past studies have reported gender differences in alcoholic subtypes, including an excess of women in internalizing subtypes and an excess of men in externalizing subtypes (Epstein et al., 2002; Moss et al., 2007, Carpenter and Hasin, 2001 and Pombo and Lesch, 2009). Findings from the present study support these sex differences. These differences correspond to gender differences in the prevalence of internalizing and externalizing disorders in the total population (Grant et al., 2004a, Grant et al., 2004b and Stinson et al., 2005). The few studies that have examined gender differences in the comorbidity of alcohol dependence have reported disparate findings (Kessler et al., 1997; Alonso et al., 2004, Kramer et al., 2008).

Another notable difference between our findings and those of Kendler et al. (2010) is the age at which alcohol specific genetic risk factors and externalizing genetic risk factors shift in their relative importance. The most dramatic shift in genetic influence on drinking frequency occurred around age 21 in Kendler's Virginia Adult Twin Study of Psychiatric and Substance Use Disorders. However, this shift occurred around age 17 in the Finnish data. This may attributable to several factors, including the ages at which the alcohol assessments were made in each of the samples (the males in Kendler's study were making retrospective reports of their drinking at a mean age of 40.3 years [SD=9.0], whereas our reports were made prospectively), Although both studies measure alcohol use across development, several studies suggest that there are important recall biases in self-reports of past drinking behavior (Labouvie et al., 1997, Engels et al., 1997; Prause et al., 2007). Lastly, there are both differences in the legal drinking age and cultural norms regarding alcohol use in Finland and the United States (Helasoja et al. 2004; Bloomfield et al, 2010).

These results should be interpreted in the context of several important limitations. First, we used hierarchical linear modeling rather than structural equations modeling in our analyses. We used this method because it allowed us to easily incorporate and interpret data on parental psychopathology in our measures of genetic risk. However, this method lacks the precision to distinguish genetic from familial environment effects. Second, although the present study uses developmentally-appropriate drinking measures across time, there were differences in both how the question was posed to the subject as well as response options available, which introduced measurement variance. We addressed this in our analyses by examining the pattern of cross-sectional effects over time, rather than fitting longitudinal growth models.

In summary, the present study replicates and extends past findings showing that two classes of genetic risk related to alcohol use changes across time. Similar to past work (Kendler et al., 2010), our findings indicate that alcohol-specific genetic risk factors increase in importance across adolescence and early adulthood; in contrast, non-specific genetic influences decrease in importance across this same period. Taken altogether, these findings highlight the importance of taking a developmental perspective on the role of genetic influences on alcohol use during adolescence and young adulthood.

References

Swendsen J, Burstein M, Case B, Conway KP, Dierker L, He J, Merikangas KR. Use and Abuse of Alcohol and Illicit Drugs in US Adolescents: Results of the National Comorbidity Survey-Adolescent Supplement. Arch Gen Psychiatry. 2012 Apr;69(4):390-8.

Kendler KS, Aggen SH, Czajkowski N, Røysamb E, Tambs K, Torgersen S, Neale MC, Reichborn-Kjennerud T. The structure of genetic and environmental risk factors for DSM-IV personality disorders: a multivariate twin study. Arch Gen Psychiatry. 2008 Dec;65(12):1438-46.

Kendler KS, Prescott CA, Myers J, Neale MC. The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. Arch Gen Psychiatry. 2003 Sep;60(9):929-37.

Hicks BM, Krueger RF, Iacono WG, McGue M, Patrick CJ. Family transmission and heritability of externalizing disorders: a twin-family study. Arch Gen Psychiatry. 2004 Sep;61(9):922-8.

Macgregor S, Lind PA, Bucholz KK, Hansell NK, Madden PA, Richter MM, Montgomery GW, Martin NG, Heath AC, Whitfield JB. Associations of ADH and ALDH2 gene variation with self report alcohol reactions, consumption and dependence: an integrated analysis. Hum Mol Genet. 2009 Feb 1;18(3):580-93.

Hicks BM, Blonigen DM, Kramer MD, Krueger RF, Patrick CJ, Iacono WG, McGue M. Gender differences and developmental change in externalizing disorders from late adolescence to early adulthood: A longitudinal twin study. J Abnorm Psychol. 2007 Aug;116(3):433-47.

Stallings MC, Corley RP, Dennehey B, Hewitt JK, Krauter KS, Lessem JM, Mikulich-Gilbertson SK, Rhee SH, Smolen A, Young SE, Crowley TJ. A genome-wide search for quantitative trait Loci that influence antisocial drug dependence in adolescence. Arch Gen Psychiatry. 2005 Sep;62(9):1042-51.

Dick DM, Bernard M, Aliev F, Viken R, Pulkkinen L, Kaprio J, Rose RJ. The role of socioregional factors in moderating genetic influences on early adolescent behavior problems and alcohol use. Alcohol Clin Exp Res. 2009 Oct;33(10):1739-48.

