Published in final edited form as:

Addict Biol. 2016 November ; 21(6): 1217–1232. doi:10.1111/adb.12286.

KAT2B Polymorphism Identified for Drug Abuse in African Americans with Regulatory Links to Drug Abuse Pathways in Human Prefrontal Cortex

Eric O. Johnson, PhD^{1,*}, Dana B. Hancock, PhD^{2,*}, Joshua L. Levy, MS³, Nathan C. Gaddis, PhD³, Grier P. Page, PhD⁴, Cristie Glasheen, PhD², Nancy L. Saccone, PhD⁵, Laura J. Bierut, MD⁶, and Alex H. Kral, PhD⁷

¹Fellow Program and Behavioral Health and Criminal Justice Division, RTI International, Research Triangle Park, NC 27709

²Behavioral and Urban Health Program, Behavioral Health and Criminal Justice Division, RTI International, Research Triangle Park, NC 27709

²Research Computing Division, RTI International, Research Triangle Park, NC 27709

³Fellow Program and Center for Genomics in Public Health and Medicine Genomics, Statistical Genetics, and Environmental Research Program, RTI International, Atlanta, GA 30341

⁴Department of Genetics, Washington University in St. Louis, MO 63110

⁵Department of Psychiatry, Washington University School of Medicine, MO 63110

⁶Behavioral and Urban Health Program, Behavioral Health and Criminal Justice Division, RTI International San Francisco, CA 94104

Abstract

Drug abuse is a common and heritable set of disorders, but the underlying genetic factors are largely unknown. We conducted genome-wide association studies of drug abuse using 7 million imputed single nucleotide polymorphisms (SNPs) and insertions/deletions in African Americans (AAs; n=3,742) and European Americans (EAs; n=6,845). Cases were drawn from the Urban Health Study of street-recruited people, who injected drugs and reported abusing opioids, cocaine, marijuana, stimulants, and/or other drugs 10 or more times in the past 30 days, and were compared to population controls. Independent replication testing was conducted in 755 AAs and 1,131 EAs from the Genetic Association Information Network. An intronic SNP (rs9829896) in the *KAT2B*

Corresponding author: Eric O. Johnson, Ph.D. (ejohnson@rti.org), RTI International, 3040 Cornwallis Road, PO Box 12194, Research Triangle Park, NC 27709.

*These authors contributed equally to this work.

Author Contributions

EOJ, DBH, LB, and AK were responsible for the study concept and design. AHK contributed to the acquisition of UHS data. EOJ was PI of grants genotyping the UHS samples and obtaining dbGaP data for controls. JLL and NCG performed QC of genotype data, genotype imputation and GWAS analysis. GPP and NLS provided statistical consultation. CG conducted data management and QC for phenotype data across contributing study cohorts. EOJ, DBH, NCG, GPP, NLS, and LJB assisted worked on interpretation of findings. EOJ and DBH drafted the manuscript. NCG, GPP, CG, NLS, LJB and AHK provided critical revision of the manuscript for important intellectual content. All authors critically reviewed content and approved final version for publication.

Eric O. Johnson, Dana B. Hancock, Joshua L. Levy, Nathan C. Gaddis, Grier P. Page, Cristie Glasheen, Nancy L. Saccone, Laura J. Bierut, and Alex H. Kral declare no conflicts of interest.

gene was significantly associated with drug abuse in AAs ($P=4.63\times 10^{-8}$) and independently replicated in AAs ($P=0.0019$). The rs9829896-C allele (frequency=12%) had odds ratios of 0.68 and 0.53 across the AA cohorts: meta-analysis $P=3.93\times 10^{-10}$. Rs9829896-C was not associated with drug abuse across the EA cohorts: frequency=36% and meta-analysis $P=0.12$. Using dorsolateral prefrontal cortex data from the BrainCloud cohort, we found that rs9829896-C was associated with reduced *KAT2B* expression in AAs ($n=113$, $P=0.050$) but not EAs ($n=110$, $P=0.39$). *KAT2B* encodes a transcriptional regulator in the cyclic adenosine monophosphate and dopamine signaling pathways, and rs9829896-C was associated with expression of genes in these pathways: reduced *CREBBP* expression ($P=0.011$) and increased *OPRM1* expression ($P=0.016$), both in AAs only. Our study identified the *KAT2B* SNP rs9829896 as having novel and biologically plausible associations with drug abuse and gene expression in AAs but not EAs, suggesting ancestry-specific effects.

