A *GABRA2* Polymorphism Improves a Model for Prediction of Drinking Initiation

Samuel Kuperman, M.D. [a], Grace Chan, Ph.D. [b], John Kramer, Ph.D. [a], Leah Wetherill, M.S. [c], Laura Acion, Ph.D. [a], Howard J. Edenberg, Ph.D. [d], Tatiana M. Foroud, Ph.D. [c], John Nurnberger Jr., M.D., Ph.D. [e], Arpana Agrawal, Ph.D. [f], Andrey Anokhin, Ph.D. [f], Andrew Brooks, Ph.D. [g], Victor Hesselbrock, Ph.D. [b], Michie Hesselbrock, Ph.D. [b], Marc Schuckit, M.D. [h], Jay Tischfield, Ph.D. [g], Xiangtao Liu, Ph.D. [a]

[a] Department of Psychiatry, University of Iowa Carver College of Medicine, Iowa City, Iowa

[b] Department of Psychiatry, University of Connecticut Health Center, Farmington, Connecticut

[c] Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, Indiana

[d] Department of Biochemistry & Molecular Biology, Indiana University School of Medicine, Indianapolis, Indiana

[e] Department of Psychiatry, Indiana University School of Medicine, Indianapolis, Indiana

[f] Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri

[g] Department of Genetics, Rutgers University, Piscataway, New Jersey

[h] Department of Psychiatry, University of California San Diego School of Medicine, La Jolla, California

Address correspondence to:
Samuel Kuperman, M.D.
Department of Psychiatry
University of Iowa, Carver College of Medicine
200 Hawkins Drive RM #2701-C JPP
Iowa City, IA 52242
Email: samuel-kuperman@uiowa.edu

Abstract

Background: Survival analysis was used to explore the addition of a single nucleotide polymorphism (SNP) and covariates (sex, interview age, and ancestry) on a previously published model's ability to predict onset of drinking. A SNP variant of rs279871, in the chromosome 4 gene encoding gamma-aminobutyric acid receptor (*GABRA2)*, was selected due to its associations with alcoholism in young adults and with behaviors that increased risk for early drinking.

Methods: A subsample of 674 adolescents (ages 14–17) participating in the Collaborative Study on the Genetics of Alcoholism (COGA) was examined using a previously derived Cox proportional hazards model containing: 1) number of non-drinking related conduct disorder (CD) symptoms, 2) membership in a high-risk alcohol-dependent (AD) family, 3) most best friends drank (MBFD), 4) Achenbach Youth Self Report (YSR) externalizing score, and 5) YSR social problems score. The above covariates along with the SNP variant of *GABRA2,* rs279871, were added to this model. Five new prototype models were examined. The most parsimonious model was chosen based on likelihood ratio tests and model fit statistics.

Results: The final model contained four of the five original predictors (YSR social problems score was no longer significant and hence dropped from subsequent models), the three covariates, and a recessive *GABRA2* rs279871 TT genotype (two copies of the high-risk allele containing thymine). The model indicated that adolescents with the high-risk TT genotype were more likely to begin drinking than those without this genotype.

Conclusions: The joint effect of the gene (rs279871 TT genotype) and environment (MBFD) on adolescent alcohol initiation is additive, but not interactive, after controlling for behavior problems (CD and YSR externalizing score). This suggests that the impact of the high-

risk TT genotype on the onset of drinking is affected by controlling for peer drinking and does

not include genotype-by-environment interactions.

Highlights

- A rs279871 SNP is associated with increased risk for problematic alcohol use.

- This SNP is associated with behavior that is related to early alcohol initiation.

- A TT genotype of this SNP improved a previous model for alcohol initiation.

- Adolescents with the TT genotype were more likely to begin drinking earlier.

- The impact of the TT genotype is affected by controlling for peer drinking.

Keywords: alcohol; drinking initiation; *GABRA2*; rs279871; survival analysis modeling; adolescent

Introduction

Alcohol use is common among adolescents (Johnston, O'Malley, Miech, Bachman, & Schulenberg, 2016) and can result in injuries, death, suicidal behavior, aggression, unprotected sex, academic failure, and social problems (Brown et al., 2008). Additional concern arises from the age at which a youth first drinks, because drinking initiation before age 15 is associated with a 4-fold increase in the likelihood of a lifetime DSM-IV alcohol dependence (AD) diagnosis (American Psychiatric Association, 1994), compared to an individual who delays drinking initiation until late adolescence or early adulthood (B. F. Grant & Dawson, 1997). This suggests that identification of predictors for adolescent drinking could aid development of prevention programs for adolescent drinking, influencing both immediate and longer-term consequences of alcohol use.

The literature delineates a number of predictors, categorical and quantifiable, linked to early drinking initiation. Such predictors include: male sex (Disney, Elkins, McGue, & Iacono, 1999; B. F. Grant, 1998); childhood psychopathology (Clark, Parker, & Lynch, 1999; Kuperman et al., 2005); poor family supervision and inconsistent/harsh discipline (Griffin, Botvin, Epstein, Doyle, & Diaz, 2000; Kuperman et al., 2001); parental separation (J. D. Grant et al., 2015; Waldron et al., 2014); positive peer attitudes toward drinking (Bekman, Cummins, & Brown, 2010; Capaldi, Stoolmiller, Kim, & Yoerger, 2009; Griffin et al., 2000; McCuller, Sussman, Dent, & Teran, 2001; Trucco, Colder, Bowker, & Wieczorek, 2011); peer use of alcohol or other substances (Bekman et al., 2010; Capaldi et al., 2009; Griffin et al., 2000; Trucco et al., 2011); parental alcohol dependence (AD) or antisocial personality disorder (Assanangkornchai, Geater, Saunders, & McNeil, 2002; Kuperman, Schlosser, Lidral, & Reich, 1999; Legrand, McGue, & Iacono, 1999); relationships with antisocial peers (Zucker, Donovan, Masten, Mattson, & Moss, 2008); and prior smoking (Chen et al., 2002). From this existing list of predictors, Kuperman