Stephens SH, Hoft NR, Schlaepfer IR, Young SE, Corley RC, McQueen MB, Hopfer C, Crowley T, Stallings M, Hewitt J, Ehringer MA. Externalizing Behaviors are Associated with SNPs in the CHRNA5/CHRNA3/CHRNB4 Gene Cluster. Behav Genet. 2011 Nov 1.

Treutlein J, Rietschel M. Genome-wide association studies of alcohol dependence and substance use disorders. Curr Psychiatry Rep. 2011 Apr;13(2):147-55. Review.

Kendler KS, Gardner C, Dick DM. Predicting alcohol consumption in adolescence from alcohol-specific and general externalizing genetic risk factors, key environmental exposures and their interaction. Psychol Med. 2011 Jul;41(7):1507-16.

Kendler & Prescott, 2006
Prescott CA, Sullivan PF, Kuo PH, Webb BT, Vittum J, Patterson DG, Thiselton DL, Myers JM, Devitt M, Halberstadt LJ, Robinson VP, Neale MC, van den Oord EJ, Walsh D, Riley BP, Kendler KS. 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. Mol Psychiatry. 2006 Jun;11(6):603-11.

Kendler KS. Family history information in biomedical research. J Contin Educ Health Prof. 2001 Fall;21(4):215-23. Review.

Shanahan MJ, Hofer SM. Social context in gene-environment interactions: retrospect and prospect. J Gerontol B Psychol Sci Soc Sci. 2005 Mar;60 Spec No 1:65-76. Review.

Meyers et al., 2011

Helasoja V, Lahelma E, Prättälä R, Klumbiene J, Pudule I, Tekkel M. Trends in the magnitude of educational inequalities in health in Estonia, Latvia, Lithuania and Finland during 1994-2004. Public Health. 2006 Sep;120(9):841-53. Epub 2006 Aug 1.

Bloomfield et al., 2010

Kaprio J, Pulkkinen L, Rose RJ. Genetic and environmental factors in health-related behaviors: studies on Finnish twins and twin families. Twin Res.

2002 Oct;5(5):366-71. Review.

Kaprio J, Koskenvuo M. Genetic and environmental factors in complex diseases: the older Finnish Twin Cohort. Twin Res. 2002 Oct;5(5):358-65. Review.

Kaprio J. Twin studies in Finland 2006. Twin Res Hum Genet. 2006 Dec;9(6):772-7.

Bucholz KK, Cadoret R, Cloninger CR, Dinwiddie SH, Hesselbrock VM, Nurnberger JI Jr, Reich T, Schmidt I, Schuckit MA. A new, semi-structured psychiatric interview for use in genetic linkage studies: a report on the reliability of the SSAGA. J Stud Alcohol. 1994 Mar;55(2):149-58.

Association, A. P. 1994, *Diagnostic and statustical manual of mental disorders*, Fourth Edition edn, American Psychiatric Association, Washington, DC.

SAS Institute, 2008

Slutske WS, Heath AC, Madden PA, Bucholz KK, Statham DJ, Martin NG. Personality and the genetic risk for alcohol dependence. J Abnorm Psychol. 2002 Feb;111(1):124-33.

Rose RJ, Viken RJ, Dick DM, Bates JE, Pulkkinen L, Kaprio J. It does take a village: nonfamilial environments and children's behavior. Psychol Sci. 2003 May;14(3):273-7.

Schuckit MA, Anthenelli RM, Bucholz KK, Hesselbrock VM, Tipp J. The time course of development of alcohol-related problems in men and women. J Stud Alcohol. 1995 Mar;56(2):218-25.

GLOBAL CONCLUSIONS

Genetic studies of alcohol phenotypes provide an excellent example of the challenges posed by the search for risk genes for complex behavioral and psychiatric disorders. Decades of twin and family studies have demonstrated that there are critical genetic and environmental components in the inheritance of substance use disorders. We now know that there are a multitude of genes, each with subtle effects influencing an individual's risk for the development of alcohol use problems that likely interact epistatically as well as with their environments (biological and external) to make an individual more susceptible to the development of these complex disorders. Also, as our understanding of substance use becomes more refined, we see that substance dependence has a complex development that starts with initiation of use, or in some respects earlier with impulsive behavior observed in adolescence (e.g. externalizing problems, conduct disorder) and continues through the individual's drinking career.

To date, researchers have had limited success in identifying all genetic variance in complex human traits ("missing heritability"). To this end, many gene-finding methodologies have been employed over the past few decades including linkage and association. Linkage, candidate gene and genome wide association techniques have provided few genetic risk variants that are consistently and robustly associated with alcohol dependence. While there is no gold standard method that has successfully led to the identification of all genetic variance in complex traits, promising new methods are currently being developed. While the task of developing and trouble shooting novel gene-finding methods for complex traits is wrought with peril, it also provides an exciting challenge for the future of the field. The present dissertation study attempts to add to this trial-and-error process by testing variations on current gene-finding methodology. There are several subtle conclusions to draw from this series of analyses; some that may inform methodology and others that speak to the specific risk for alcohol use and problems.