Keywords

African American; gene expression; GWAS; *KAT2B* and Substance Abuse

INTRODUCTION

Drug abuse is a major public health burden with profound personal and societal consequences. Abuse of drugs is both common and heritable, with approximately 50% of the variability in the population being attributable to genetic factors.(Agrawal et al., 2012) Although heritability estimates vary somewhat by specific drug,(Verweij et al., 2010) twin studies suggest that common genetic pathways are likely shared across multiple drug categories, without excluding the possibility that other genetic factors influence the risk of abusing specific drugs.(Kendler et al., 2003; Tsuang et al., 2001)

Genetic associations have been established for *CHRNA5/A3/B4* polymorphisms with cigarette smoking(Saccone et al., 2007; Thorgeirsson et al., 2010) and cocaine dependence, (Gruza et al., 2008; Sherva et al., 2010) and *ADH1B* and *ALDH2* polymorphisms with alcohol dependence.(Frank et al., 2012; Gelernter et al., 2014a; Park et al., 2013; Quillen et al., 2014) Other plausible genetic associations for drug abuse behaviors have been identified but have not been conclusively replicated.(Wang et al., 2012) One challenge to identifying genetic polymorphisms contributing to drug abuse has been development of large sample sizes for genome-wide association studies (GWAS). The largest prior GWAS ($N=3,318$ African Americans [AAs] and 2,379 European Americans [EAs]) identified significant single nucleotide polymorphism (SNP) associations with opioid (Gelernter et al., 2013) or cocaine dependence (Gelernter et al., 2014b), which await independent replication. Other GWAS of drug abuse behaviors, containing up to 3,053 subjects, identified only nominally significant SNP associations.(Agrawal et al., 2014; Liu et al., 2013; Nielsen et al., 2010; Uhl et al., 2008) In this study, we assembled a cohort of 10,587 subjects (3,742 AAs and 6,845 EAs) to conduct the largest GWAS of drug abuse, other than nicotine or alcohol, to date. Drug abuse cases were street-recruited people who inject drugs (PWID) and who reported abusing opioids, cocaine, marijuana, stimulants, and/or other drugs 10 or more times in the past 30 days – a phenotype highly predictive of clinical abuse/dependence (see Methods).

Resembling approaches taken elsewhere (Cross-Disorder Group of the Psychiatric Genomics, 2013; Ruderfer et al., 2013; Steinberg et al., 2014; Wetherill et al., 2014; Yu et al., 2014), combining cases who abuse drugs across one or more categories for GWAS analyses focuses discovery on identifying variants that contribute the genetic susceptibility shared across drugs.

MATERIALS and METHODS

Drug Abuse Cases from the Urban Health Study (UHS)

Drug abuse cases were drawn from the UHS, one of the largest studies of street-recruited PWID in North America. Details on the study are provided in the Supplemental Information and elsewhere.(Kral et al., 2003) Briefly, eligibility criteria for study entry included injection of a drug in the past 30 days as verified by signs of venipuncture, ability to provide informed consent, age 18 or older, and ability to speak English or Spanish. Stored serum samples from UHS participants were selected for DNA extraction and genotyping on the Illumina Omni1-Quad BeadChip (see Supplemental Information). All of the genotyped UHS subjects were PWID but varied by reported drugs abused (Table 1). Past 30 day use of illicit drugs was assessed by self-report among the participating PWIDs, no confirmatory toxicology was performed. Most subjects were defined as heroin or other opioid abusers, who met the Office of National Drug Control Policy definition of heroin abuse by using 10 or more times in the past 30 days.(Morral et al., 2000; Rhodes et al., 2000) Other genotyped UHS subjects reported abuse (using 10 or more times in the past 30 days) of cocaine, marijuana, or amphetamine/stimulants specifically. UHS subjects included as “other” captured PWIDs reporting lower levels of abuse across a variety of illicit drugs, which cumulatively were used 10 or more times in the past 30 days.

To assess the correspondence between definition of a drug abuse case used here (past 30 day injection use of an illicit drug, plus use of an illicit drug 10 or more times in the past 30 days) and the phenotype of any DSM-IV drug abuse/dependence, we analyzed data from the 2008 through 2012 National Surveys on Drug Use and Health (NSDUHs). Eighty-seven percent of those who mirrored our study case criteria (unweighted N=225) met a diagnosis of DSM-IV abuse/dependence for any illicit drug in the past 12 months (see Supplemental Information and Table S1). Thus, it is very likely that the drug abuse cases from UHS have clinically important abuse/dependence and that any findings from this study should be applicable to and testable in cohorts with clinical diagnostic phenotypes.

Population Controls

Drug abuse cases from the UHS were compared to population controls assembled using the pipeline presented in Figure 1. Controls were obtained from the database of Genotypes and Phenotypes (dbGaP). Six cohorts passed all steps in our selection pipeline (see Supplemental Information); see Tables S2 and S3 for dbGaP accession numbers and consent groups. Controls were genotyped on one of three Illumina BeadChip arrays: Omni1-Quad, 1M-Duo, or Omni2.5.

Two of the selected control cohorts were ascertained for addiction phenotypes: Collaborative Genetic Study of Nicotine Dependence (COGEND) and Study of Addiction: Genetics and Environment (SAGE). We excluded subjects with a positive report of opioid, marijuana, cocaine, or other drug dependence, as defined by the Diagnostic and Statistical Manual of Mental Disorders (4th ed.; DSM-IV). None of the other population control samples were ascertained for drug abuse or similar phenotypes and did not have drug use histories available.