et al. (2005) identified 63 contained within the adolescent version of the Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA) interview. These variables were used in a series of multiple regression models to identify those significantly related to the self-reported age of drinking initiation in a subsample of 440 adolescent drinkers from the Collaborative Study on the Genetics of Alcoholism (COGA). From this pool, 3 variables – age at interview, number of DSM-IV non-drinking related conduct disorder symptoms (referred to as CD from this point), and the number of adult alcohol-dependent (AD) siblings – formed the most parsimonious model and explained 45% of the variance of age of first drink. Kuperman et al. (2013) subsequently employed these variables along with two additional SSAGA variables (whether most of a subject's best friends drank, and if the subject smoked a cigarette before initiating drinking) and eight scale scores from the commonly used Achenbach Youth Self Report (Achenbach & Rescorla, 2001) to predict the onset of alcohol initiation in an independent sample of 820 adolescent COGA subjects (~60% had not initiated drinking) utilizing the Cox proportional hazards models. Four of these Achenbach Youth Self Report (YSR) scales were hypothesized to be protective and decrease the risk for drinking initiation (positive qualities, activities competence, social competence, and school competence), and four were predicted to be harmful and increase the risk for alcohol initiation in this young cohort (externalizing, attentional, social problems, and internalizing scores). The most parsimonious model included the following: 1) most best friends drank (MBFD), 2) high-risk AD family membership, 3) number of CD symptoms, 4) YSR externalizing score (a non-diagnostic 32-item scale of a wide array of problematic behavior defined as being aggressive, hostile, destructive, defiant in nature, and at odds with accepted societal norms), and 5) YSR social problems score (a non-diagnostic 11-item scale that assesses difficulties with friends and peers).

The current study's goal was to explore whether the inclusion of genetic information would further improve this model's ability to predict alcohol initiation. Variation in the gene encoding gamma-aminobutyric acid receptor (*GABRA2*) on chromosome 4 has been one of the most replicated predictors of alcohol-related phenotypes (Enoch, 2013). *GABRA2* studies have been conducted on both populations outside of (Enoch, 2008; Fehr et al., 2006; Philibert et al., 2009; Soyka et al., 2008) as well as within COGA (Covault, Gelernter, Hesselbrock, Nellissery, & Kranzler, 2004; Dick, Agrawal, et al., 2006; Dick, Bierut, et al., 2006; Edenberg et al., 2004). Previous COGA studies, using a different study population than that of the current investigation, found that the adenine (A) nucleotide for SNP rs279871 was associated with an increased risk for adult AD (Edenberg et al., 2004), with the greatest risk occurring in individuals who were homozygous for this allele (Dick, Agrawal, et al., 2006). These data suggested that the A allele (the T allele on the complementary strand examined in the current study) could potentially be a risk allele for early initiation of alcohol. The SNP rs279871 is in high linkage disequilibrium with other SNPs across *GABRA*2 and captures the majority of genetic variability across the gene (Edenberg et al., 2004). It is therefore optimal in the proposed model. Although the *GABRA2* studies cited above show an association of this SNP with several problematic alcohol-use related phenotypes, other recent studies have also suggested a potential relationship between rs279871 and disruptive behavior in adolescence. Dick, Bierut, et al. (2006) have demonstrated an association between the A allele of this SNP (again, the T allele examined on the complementary strand in the current study) and the number of CD symptoms. However, in a later study, this group did not replicate this relationship but did demonstrate an association between this SNP and the YSR externalizing score (Dick, Aliev, Latendresse, Hickman, et al., 2013). This potential relationship between disruptive behavior and rs279871 builds upon our previous work

demonstrating significance for both the number of CD symptoms and the YSR externalizing score as predictors for the age of drinking initiation (Kuperman et al., 2013) and provides further support for using rs279871 to determine whether it improves our previous model.

Materials and methods

Subjects and variables

Subjects were participants in the multi-site Collaborative Study on the Genetics of Alcoholism (COGA), a project designed to explore behavioral, biochemical, genetic, neuropsychological, environmental, and neurophysiologic contributions to AD in high-risk families (defined as having at least one adult proband treated for AD) and community comparison families (Begleiter et al., 1995). Since 2005, COGA has used a prospective design focusing on participants who were 12 to 21 years of age at the start of this phase; these participants are the offspring and non-first-degree family members (e.g., nieces, nephews, grandchildren) of the original probands in high-risk and comparison families, and are followed-up every 2 years. The subject pool for this study consisted of the 820 adolescents from our previous study (Kuperman et al., 2013) who were 14 to 17 years old at the time of their assessments during the years of 2005–2007. Institutional Review Boards at all sites reviewed and approved study design. Parents provided consent for all offspring below age 18; children age 13 and older also provided consent, and children age 12 provided assent.

To achieve our goal, the *GABRA2* SNP rs279871 was selected following the rationale detailed in the introduction. Of the 820 subjects in the prior study (Kuperman et al., 2013), 691 had been genotyped for this SNP on chromosome 4 using the Sequenom MassArray (Sequenom, San Diego, CA, USA). Plink (Purcell et al., 2007) was used to confirm family structure, while PEDCHECK (O'Connell & Weeks, 1998) was used to correct for Mendelian inconsistencies.