First, I believe that we can use information gained from other fields and methods, such as behavioral genetics, developmental psychology and epidemiology to inform genetic association studies for complex behavioral traits. While the twin and family literature currently exists somewhat separately from the gene-identification literature, I believe that this gap can be narrowed if new methodology is developed to combine the strengths of these two methods. Hopefully, the analyses presented in this dissertation have demonstrated novel ways in which these two methods can inform each other both indirectly, by testing the same research question using two different methods, and directly, by using genetic factor loadings from twin analyses as the outcome in genetic association studies. A second overall conclusion that can be made from this series of analyses is that different aspects of alcohol use appear to be mediated by different genetic risk variants. We have demonstrated this at the latent genetic level in twin studies as well as with molecular genetic data in GWAS. As scientists, we tend to compartmentalize and potentially over simplify complex concepts in an effort to make them measurable. This has been very useful in the context of understanding and recognizing patterns in human behavior. However, it is likely that our biology does not respect these categories and distinctions. Further, it appears likely that several aspects (and measures) of the “same” behavior or disorder are not necessarily equal. Moreover, alcohol dependence symptoms are both phenotypically and genetically heterogenous. There are many different routes to a disease like alcoholism. The likely possibility that for every developmental trajectory that leads to alcoholism, there may be an equivalent “biological-course,” indicates a degree of heterogeneity that is rarely modeled/tested. Using biologically informed alcohol phenotypes (eg. genetic factor scores) may improve the ability to detect genetic association by reducing some of this heterogeneity. Perhaps the most important conclusion to draw from this dissertation study is that certain environments moderate genetic influences on alcohol use and/or dependence. Environments have the capacity to both mask and exacerbate genetic influences. This is of immense importance to a disease like alcoholism, which specifically requires an individual to initiate drinking behavior. While methodology and statistical considerations required to properly test this have not yet been fully developed, excluding gene-environment interactions from our models may pose serious challenges to truly characterizing risk for alcohol use phenotypes.

In summary, the field of psychiatric genetics is trouble-shooting effective methodologies for the identification of genetic risk variants that predispose individuals to the development of complex behavioral disorders. Several challenges related to the complex and polygenic nature of these phenotypes, must be considered. This dissertation study sought to address these important challenges in the context of alcohol use disorders and related phenotypes. In this dissertation several studies were described that integrated twin methodologies into gene identification studies in an effort to 1) reduce heterogeneity (both phenotypic and genotypic), 2) elucidate environmental constructs that moderate genetic influences, and 3) enhance our ability to detect the subtle genetic influences on alcohol use and problems.

This dissertation has offered the field a novel approach to characterizing genetic and environmental risk that integrates quantitative and molecular genetic methodologies through a variety of analyses conducted in the longitudinal Finnish Twin Studies. This study has integrated twin methodology and genome wide association to offer a method that directly utilizes information gained at the latent genetic and environmental level in genetic association studies. This latent information includes genetic factors derived from twin models, which can be harnessed in genetic association studies and has the capability of reducing the heterogeneity present in measures of alcohol consumption and problems. In addition, these analyses suggest a new way to move the study of gene environment interaction forward in testing for moderation at the level of aggregate molecular genetic risk. In doing so, we examine the interaction between aggregate molecular genetic risk and the environment that allows us to parallel the established latent gene-environment interaction effects reported from twin studies. This method also allows us to begin to more systematically characterize the specific environments that are critical for moderating the importance of a genetic predisposition, and the ages and developmental stages at which these gene environment interactions operate. This will advance our understanding of how genetic risk unfolds across time, and how to reduce risk among individuals carrying genetic predispositions associated with substance use outcomes, which could be useful for prevention and intervention efforts.

There are several future directions that each of these studies could improve each of these studies and their related areas of research. First, further characterization of the latent genetic factors derived from measures of alcohol consumption and problems should be carried out. To this end, the relationship between genetic factors and externally validating variables (related behaviors and disorders) should be explored in order to more precisely interpret the role each of the genetic variants implicated by the genetic association analyses has in the risk for alcohol use and/or problems. The gene-environment interaction analyses carried out in this dissertation relied heavily on self-reported environmental constructs, measured only once in adolescence. Ideally, future work would include more carefully considered environmental constructs measured at the relevant stage of development for the outcome. While *Finntwin12* is a rich longitudinal twin sample, its utility in identifying individual genetic risk variants of small effect is limited by the number of individuals in which molecular genetic data is available on. Future directions should involve the inclusion of all twins in order to increase the potential to detect genetic risk variants for alcohol use phenotypes. Most importantly, all results including latent twin models (chapters 1 & 4), specific associated genetic variants (chapters 2 & 3), aggregate genetic risk scores, and environmental moderation effects (chapters 5 & 6), presented in this dissertation should be replicated in an independent sample with comparable measures and ages of assessment.

VITA

Jacquelyn Leigh Meyers was born on October 30, 1986, in Oceanside County, New York, and is an American citizen. She graduated from North Broward Preparatory School, Coconut Creek, Florida in 2004. She received her Bachelors of Science in Psychology, with a minor in English Literature from Florida State University, Tallahassee, Florida in 2007. From 2004 - 2007, she was employed at the Florida State University Psychology Clinic.