Using unassessed population controls carries the possibility of misclassifying true cases as controls, which is a key potential threat to the validity of using population controls and would reduce power to detect true associations. (Ho and Lange, 2010; Wellcome Trust Case Control, 2007) A recent analysis of nationally representative data estimates that 2.6% of the U.S. population 13 years of age or older has a lifetime history of injecting drugs. (Lansky et al., 2014) Thus, we expect the rate of misclassification of true cases (injection drug use in the past 30 days, plus use of an illicit drug 10+ times in the past 30 days) as population controls to not exceed 2.6%. This level of misclassification seems unlikely to affect study results but to the extent it does, it would increase risk of false negative findings and reduce risk of false positive findings. Additionally, our replication data set includes assessed controls (see *Replication Study Cohort and Analyses*), which helps to minimize any potential bias from using population controls affecting the conclusions of this study.

Quality Control and Genotype Imputation

Prior evaluations have shown that GWAS can be validly conducted with cases compared to population controls. (Garner et al., 2014; Ho and Lange, 2010; Mukherjee et al., 2011; Wellcome Trust Case Control, 2007; Zhuang et al., 2010) Our quality control procedures, as outlined in the Supplementary Information, were modeled after other GWAS that used population controls. (Luca et al., 2008; Silverberg et al., 2009) Subject-level and SNP-level quality control are detailed in Tables S2 and S3, and Tables S4 and S5, respectively.

Before combining the UHS drug abuse cases in each ancestry group with population controls of the same ancestry, we needed to ensure that there were no underlying biases driven by any one of the control data sets. To check for such bias, we made pairwise comparisons between the control data sets based on overlapping genotypes and arbitrary case assignments to subjects from one data set and control assignments to subjects from the other data set. Using SNP association results with the arbitrary case/control assignments, genomic control inflation factors (λ_{gc}) were computed. Control data sets used in the final analysis data set had λ_{gc} 1.05 for all pairwise comparisons (see Supplemental Information, Tables S6 and S7).

Our final analysis data set included 3,742 AAs (2,017 UHS cases and 1,725 population controls) and 6,845 EAs (1,142 UHS cases and 5,703 population controls). We previously found that combining imputed SNPs derived from cases and controls genotyped on different arrays has the potential to create artifactual differences leading to biased genotype-phenotype associations. (Johnson et al., 2013) To circumvent this potential bias, imputation was conducted using the common set of genotyped SNPs available across all EAs (344,736

SNPs) or AAs (336,412 SNPs) as the input genotypes. We have shown this approach eliminates this potential bias and retains high imputation quality. (Johnson et al., 2013)

Genome-wide genotype imputation of SNPs and insertion/deletion polymorphisms (indels) was conducted separately by ancestry group using IMPUTE2 (Howie et al., 2011) with reference to all 1000 Genomes haplotype panels (see Supplemental Information). (Durbin et al., 2010) For each imputed SNP and indel, a dosage value (a fractional value between 0 and 2 indicating the expected number of minor allele copies) and an “info” metric (a fractional value typically between 0 and 1 with higher values indicating SNPs imputed with higher certainty) were generated. We filtered imputed SNPs and indels based on minor allele frequency (MAF)<1% in the 1000 Genomes populations of African (AFR) or European (EUR) descent, leaving 6,809,218 SNPs and 739,119 indels in the AAs and 6,727,359 SNPs and 729,634 indels in the EAs.

Discovery GWAS Analyses

GWAS analyses were conducted in each ancestry group to test the imputed SNP and indel genotype dosages for association with drug abuse using logistic regression models in ProbABEL (Aulchenko et al., 2007). Drug abuse was coded as a dichotomous variable: ‘1’ corresponding to UHS cases who reported abusing drugs (regardless of type of drug category or frequency of abuse) and ‘0’ corresponding to population controls. The regression models were adjusted for sex and selected principal component eigenvectors to remove potential bias due to population stratification. For each ancestry group, 10 eigenvectors were generated using EIGENSTRAT (Price et al., 2006) analyses with a pruned set of genotyped SNPs in linkage equilibrium ($r^2 < 0.5$). We selected the number of eigenvectors needed to cumulatively explain >90% of the phenotypic variance, resulting in the inclusion of three eigenvectors as covariates for each ancestry group (Table S8). The ancestry-specific GWAS results were combined via fixed-effects, sample-size weighted meta-analysis, as done in other multi-ancestry GWAS meta-analyses. (Guo et al., 2013) using the METAL program. (Willer et al., 2010) Discovery analyses could not account for smoking or alcohol abuse history because these data were not collected in the UHS and not available in some of control data sets.