The SNP had a genotyping rate >96.5% and was in Hardy-Weinberg equilibrium in both the EA ($p = 0.27$) and the AA ($p = 0.64$) samples. The risk allele for rs279871 contained the nucleotide thymine (T) (equivalent to the A nucleotide on the complementary strand in Edenberg et al., 2004). A chi-squared test was employed to examine whether the allele frequency of the T allele was different between high-risk and control families. To ensure an independent sample, one person per family was chosen at random to be included in the test.

To determine the contributions of interview age, sex, and ancestry, these variables were included in the *new* models as covariates. Ancestry was determined through the use of a set of 64 ancestry informative markers (AIM), which were genotyped as part of a 96 SNP Biorepository Panel by the Rutgers University Cell and DNA Repository. SNPrelate, a library function in R (R Development Core Team, 2013), was applied to these markers to estimate principal components in order to assign ancestry groups at family and individual levels. HapMap3 populations were included as reference groups. Based on the first two principal components, individuals were assigned to one of three ancestry groups: EA, AA, or Other. Seventeen subjects were assigned to the "Other" group; due to its small size, this group was removed from further analysis. This resulted in a final sample size of 674 subjects, with 451 (66.9%) identified as EA and 223 (33.1%) as AA. There were no significant differences with respect to gender, interview age, or predictor scores for subjects with genotype data (674) and those who were dropped (146).

Statistical analyses

Cox proportional hazards modeling using SAS PROC PHREG (SAS, 2011), with the COVSANDWICH (AGGREGATE) option to adjust for correlated familial data, was used to study the effect of predictor variables on the probability of alcohol initiation. Use of Cox proportional hazards modeling also accounts for potential variable confounds and reports the

unique contribution of each to the model. All model assumptions were checked for violations of the proportional hazard assumption and overall model adequacy (Box-Steffensmeier & Jones, 2004). As a preliminary step, single-predictor models were used to determine each predictor's individual effect on the likelihood of alcohol initiation. The main analysis considered five multi-predictor Cox proportional hazards models. **Model 1**, our initial reference model, contained the five factors from our previous study (Kuperman et al., 2013): 1) most best friends drank (MBFD: Yes versus No); 2) number of CD symptoms (0 to 15); 3) YSR externalizing score (0 to 64); 4) YSR social problems score (0 to 22); and 5) member of a high-risk AD family (Yes versus No). **Model 2** removed any non-significant predictors from Model 1. **Model 3** added the covariates of: 1) sex (Male = 1 versus Female = 0); 2) interview age (14, 15, 16, or 17); and 3) ancestry (EA versus AA) to Model 2 because the previous study (Kuperman et al., 2013) did not examine their effects. **Model 4** included all variables in Model 3 plus the number of T alleles (0, 1, or 2) at rs279871 (an additive genetic model for the high-risk T allele). **Model 5** also included all variables in Model 3 but changed the genetic predictor from an additive to a recessive genetic model indicator for a homozygous TT genotype (two copies of the T allele versus 0 or 1 copy). Likelihood Ratio Test (LRT) and model fit statistics Akaike information criterion (AIC) and Schwarz Bayesian criterion (SBC) were used to determine the final model. After the final model was determined, exploratory analysis for a possible genetic by environmental (G X E) interaction was performed.

Results

Subject characteristics

Table 1 provides the descriptive statistics of the nine variables of interest in the final sample of 674. Most subjects (86%) were members of high-risk AD families. About 40% of subjects endorsed one or more CD symptoms. Slightly less than 20% of the sample endorsed

MBFD. The average interview age was approximately 15.5 years; interview age was evenly

distributed across the 4-year age range of subjects. Using only one subject per family, the

T allele was absent in 14% of subjects, 48% had a single T allele, and 38% had two T alleles.

The frequency of the T allele was not different between control and high-risk families ($p = 0.91$);

this was true for both AE and AA control and high-risk families ($p = 0.78$ and $p = 0.62$,

respectively). A first drink was reported by slightly over 40% of subjects; on average, subjects

who drank did so 2.0 ± 1.7 years prior to being assessed.

Preliminary Analysis

Table 2 presents single-predictor Cox proportional hazards models for each of the five

original predictors, the three covariates, and the SNP variant. All five of the original predictors,

two of the covariates (age and ancestry), and the homozygous TT genotype (but not the number

of T alleles) each had a significant unadjusted hazard ratio (uHR) that was positively related to

the risk for initiating drinking.

Main Analysis

The results of the five multi-predictor Cox proportional hazards models are shown in

Table 3. **Model 1** was the initial reference model and contained the five predictors from our

previous study (Kuperman et al., 2013); the outcome of this model is based on the 674 subjects

with available genotyping in the current study sample. Four of the five variables had adjusted

hazard ratios (aHRs) that suggested an increase in the risk for drinking initiation: the number of

CD symptoms, MBFD, YSR externalizing score, and being a member of a high-risk AD family;

the YSR social problems score was no longer significant, and was dropped from Model 1. This

resulted in a simplified reference model, **Model 2**. **Model 3** contained the four predictors in

Model 2 plus three covariates (sex, interview age, and ancestry, though only ancestry was

significant). This model became the new reference model when evaluating the impact of the chosen candidate SNP. **Model 4** contained all predictors from Model 3 plus the number of T alleles at rs279871 (additive genetic model). The likelihood of initiating drinking was 28% higher for each additional copy of the T allele (aHR = 1.28 fold, p = 0.01). **Model 5** contained all predictors from Model 3 plus the presence of a homozygous TT genotype (recessive genetic model for the high-risk T allele). The likelihood of initiating drinking was 54% higher for those with the TT genotype (aHR = 1.54-fold, $p < 0.001$). The effects of predictors and covariates common to Models 3–5 were very similar in magnitude. LRT between nested models concluded that Model 3 should be rejected in favor of both Model 4 (χ^2[df = 1] = 6.63, p = 0.001) and Model 5 (χ^2[df = 1] = 11.07, $p < 0.001$). Model fit statistics of AIC and SBC were both lower for Model 5 compared to Model 4, suggesting that Model 5 was the best overall model.