The GWAS resulted in $\lambda_{gc}=1.036$ for AAs (Figure S1a) and $\lambda_{gc}=1.069$ for EAs (Figure S1b). Given the slight λ_{gc} inflation in EAs, genomic control was applied to the EA-specific GWAS results prior to combining with the AA-specific GWAS results, resulting in $\lambda_{gc}=1.020$ for the multi-ancestry GWAS meta-analysis (Figure S1c). $P < 5 \times 10^{-8}$ was used to declare genome-wide statistical significance.

Replication Study Cohort and Analyses

One dbGaP cohort with drug abuse data available on both AA and EA participants was utilized for independent replication of genome-wide significant SNP associations: the Genetic Association Information Network (GAIN) “GWAS of Schizophrenia” (dbGaP accession number phs000021.v3.p2). We focused only on the schizophrenia controls consented for general research (979 AAs and 1,442 EAs) and processed them using the same quality control procedures as used for the discovery cohorts.

As described in Shi et al.,(Shi et al., 2009) GAIN controls were recruited from the general population using random digit dialing sampling and administered questionnaires to ascertain their psychiatric health histories, which included DSM-IV assessment for drug abuse and dependence. Exclusions were made for reporting any psychotic or bipolar disorder, leaving a control set representative of the general population. We used these data to define illicit drug abuse cases and controls, who according to the GAIN recruitment criteria had no co-morbid psychiatric disease. We defined illicit drug abuse cases as follows: (1) ever using at least one drug (heroin/opioids, cocaine, marijuana, amphetamines/stimulants, or other drugs) and (2) reporting at least one DSM-IV symptom of drug dependence. Controls were defined as those who reported never using any drug. Replication analyses included 755 AAs (259 cases and 496 controls) and 1,131 EAs (273 cases and 858 controls), who passed quality control, had complete DSM-IV and covariate data available, and met our case/control definitions: see Table 1. We also conducted a sensitivity analysis in which we increased the number of dependence symptoms from 1 to 3 to define a case, matching the threshold for a diagnosis but reducing sample size. Consistency of effect across the number of symptoms required, despite lower sample size and power, would support the conclusion that SNPs associated with any illicit drug abuse in the discovery analyses that replicate in GAIN are generalizable to clinical phenotypes such as DSM-IV.

The GAIN EAs and AAs were analyzed separately. Genotyped SNPs passing quality control were used as input for 1000 Genomes imputation, following the same procedure outlined in the Supplemental Information. A subset of genotyped SNPs in linkage equilibrium ($r^2 < 0.5$) were used to compute eigenvectors, from which we selected the number needed to explain >90% of the phenotype variance: 4 for GAIN AAs and 6 for GAIN EAs (Table S8). Logistic regression models were run to replicate SNP associations with illicit drug abuse. Additional data were available on alcohol abuse and nicotine dependence in the GAIN cohort. To focus on SNP/indel associations with illicit drug abuse, the replication analysis included sex, selected eigenvectors, alcohol abuse, and nicotine dependence as covariates (see Supplemental Information).

Bioinformatics Analyses

Linkage disequilibrium patterns were discerned using LocusZoom (Pruim et al., 2010) with populations of EUR or AFR ancestry. Novel genes implicated in our GWAS analyses were input into GeneMANIA (Zuberi et al., 2013) to generate hypotheses about function by building networks of genes that share pathways, physical interactions, protein domains, and other features.

Expression Quantitative Trait Loci (eQTL) Analyses

To further characterize SNP regulatory potential, we utilized genome-wide SNP genotype (Illumina Human1M-Duo [version 3] and HumanHap650Y [version 3]) and gene expression (Illumina Human 49K Oligo) data made available to the scientific community by the BrainCloud project and dbGaP (accession number phs000417.v2.p1).(Colantuoni et al., 2011) These data enabled us to implicate expression quantitative trait loci (eQTLs) by testing SNP associations with mRNA expression levels. Post-mortem dorsolateral prefrontal cortex samples were retrieved, as previously described (Lipska et al., 2006), from 113 AAs

and 110 EAs, ranging in age from 0 to 78 years old, who had no neuropathological or neuropsychiatric diagnoses. Their drug abuse histories are unknown. We did not analyze the fetal samples, given the observed differences in gene expression patterns between fetal and postnatal periods.(Colantuoni et al., 2011)

We used the BrainCloud software application to test additive SNP genotypes for association with expression level separately by ancestry using the best fit procedure under a general linear model, as described at <http://braincloud.jhmi.edu/>. SNP-expression associations were tested for the most proximally located gene (*cis*-acting effect) and distally located genes of interest (*trans*-acting effect).

RESULTS

Table 1 presents the drug abuse histories reported by cases from UHS and the GAIN replication cohort. Among the 3,159 UHS drug abuse cases, the majority (63.4%) reported abusing heroin/opioids 10 or more times in the past 30 days; on average, they used heroin/opioids 80.9 times over the past 30-day period, indicating their high-degree of abuse. The remainder of the UHS cases reported past 30-day abuse of cocaine, marijuana, or amphetamine/stimulants or abuse of other drugs. Nearly one-third of the UHS cases reported abusing more than one drug in the past 30 days. Among the cases from GAIN, a higher percentage abused more than one drug, as expected, given that abuse was based on lifetime history here compared to past 30 days in UHS. Marijuana was the most frequently reported drug abuse in GAIN.

Rs9829896 Identified in the AA-specific GWAS and Replicated

SNP/indel associations with drug abuse at $P < 5 \times 10^{-6}$ are presented from the AA-specific (Table S9), EA-specific (Table S10), and meta-analysis (Table S11) analyses. No SNPs or indels met genome-wide statistical significance in the EA-specific GWAS (Figure S2b) or the GWAS meta-analysis of both ancestry groups (Figure S2c). However, in the AA-specific GWAS, a genome-wide significant SNP association was observed for the rs9829896 SNP on chromosome 3 ($P = 4.63 \times 10^{-8}$, Figure S2a). The rs9829896-C allele was associated with reduced risk of drug abuse: odds ratio (OR)=0.68 (Table 2). Rs9829896 is an intronic SNP in the gene encoding K(lysine) acetyltransferase 2B (*KAT2B*). Rs9829896 is in low to moderate linkage disequilibrium ($r^2 < 0.6$ in the 1000 Genomes AFR panel) with nearby SNPs (Figure S3).

The rs9829896 association with drug abuse was replicated at $P = 0.0019$ in AAs from GAIN, and its OR of 0.53 was somewhat stronger than the discovery cohort (Table 2). Meta-analysis of rs9829896 across the AA cohorts resulted in $P = 3.93 \times 10^{-10}$. Rs9829896 was genotyped in the UHS cohort and imputed with high quality in the GAIN cohort (info 0.99, Table 2). Sensitivity analysis, in which we increased the required number of symptoms from 1 to 3 in the GAIN cohort to reach the threshold for a dependence diagnosis, maintained statistical significance ($p = 0.015$) and a nearly identical OR (0.55) despite the reduced number of cases and statistical power (Table S12).

To assess the degree to which rs9829896 association with drug abuse was a general effect or driven by abuse of a specific drug, we tested rs9829896 for association with each illicit drug category separately. As shown in Table 3, the rs9829896-C allele was significantly and consistently associated with reduced risk of abuse for all drug categories in the meta-analysis of UHS and GAIN AAs, with degree of significance varying largely by the number of cases available by illicit drug category and thus reflecting reduced statistical power (Table S13).

In EAs, the directionality was consistent with the AAs, the rs9829896-C allele being associated with reduced risk of drug abuse; however, the effect size was weaker, and the association was not significant in either the discovery or replication cohort (Table 2): meta-analysis $P=0.12$. The rs9829896-C allele occurred less frequently in AAs (frequency=12%) than in EAs (frequency=34%–36%). These allele frequencies are very similar to those observed in the 1000 Genomes AFR (10.2%) and EUR (35.5%) reference populations.

Rs9829896 as a Cis-acting and Trans-acting eQTL in Human Prefrontal Cortex

We evaluated the regulatory potential of rs9829896 as a *cis*-eQTL for *KAT2B* using BrainCloud. As shown in Figure 2a, the rs9829896-C allele was associated with lower *KAT2B* expression in AAs ($P=0.050$) but not in EAs ($P=0.39$).

Because *KAT2B* is a novel genetic susceptibility factor for drug abuse, we built a GeneMANIA (Zuberi et al., 2013) network centered around *KAT2B* to find related genes that may have functional links to drug abuse. Figure 3 shows that 20 genes were implicated in the *KAT2B* network, based on their co-expression, physical interactions, predicted functional relationships (e.g., protein interactions), and pathway data. We investigated rs9829896 as a *trans*-eQTL for transcripts encoded by the 20 genes identified in the *KAT2B* network via GeneMANIA. Two gene transcripts were associated with rs9829896 at $P<0.05$: the C allele being associated with lower expression of *MAML1* ($P=0.043$) and *CREBBP* ($P=0.011$) in AAs only. *MAML1* functions as a transcriptional co-activator of the Notch signaling pathway that is involved in cell fate determination (Wu et al., 2000). Notch signaling has been indicated for altering drug reward properties in *Drosophila melanogaster* (Kaun et al., 2011), but the function of *MAML1* in relation to drug abuse is unclear. In contrast, CREB binding protein (CREBBP) is a member of the cyclic adenosine monophosphate (cAMP) pathway that has been widely studied for drug addiction. Figure 2b shows that rs9829896-C was associated with lower *CREBBP* expression in AAs ($P=0.011$) but not in EAs ($P=1.00$).