Exploratory GXE interactions were performed for the variables and covariates in **Model 5**. After controlling for behavior problems – the number of CD symptoms and the YSR externalizing score – no significant interactive effect of G (rs279871 TT genotype) by E (most best friends drank) was found for the risk of alcohol initiation.

Closer examination of Model 5 (the last column in Table 3) indicated that the four binary predictor variables have adjusted hazard ratios (aHRs) that are higher than those for the two quantitative predictors. This is due to the fact that the two non-binary predictors (the number of CD symptoms and the YSR externalizing score) each had a wider observed range. The aHR was increased 17% for each additional CD symptom (observed range of 0–10) and 3% for each unit increase of the YSR externalizing score (observed range of 0–50). Standardized aHRs (saHR) provide a better way to judge the relative impact of these predictors: (from the most influential predictor to the least) MBFD (saHR = 1.40), externalizing score (saHR = 1.32), EA ancestry

(saHR = 1.25), homozygous TT genotype of rs279871 (saHR = 1.23), number of CD symptoms (saHR = 1.22), and high-risk family membership (saHR = 1.17).

Using Model 5 to plot fitted cumulative incidence curves, Fig. 1 visually illustrates the relative contributions of the significant predictors of MBFD, the presence of the high-risk TT genotype, and ancestry. Other covariates and predictors were held constant in this fitted model using the following constraints: sex = male (1), age = 17 (sample maximum), the number of CD symptoms = 0 (sample median), membership in a high-risk family = yes (sample mode), and YSR externalizing score = 11 (sample median). The top half of Fig. 1 shows the curves for EA males, with the high-risk TT genotype present in panel A and absent in panel B. In both of these panels, the two curves represent the mean cumulative risk, and the 95% confidence interval, of initiating drinking, for subjects with and without MBFD. Analogous cumulative incidence curves are shown in the bottom half of Fig. 1 (panels C and D) for African-Americans. In Fig. 1, an EA male with both MBFD and the high-risk TT genotype had the highest probability of starting drinking at any age; specifically, Model 5 predicts that 50% of such EA males will have had their first drink by an age of 14.2 years. The median age of first drink was delayed for an EA male with MBFD and without the high-risk TT genotype to just before their 15th birthday; to soon after their 15th birthday for an AA male with both MBFD and the high-risk TT genotype; to about 15.8 years for an EA male without MBFD and with the high-risk TT genotype; and to just after their 16th birthday for an AA male with MBFD and without the high-risk TT genotype. Three groups did not reach the 50% threshold for cumulative probability of drinking by age 17. These included EA males with neither MBFD nor the high-risk TT genotype (43%), AA males without MBFD but with the high-risk TT genotype (42%), and AA males with neither MBFD nor the high-risk genotype (29%). Changing the covariate of sex to female, while leaving age

and all other predictors unchanged, results in a similar pattern of cumulative incidence curves; these curves are not reproduced.

Discussion

The inclusion of the high-risk TT genotype of *GABRA2* rs279871 resulted in model improvement even after adjusting for other substantial predictors, such as MBFD, number of CD symptoms, externalizing problems, a family history of AD, and covariates of sex, age, and ancestry. The recessive genetic model performed better than the additive genetic model. G X E interactions were not supported.

There was no evidence of a gender difference in alcohol initiation, either by itself or after adjusting for other predictors/covariates. Interview age significantly contributed only in the single-predictor model and not in multiple-predictor models. Ancestry was significant in both the single- and multiple-predictor models. Our findings that EA subjects are at higher risk for initiating drinking than AA is consistent with previous studies of age of actual drinking initiation (Alvanzo et al., 2011; Jackson, 2010; Zapolski, Pedersen, McCarthy, & Smith, 2014).

The major contribution of this study is the demonstration of the association between rs279871 and the onset of drinking, even though several studies have failed to link rs279871 with actual AD in this age group (Dick, Agrawal, et al., 2006; Dick, Aliev, Latendresse, Porjesz, et al., 2013; Melroy et al., 2014; Sakai et al., 2010). The lack of a relationship between rs279871 and adolescent AD may be due to the limited availability of alcohol to adolescents and the time required for progression from initiation of drinking, escalation of its use, and development of alcohol dependence. Two lines of evidence support this. First, the Monitoring the Future Study (Johnston et al., 2016) reports an annual prevalence for alcohol use across 8th, 10th, and 12th graders (ages 14–18, the approximate age range of our sample) of 40%; this diminishes to a

30-day prevalence of 22%, and further decreases to a daily prevalence of <1%. Second, data

from Wave 1 of the National Epidemiologic Survey on Alcohol and Related Conditions

(NESARC) suggest that, on average, it takes 5.2 years to progress from initiation of alcohol use

to the *first* symptom of AD (Alvanzo et al., 2011). In the current study, the mean age of first

drink in the 42% of subjects who reported drinking is 14.0 ± 1.9 years of age. If the NESARC

time frame is applied to this group of drinkers, development of AD would likely occur when

these individuals were in their early 20s. Of note, this is consistent with a previous COGA study

that found no relationship between AD and *GABRA2* until individuals with AD were in their

mid-twenties (Dick, Bierut, et al., 2006).