We additionally tested rs9829896 as a *trans*-eQTL for the μ -opioid receptor (*OPRM1*) gene transcript, even though it is not part of the GeneMANIA identified gene network, given this gene's suggested role in the cAMP pathway and drug addiction (Gris et al., 2010; Zhang et al., 2013). Similar to the eQTL patterns observed for *KAT2B* and *CREBBP*, rs9829896 was associated with *OPRM1* expression in AAs ($P=0.016$) but not in EAs ($P=0.12$). The drug abuse protective rs9829896-C allele was associated with higher *OPRM1* expression (Figure 2c).

In silico evaluation of SNPs previously implicated in prior GWAS of drug abuse phenotypes

We also evaluated associations of top SNP findings from prior GWAS of specific drugs, including heroin/opioids (Gelernter et al., 2013; Liu et al., 2013; Nielsen et al., 2010) cocaine (Gelernter et al., 2014b) and marijuana (Agrawal et al., 2014). Results for 96 of the 127 previously reported SNPs are shown in Table S14. The other 31 SNPs were not available due to $MAF < 1\%$ in either the 1000 Genomes AFR or EUR panels. We found nominally significant associations (meta-analysis $P < 0.05$) for three SNPs previously implicated for marijuana: rs9891146 (a missense SNP in *C17orf58*, $P = 0.0059$), rs17552189 (an intronic SNP in *FLJ30838*, $P = 0.020$), and rs2381356 (an intronic SNP in the kynureninase [*KYNU*] gene, $P = 0.040$): see Table S14 for details. Their directions of association were consistent with the prior GWAS of marijuana use disorder (Agrawal et al., 2014). The rs9891146 association was driven by EAs, mimicking its prior implication in this ancestry group; the associations for rs17552189 and rs2381356, which were originally implicated in a combined analysis of EAs and AAs, were strongest in AAs.

The prior GWAS of marijuana use disorder were conducted in the SAGE cohort (Agrawal et al., 2014). Since a subset of the population controls used for our GWAS were drawn from SAGE, we repeated our analyses after excluding controls from SAGE and the similarly ascertained COGEND cohort. Despite the reduced sample size — nearly 3,600 fewer controls, rs9891146 was associated with drug abuse at meta-analysis $P = 0.026$, thus providing independent nominally significant evidence for association: OR (95% CI) = 1.13 (1.00–1.27) in EAs and 1.08 (0.94–1.24) in AAs. Rs8071463 and rs9897982 both had meta-analysis P values of 0.20 in the reduced sample but consistent directions of association with the full sample.

Evaluation of rs9891146 in the BrainCloud cohort showed that its minor allele (C) was associated with higher *C17orf58* expression ($P = 9.32 \times 10^{-4}$ in AAs and $P = 0.043$ in EAs, Figure S4).

DISCUSSION

Our study identified the *KAT2B* SNP rs9829896 as having a novel and statistically significant association with drug abuse among AAs. We replicated the rs9829896 association in an independent cohort of AAs and found that its C allele was associated with reduced risk of drug abuse, showing consistent ORs (0.68 and 0.53) across the AA cohorts. This protective effect was shared across multiple drug categories, rather than being driven by any specific drug. We implicated rs9829896 as a *cis*-eQTL for the *KAT2B* gene and as a *trans*-eQTL for the *CREBBP* and *OPRM1* genes, using mRNA expression levels measured in human prefrontal cortex. The proteins encoded by *CREBBP* and *OPRM1* are members of the cAMP and dopamine pathways, respectively, and both pathways have been widely studied for their roles in drug addiction. Our regulatory evidence suggests that rs9829896 may be a biologically important SNP that may contribute to differential expression of cAMP and dopamine pathway genes in human prefrontal cortex among AAs: pathways by which the C allele may have its observed protective effect on risk of drug abuse also among AAs.

The cAMP pathway is activated by drugs of abuse,(Nestler, 2001) and the protein encoded by *KAT2B* is a direct regulator of this pathway.(Ravnskjaer et al., 2013) *KAT2B* promotes transcription activation by inducing epigenetic changes (specifically, increased histone H3K9 acetylation) that enhance the occupancy of promoter binding sites by the CREB protein and its cofactor (CREB regulated transcription coactivator 2, CRTC2), leading to increased expression of target genes.(Ravnskjaer et al., 2013) The CREB protein is bound by another transcriptional regulator, known as CREBBP (or CBP). CREB and CREBBP were recently highlighted in a pathway analysis using GWAS results for opioid dependence. (Gelernter et al., 2013) Our GeneMANIA network drew a biological connection between *KAT2B* and *CREBBP* based on their physical interaction, pathways, and shared protein domains, and we found that expression of both genes was down-regulated by the protective allele for drug abuse, rs9829896-C, specifically in AAs. The *trans*-eQTL pattern between the protective rs9829896-C allele and reduced *CREBBP* gene expression is consistent with prior indications related to drug-seeking behavior. Mice with *CREBBP* depleted in the nucleus accumbens show impaired sensitivity and reward from cocaine exposure, suggesting that behavioral response to cocaine is regulated, at least in part, by *CREBBP*.(Malvaez et al., 2011) Assuming that the same pattern holds in humans, we expect that attenuation in drug rewarding effects due to reduced *CREBBP* expression would decrease susceptibility to abusing drugs. Altogether, these gene expression patterns for *KAT2B* and *CREBBP* suggest a common pathway leading to downstream effects on CREB and the broader cAMP pathway.

The protective rs9829896-C allele was also associated with increased *OPRM1* gene expression. The *OPRM1* gene has long been suspected to influence susceptibility to drug addiction, given that the protein encoded by *OPRM1* is a principal target of natural and synthetic opioids. Our observed pattern for *OPRM1* expression is consistent with prior reports showing lower *OPRM1* expression levels in putamen(Sullivan et al., 2013) and frontal cortex(Ferrer-Alcon et al., 2004) of heroin abusers compared to normal controls. Down-regulation of *OPRM1* may play a role in increasing risk of heroin abuse, by depriving the opioid system of a key compensatory response to heroin exposure.(Oertel et al., 2012) In contrast, up-regulation of *OPRM1* may lead to decreased risk of heroin and other drug abuse, as suggested by the trans-regulatory effect of rs9829896-C.

The *KAT2B* gene region has not been previously implicated for any substance abuse-related phenotypes, but SNPs in this region were nominally significant in two GWAS of psychiatric traits: (1) schizophrenia and bipolar disorder using GAIN EAs(Wang et al., 2010) and (2) temperament in bipolar disorder using EAs from the Bipolar Genome Study (BiGS). (Greenwood et al., 2012) None of the seven SNPs implicated in the GAIN GWAS were associated with drug abuse in our study (Table S15). However, all four of the SNPs implicated in the BiGS GWAS were associated at $P < 0.05$ in our multiancestry GWAS meta-analysis of drug abuse (lowest $P = 1.20 \times 10^{-4}$, Table S15), and their associations had smaller P values and stronger magnitudes of association in AAs compared to EAs. Linkage disequilibrium patterns between rs9829896 and each of the 11 previously implicated SNPs (Figures S5 and S6) show high D' values (up to 1 in both EUR and AFR) but low to modest r^2 values (r^2 0.41 in EUR and r^2 0.25 in AFR). These prior implications give credence to *KAT2B* as a genetic risk factor for drug abuse behavior and mental disorders more broadly,

but suggest that different alleles across *KAT2B* may give rise to disorder-specific associations.

Despite its consistent association in AAs, rs9829896 was not associated with drug abuse in EAs. Several other examples of ancestry-specific genetic associations have been reported for complex diseases.(Keebler et al., 2009; Namjou et al., 2013; Torgerson et al., 2011) Heterogeneous SNP associations across ancestry groups may reflect differing allele frequencies and effect sizes, varied linkage disequilibrium patterns with causal variants, modifying genetic backgrounds, or environmental exposures. Our findings implicated rs9829896 as both a *cis*- and *trans*-eQTL, altering gene expression levels in AAs but not in EAs. Prior comparisons across the diverse HapMap populations showed that *cis*-eQTL associations are often shared across ancestry groups but that the majority of *cis*-eQTL associations are specific to only one ancestry group.(Stranger et al., 2012) It is plausible that the ancestry-specific SNP association results observed between rs9829896 and drug abuse were driven by *cis*- and *trans*-eQTL effects that differ between AAs and EAs.

The primary limitation of this study is combining drug abuse cases from one cohort and population controls drawn from several other cohorts. This design enabled us to construct the largest GWAS to date for a drug abuse phenotype, other than nicotine and alcohol; and there are other examples of GWAS comparing cases to population controls.(Luca et al., 2008; Silverberg et al., 2009; Wellcome Trust Case Control, 2007) Nonetheless, we recognize the caveats of using population controls. Potential phenotype misclassification among controls may reduce statistical power, generating false negative results. However, the relatively low prevalence of our phenotype (injection drug use in the past 30 days plus using illicit drugs 10+ times in the past 30 days) in the U.S. population (lifetime history of injecting drugs of 2.6% (Lansky et al., 2014)) suggests that phenotype misclassification was unlikely to substantial effect the study results.

Population stratification or other underlying differences (e.g., genotyping batch effects) between the different cases and controls cohorts may give rise to false positive associations. To minimize this possibility, we followed a carefully constructed pipeline (Figure 1) that included stringent quality control, pairwise comparisons of control cohorts to ensure that λ_{gc} values were not inflated, use of only SNPs genotyped across all cohorts as inputs for imputation,(Johnson et al., 2013) adjustment for eigenvectors to circumvent population stratification, and application of genomic control to reduce residual inflation. The λ_{gc} values from our GWAS analyses (e.g., $\lambda_{gc}=1.02$ for the GWAS meta-analysis) suggest that our pipeline minimized the potential for false positive associations.