There are several potential limitations to this study. First, the majority of subjects were

offspring/descendants of the original COGA high-risk families, which may limit the ability to

generalize the results to other populations. Second, due to the number of subjects as well as

multiple testing issues, we did not perform a genome-wide association study (GWAS) that

specifically identified *GABRA2* rs279871 as the best choice for inclusion in the model. Third,

although several studies have identified both the minor and major alleles of SNPs in this region

to be the risk allele for alcohol dependence, two meta-analyses of these publications (Li et al.,

2014; Zintzaras, 2012) found the less common allele to be the risk allele. Additionally, both

meta-analyses report the distribution of the risk allele to be in the opposite directions for EA

compared to AA subjects. Fourth, while there are no gender differences in our model's

prediction of age of first drink (similar to Miech, Johnston, O'Malley, Bachman, & Schulenberg,

2016), this may be limited by our use of this covariate as a binary variable; our sample size is too

small to explore this in more detail by modeling males and females independently across race.

Fifth, though the predictors in **Model 5** have been identified as also having predictive ability for

later alcohol-use problems, the question of "how early age of first drink" actually predicts early adulthood AD has not been determined (Kuntsche, Rossow, Engels, & Kuntsche, 2016).

These identified limitations are balanced by study strengths. The selection of a *GABRA2*-SNP variant for inclusion in our model was based on an extensive literature review that suggested this gene's association with a wide range of alcohol-related problems, including a relationship with externalizing behavior, a known precursor of both drinking initiation and alcohol-use problems. Furthermore, two recent studies (Demers, Bogdan, & Agrawal, 2014; Villafuerte, Strumba, Stoltenberg, Zucker, & Burmeister, 2013) suggest that the association of *GABRA2* SNPs (including rs279871) with alcohol-related phenotypes may be mediated by impulsivity, which is itself related to externalizing behavior (Eisenberg et al., 2009). Within *GABRA2*, the rs279871 SNP was selected because in the COGA sample it is in high linkage disequilibrium with other SNPs associated with alcohol dependence in adults, and therefore represents most of the relevant genetic variability within this gene. Our finding that the major allele (T) is the risk allele is at odds with the two meta-analyses cited above; however, the relative frequencies of the major and minor alleles in our study were not significantly different in the overall sample or the results after the sample was divided into individual ancestry type. Potential causes for this discrepancy may be the much larger size of the two meta-analyses versus the sample used in the current study, though our finding are consistent with other samples in the COGA study. Because alcohol-use disorders are complex (Morozova, Goldman, Mackay, & Anholt, 2012), they may be influenced by a combination of major allele effects of some genes and minor allele effects from other genes.

Additionally, because the current study's population is re-evaluated, we have the unique capacity to determine whether our identified high-risk genotype for early alcohol initiation is

linked to young adult AD, as has been reported in previous cross-sectional COGA studies (Dick, Agrawal, et al., 2006; Dick, Bierut, et al., 2006; Edenberg et al., 2004). Because 60% of the current subjects have not yet had a drink, the 2-year follow-up design of this study will allow collection of onset age of multiple alcohol milestones (first drink, first intoxication, onset of regular drinking, onset of first DSM symptom, first DSM diagnosis, etc.) with minimal recollection bias. It will also provide some clarification as called for by Kuntsche et al. (2016) in describing the relationships of early first drink to the progression from low-level drinking to more problematic drinking in young adults. The use of survival analysis techniques will aid the understanding of how milestone progression affects the onset of AD and may help identify more specific prevention applications.

Financial Disclosure

Funding Source

Figure legend

Fig. 1. Cumulative incidence curves for the estimated likelihood of initiating drinking using Model 5 for males. European-Americans (EA) are more likely to start drinking than African-Americans (AA) at any age for the same conditions of whether most best friends drank (MBFD) and the presence of the TT genotype. For either ancestry, MBFD has a greater effect on initiating drinking than the presence of the TT genotype.

References

Achenbach, T. M., & Rescorla, L. (2001). Manual for the ASEBA school-age forms & profiles: an integrated system of multi-informant assessment. (Burlington, VT: ASEBA).

Alvanzo, A. A., Storr, C. L., La Flair, L., Green, K. M., Wagner, F. A., & Crum, R. M. (2011). Race/ethnicity and sex differences in progression from drinking initiation to the development of alcohol dependence. *Drug and Alcohol Dependence, 118*, 375–382. doi:10.1016/j.drugalcdep.2011.04.024

American Psychiatric Association. (1994). Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV), 4th edn (Washington, DC: American Psychiatric Association).

Assanangkornchai, S., Geater, A. F., Saunders, J. B., & McNeil, D. R. (2002). Effects of paternal drinking, conduct disorder and childhood home environment on the development of alcohol use disorders in a Thai population. *Addiction, 97*, 217–226.

Begleiter, H., Reich, T., Hesselbrock, V., Porjesz, B., Li, T. K., Schuckit, M., et al. (1995). The Collaborative Study on the Genetics of Alcoholism. *Alcohol Health & Research World, 19*, 228–236.

Bekman, N. M., Cummins, K., & Brown, S. A. (2010). Affective and personality risk and cognitive mediators of initial adolescent alcohol use. *Journal of Studies on Alcohol and Drugs, 71*, 570–580.

Box-Steffensmeier, J. M., & Jones, B. S. (2004). Event History Modeling: A Guide for Social Scientists. (West Nyack, NY: Cambridge University Press).

Brown, S. A., McGue, M., Maggs, J., Schulenberg, J., Hingson, R., Swartzwelder, S., et al. (2008). A developmental perspective on alcohol and youths 16 to 20 years of age. *Pediatrics, 121 Suppl 4*, S290–310. doi:10.1542/peds.2007-2243D

Capaldi, D. M., Stoolmiller, M., Kim, H. K., & Yoerger, K. (2009). Growth in alcohol use in at-risk adolescent boys: two-part random effects prediction models. *Drug and Alcohol Dependence, 105*, 109–117. doi:10.1016/j.drugalcdep.2009.06.013

Chen, X., Unger, J. B., Palmer, P., Weiner, M. D., Johnson, C. A., Wong, M. M., et al. (2002). Prior cigarette smoking initiation predicting current alcohol use: evidence for a gateway drug effect among California adolescents from eleven ethnic groups. *Addictive Behaviors, 27*, 799–817.

Clark, D. B., Parker, A. M., & Lynch, K. G. (1999). Psychopathology and substance-related problems during early adolescence: a survival analysis. *Journal of Clinical Child Psychology, 28*, 333–341. doi:10.1207/S15374424jccp280305

Covault, J., Gelernter, J., Hesselbrock, V., Nellissery, M., & Kranzler, H. R. (2004). Allelic and haplotypic association of GABRA2 with alcohol dependence. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics, 129B*, 104–109. doi:10.1002/ajmg.b.30091

Demers, C. H., Bogdan, R., & Agrawal, A. (2014). The Genetics, Neurogenetics and Pharmacogenetics of Addiction. *Current Behavioral Neuroscience Reports, 1*, 33–44. doi:10.1007/s40473-013-0004-8

Dick, D. M., Agrawal, A., Schuckit, M. A., Bierut, L., Hinrichs, A., Fox, L., et al. (2006). Marital status, alcohol dependence, and GABRA2: evidence for gene-environment correlation and interaction. *Journal of Studies on Alcohol, 67*, 185–194.

Dick, D. M., Aliev, F., Latendresse, S., Porjesz, B., Schuckit, M., Rangaswamy, M., et al. (2013). How phenotype and developmental stage affect the genes we find: GABRA2 and impulsivity. *Twin Research and Human Genetics, 16*, 661–669. doi:10.1017/thg.2013.20

Dick, D. M., Aliev, F., Latendresse, S. J., Hickman, M., Heron, J., Macleod, J., et al. (2013). Adolescent alcohol use is predicted by childhood temperament factors before age 5, with mediation through personality and peers. *Alcoholism: Clinical and Experimental Research, 37*, 2108–2117. doi:10.1111/acer.12206

Dick, D. M., Bierut, L., Hinrichs, A., Fox, L., Bucholz, K. K., Kramer, J., et al. (2006). The role of GABRA2 in risk for conduct disorder and alcohol and drug dependence across developmental stages. *Behavior Genetics, 36*, 577–590. doi:10.1007/s10519-005-9041-8

Disney, E. R., Elkins, I. J., McGue, M., & Iacono, W. G. (1999). Effects of ADHD, conduct disorder, and gender on substance use and abuse in adolescence. *American Journal of Psychiatry, 156*, 1515–1521.

Edenberg, H. J., Dick, D. M., Xuei, X., Tian, H., Almasy, L., Bauer, L. O., et al. (2004). Variations in GABRA2, encoding the alpha 2 subunit of the GABA(A) receptor, are associated with alcohol dependence and with brain oscillations. *American Journal of Human Genetics, 74*, 705–714. doi:10.1086/383283

Eisenberg, N., Valiente, C., Spinrad, T. L., Liew, J., Zhou, Q., Losoya, S. H., et al. (2009). Longitudinal relations of children's effortful control, impulsivity, and negative emotionality to their externalizing, internalizing, and co-occurring behavior problems. *Developmental Psychology, 45*, 988–1008. doi:10.1037/a0016213

Enoch, M. A. (2008). The role of GABA(A) receptors in the development of alcoholism. *Pharmacology, Biochemistry, and Behavior, 90*, 95–104. doi:10.1016/j.pbb.2008.03.007

Enoch, M. A. (2013). Genetic influences on the development of alcoholism. *Current Psychiatry Reports, 15*, 412. doi:10.1007/s11920-013-0412-1

Fehr, C., Sander, T., Tadic, A., Lenzen, K. P., Anghelescu, I., Klawe, C., et al. (2006). Confirmation of association of the GABRA2 gene with alcohol dependence by subtype-specific analysis. *Psychiatric Genetics, 16*, 9–17.

Grant, B. F. (1998). The impact of a family history of alcoholism on the relationship between age at onset of alcohol use and DSM-IV alcohol dependence: results from the National Longitudinal Alcohol Epidemiologic Survey. *Alcohol Health & Research World, 22*, 144–147.

Grant, B. F., & Dawson, D. A. (1997). Age at onset of alcohol use and its association with DSM-IV alcohol abuse and dependence: results from the National Longitudinal Alcohol Epidemiologic Survey. *Journal of Substance Abuse, 9*, 103–110.

Grant, J. D., Waldron, M., Sartor, C. E., Scherrer, J. F., Duncan, A. E., McCutcheon, V. V., et al. (2015). Parental Separation and Offspring Alcohol Involvement: Findings from Offspring

of Alcoholic and Drug Dependent Twin Fathers. *Alcoholism: Clinical and Experimental Research, 39*, 1166–1173. doi:10.1111/acer.12766

Griffin, K. W., Botvin, G. J., Epstein, J. A., Doyle, M. M., & Diaz, T. (2000). Psychosocial and behavioral factors in early adolescence as predictors of heavy drinking among high school seniors. *Journal of Studies on Alcohol, 61*, 603–606.

Jackson, K. M. (2010). Progression through early drinking milestones in an adolescent treatment sample. *Addiction, 105*, 438–449. doi:10.1111/j.1360-0443.2009.02800.x

Johnston, L. D., O'Malley, P. M., Miech, R. A., Bachman, J. G., & Schulenberg, J. E. (2016). *Monitoring the Future: National Survey Results on Drug Use, 1975–2015: Overview; Key Findings on Adolescent Drug Use.* Ann Arbor: Institute for Social Research: The University of Michigan, 630 pp.

Kuntsche, E., Rossow, I., Engels, R., & Kuntsche, S. (2016). Is 'age at first drink' a useful concept in alcohol research and prevention? We doubt that. *Addiction, 111*, 957–965. doi:10.1111/add.12980

Kuperman, S., Chan, G., Kramer, J. R., Bierut, L., Bucholz, K. K., Fox, L., et al. (2005). Relationship of age of first drink to child behavioral problems and family psychopathology. *Alcoholism: Clinical and Experimental Research, 29*, 1869–1876.

Kuperman, S., Chan, G., Kramer, J. R., Wetherill, L., Bucholz, K. K., Dick, D., et al. (2013). A model to determine the likely age of an adolescent's first drink of alcohol. *Pediatrics, 131*, 242–248. doi:10.1542/peds.2012-0880

Kuperman, S., Schlosser, S. S., Kramer, J. R., Bucholz, K., Hesselbrock, V., Reich, T., et al. (2001). Risk domains associated with an adolescent alcohol dependence diagnosis. *Addiction, 96*, 629–636. doi:10.1080/09652140020031674

Kuperman, S., Schlosser, S. S., Lidral, J., & Reich, W. (1999). Relationship of child psychopathology to parental alcoholism and antisocial personality disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 38*, 686–692. doi:10.1097/00004583-199906000-00015

Legrand, L. N., McGue, M., & Iacono, W. G. (1999). Searching for interactive effects in the etiology of early-onset substance use. *Behavior Genetics, 29*, 433–444.

Li, D., Sulovari, A., Cheng, C., Zhao, H., Kranzler, H. R., & Gelernter, J. (2014). Association of gamma-aminobutyric acid A receptor α2 gene (GABRA2) with alcohol use disorder. *Neuropsychopharmacology, 39*, 907–918. doi:10.1038/npp.2013.291

Mcculler, W. J., Sussman, S., Dent, C. W., & Teran, L. (2001). Concurrent prediction of drug use among high-risk youth. *Addictive Behaviors, 26*, 137–142. doi:S0306-4603(00)00082-4 [pii]

Melroy, W. E., Stephens, S. H., Sakai, J. T., Kamens, H. M., McQueen, M. B., Corley, R. P., et al. (2014). Examination of genetic variation in GABRA2 with conduct disorder and alcohol abuse and dependence in a longitudinal study. *Behavior Genetics, 44*, 356–367. doi:10.1007/s10519-014-9653-y

Miech, R. A., Johnston, L. D., O'Malley, P. M., Bachman, J. G., & Schulenberg, J. E. (2016). *Monitoring the Future: National Survey Results on Drug Use, 1975–2015: Volume I,*

Secondary school students (pp. 636). Retrieved from
http://monitoringthefuture.org/pubs.html#monographs

Morozova, T. V., Goldman, D., Mackay, T. F., & Anholt, R. R. (2012). The genetic basis of alcoholism: multiple phenotypes, many genes, complex networks. *Genome Biology, 13*, 239. doi:10.1186/gb-2012-13-2-239

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*, 259–266. doi:10.1086/301904

Philibert, R. A., Gunter, T. D., Beach, S. R., Brody, G. H., Hollenbeck, N., Andersen, A., et al. (2009). Role of GABRA2 on risk for alcohol, nicotine, and cannabis dependence in the Iowa Adoption Studies. *Psychiatric Genetics, 19*, 91–98. doi:10.1097/YPG.0b013e3283208026

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics, 81*, 559–575. doi:10.1086/519795

R Development Core Team. (2013). R: A language and environment for statistical computing. (Vienna, Austria: R Foundation for Statistical Computing).

Sakai, J. T., Stallings, M. C., Crowley, T. J., Gelhorn, H. L., McQueen, M. B., & Ehringer, M. A. (2010). Test of association between GABRA2 (SNP rs279871) and adolescent conduct/alcohol use disorders utilizing a sample of clinic referred youth with serious substance and conduct problems, controls and available first degree relatives. *Drug and Alcohol Dependence, 106*, 199–203. doi:10.1016/j.drugalcdep.2009.08.015

SAS. (2011). Base SAS 9.2 Procedures Guide. (Cary, NC: SAS Institute, Inc.).

Soyka, M., Preuss, U. W., Hesselbrock, V., Zill, P., Koller, G., & Bondy, B. (2008). GABA-A2 receptor subunit gene (GABRA2) polymorphisms and risk for alcohol dependence. *Journal of Psychiatric Research, 42*, 184–191. doi:10.1016/j.jpsychires.2006.11.006

Trucco, E. M., Colder, C. R., Bowker, J. C., & Wieczorek, W. F. (2011). Interpersonal Goals and Susceptibility to Peer Influence: Risk Factors for Intentions to Initiate Substance Use during Early Adolescence. *The Journal of Early Adolescence, 31*, 526–547. doi:10.1177/0272431610366252

Villafuerte, S., Strumba, V., Stoltenberg, S. F., Zucker, R. A., & Burmeister, M. (2013). Impulsiveness mediates the association between GABRA2 SNPs and lifetime alcohol problems. *Genes, Brain, and Behavior, 12*, 525–531. doi:10.1111/gbb.12039

Waldron, M., Grant, J. D., Bucholz, K. K., Lynskey, M. T., Slutske, W. S., Glowinski, A. L., et al. (2014). Parental separation and early substance involvement: results from children of alcoholic and cannabis dependent twins. *Drug and Alcohol Dependence, 134*, 78–84. doi:10.1016/j.drugalcdep.2013.09.010

Zapolski, T. C., Pedersen, S. L., McCarthy, D. M., & Smith, G. T. (2014). Less drinking, yet more problems: understanding African American drinking and related problems. *Psychological Bulletin, 140*, 188–223. doi:10.1037/a0032113

Zintzaras, E. (2012). Gamma-aminobutyric acid A receptor, α-2 (GABRA2) variants as individual markers for alcoholism: a meta-analysis. *Psychiatric Genetics, 22*, 189–196. doi:10.1097/YPG.0b013e328353ae53

Zucker, R. A., Donovan, J. E., Masten, A. S., Mattson, M. E., & Moss, H. B. (2008). Early developmental processes and the continuity of risk for underage drinking and problem drinking. *Pediatrics, 121 Suppl 4*, S252–272. doi:10.1542/peds.2007-2243B

Table 1. Descriptive Statistics of Predictors and Covariates

Source	Predictor/Covariate	Observed Range or High Risk Value	Mean (SD) / N (%)
Original	Membership High-Risk Alcohol Dependent Family	Yes	578 (85.8%)
	Number of non-Alcohol Related Conduct Disorder Symptoms	0-10	0.8 (1.3)
	Most Best Friends Drank	Yes	130 (19.3%)
	Achenbach Youth Self Report Externalizing Score	0-50	12.6 (8.4)
	Achenbach Youth Self Report Social Problems Score	0-16	3.4 (2.9)
Covariates	Interview Age in Years	14-17	15.6 (1.1)
	Sex	Male	323 (47.9%)
	Ancestry	European-American (versus African-American)	451 (66.9%)
GABRA2 rs279871	Number of Thymine Alleles	0 1 2	92 (13.6%) 324 (48.1%) 258 (38.3%)
	Homozygous Thymine Genotype	Yes	258 (38.3%)

Table 2. Single Predictor Cox Proportional Hazards Models. The unadjusted hazard ratio (uHR) predicts the change of the likelihood of alcohol initiation corresponding to a unit increase in each predictor. A predictor with a Hazard Ratio >1 increases (Hazard Ratio <1 decreases) the likelihood.

Source	Predictor Variables	Unadjusted HR (95% CI) p value
Original	Number of non-Alcohol Related Conduct Disorder Symptoms	1.23 (1.11, 1.35)***
	Membership High-Risk Alcohol Dependent Family	1.63 (1.13, 2.34)*
	Most Best Friends Drank	3.25 (2.64, 4.00)***
	Achenbach Youth Self Report Externalizing Score	1.06 (1.05, 1.07)***
	Achenbach Youth Self Report Social Problems Score	1.05 (1.01, 1.09)**
Covariates	Male versus Female	1.16 (0.93, 1.45)
	Interview Age in Years	1.20 (1.07, 1.35)**
	Ancestry: European versus African Americans	1.32 (1.03, 1.70)*
GABRA2 rs279871	Number of Thymine Alleles	1.12 (0.94, 1.33)
	Homozygous Thymine Genotype	1.27 (1.02, 1.58)*

*$p < 0.05$; **$p < 0.01$, ***$p < 0.001$

Table 3. Five Multi-Predictor Cox Proportional Hazards Models

	Predictor Observed Range	Estimated Adjusted Hazard Ratio (95% Confidence Interval) *p* value				
		Model 1 (Reference)	Model 2 (New Reference)	Model 3 (Model 2 + Covariates)	Model 4 (Model 3 + Number of rs279871 Thymine Alleles)	Model 5 (Model 3 + Homozygous Thymine Genotype of rs279871)
Predictors	Number of non-Alcohol Related Conduct Disorder Symptoms Range 0-10	1.14 (1.04, 1.24)**	1.14 (1.05, 1.25)**	1.17 (1.07, 1.28)***	1.17 (1.07, 1.28)***	1.17 (1.07, 1.28)***
	Membership High-Risk Alcohol Dependent Family No=0, Yes=1	1.41 (0.99, 2.00)*	1.41 (0.99, 2.00)*	1.49 (1.05, 2.12)*	1.54 (1.09, 2.19)*	1.56 (1.11, 2.21)*
	Most Best Friends Drank No=0, Yes=1	2.50 (1.98, 3.16)***	2.58 (2.04, 3.25)***	2.34 (1.82, 3.01)***	2.35 (1.83, 3.02)***	2.34 (1.83, 3.00)***
	Achenbach Youth Self Report Externalizing Score Range 0-50	1.04 (1.02, 1.06)***	1.03 (1.02, 1.04)***	1.03 (1.02, 1.04)***	1.03 (1.02, 1.05)***	1.03 (1.02, 1.05)***
	Achenbach Youth Self Report Social Problems Score Range 0-16	0.97 (0.92, 1.01)				
	Number of rs279871 Thymine Alleles (0, 1, 2)				1.28 (1.07, 1.54)**	
	Homozygous Thymine Genotype No=0, Yes=1					1.54 (1.22, 1.93)***
Covariates	Sex Female=0, Male=1			0.96 (0.76, 1.20)	0.96 (0.77, 1.20)	0.97 (0.77, 1.21)
	Age Range 14-17			1.07 (0.94, 1.21)	1.07 (0.94, 1.21)	1.08 (0.95, 1.22)
	Ancestry African=0, European=1			1.41 (1.07, 1.86)*	1.57 (1.18, 2.09)**	1.61 (1.21, 2.15)**
Model Fit Statistics	−2 log L (df)	3286.64 (5)	3288.76 (4)	3281.61 (7)	3274.98 (8)	3270.54 (8)
	AIC	3296.64	3296.76	3295.61	3290.98	3286.54
	SBC	3314.81	3311.30	3321.05	3320.06	3315.62

*$p < 0.05$; **$p < 0.01$, ***$p < 0.001$

with TT Genotype   without TT Genotype

A   B

European Americans

C   D

African Americans