The most compelling evidence for lack of bias is that our genome-wide significant finding was replicated in the independent GAIN cohort, which had data available on lifetime drug abuse histories for all of the defined cases and controls. Our discovery cohort also found nominal evidence of replication for a SNP previously reported to be associated with DSM-IV marijuana in a cohort using DSM diagnoses for identifying both cases and controls, further suggesting that the use of population controls functions acceptably for the study of drug abuse phenotypes.

Another limitation of this study involved the limited availability of drug, smoking, and alcohol histories for use in the discovery GWAS analyses. Drug abuse cases were defined by reported injection and use of illicit drugs 10 or more times in the past 30 days; lifetime drug abuse history was not known, and DSM-IV or other diagnostic criteria were not assessed. However, the ascertainment of cases from UHS (a street-recruited cohort of PWID) and the high correlation between the study threshold and clinical levels of any abuse/dependence, a positive predictive value of 87% in analysis of NSDUH data (see Methods), suggests that the cases very likely meet diagnostic criteria for abuse/dependence. Furthermore, sensitivity analyses in our replication cohort (GAIN) showed that increasing the symptom threshold from 1 to 3 maintained significant replication in spite of the reduced number of cases and statistical power ($P=0.0019$ to 0.015 , Table S12) and very similar magnitudes of association ($OR=0.53$ to 0.55 , Table S12). These results support the generalizability of this study's genome-wide significant and replicated association to other drug abuse phenotypes, including DSM criteria. Other AA cohorts with a large number of DSM-IV or other defined abuse/dependence cases and nondependent controls would be appropriate for future replication testing of rs9829896.

The final limitation of this study is that the rs9829896-*KAT2B* expression analysis of the BrainCloud data resulted in a P value that fell just at the 0.05 threshold for declaring statistical significance. We are not aware of any other currently available AA data sets to corroborate our eQTL finding. However, the consistency of rs9829896 association of expression across *KAT2B*, *CREBBP*, and *OPRM1* among AAs and not in EAs lend some support to the potential of this SNP as tagging differences in expression of these drug abuse associated genes. Future gene expression studies using a large AA sample of prefrontal cortex and other addiction-relevant brain regions are needed.

The strengths of this study come from combining multiple-ancestry discovery, replication, and functional assessment of genetic variants in the relevant tissue for drug abuse, the human brain. This approach was only possible through broad sharing of data in the genetic research community. In our opinion, keys for further advances in the genetics of addiction include continuing to take on the valuable but complex tasks of leveraging shared data for much larger sample sizes, as exemplified with large-scale GWAS meta-analyses for psychiatric disorders (Cross-Disorder Group of the Psychiatric Genomics, 2013; Schizophrenia Working Group of the Psychiatric Genomics, 2014), and integrating the power of GWAS with functional studies to highlight potential biological processes underlying statistical associations of variants with phenotypes.

In summary, the rs9829896-C allele was associated with reduced risk of drug abuse, reduced *KAT2B* and *CREBBP* expression, and increased *OPRM1* expression, all specifically in AAs. Several lines of evidence suggest that the newly implicated *KAT2B* gene and its rs9829896 SNP have biological plausibility for influencing abuse of drugs, including prior implication for psychiatric traits and direct connections to members of the cAMP and dopamine pathways (*CREBBP* and *OPRM1*, respectively) that play important roles in drug addiction. Our study also provided support to prior GWAS findings for marijuana use disorder, (Agrawal et al., 2014) showing that the missense SNP rs9891146-T allele in the *C17orf58* gene was associated with increased risk of drug abuse in general, corroborating its

association primarily in EAs, and adding novel regulatory evidence that this SNP correlates with *C17orf58* expression. Altogether, our findings highlight *KAT2B* and *C17orf58* as genetic susceptibility factors for drug abuse and demonstrate the importance of considering their effects across ancestry groups.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Our GWAS of drug abuse, with its use of Urban Health Study participants, was supported by the National Institute of Drug Abuse (NIDA) grant numbers R33 DA027486 and R01 DA026141.

Study controls were drawn from the following six cohorts in the database of Genotypes and Phenotypes (dbGaP, <http://www.ncbi.nlm.nih.gov/gap>).

(1) Funding support for the Study of Addiction: Genetics and Environment (SAGE) was provided through the National Institutes of Health (NIH) Genes, Environment and Health Initiative (GEI) (U01 HG004422). SAGE is one of the GWAS funded as part of the Gene Environment Association Studies (GENEVA) under GEI. Assistance with phenotype harmonization and genotype cleaning, as well as with general study coordination, was provided by the GENEVA Coordinating Center (U01 HG004446). Assistance with data cleaning was provided by the National Center for Biotechnology Information (NCBI). Support for collection of datasets and samples was provided by the Collaborative Study on the Genetics of Alcoholism (COGA; U10 AA008401), the Collaborative Genetic Study of Nicotine Dependence (COGEND; P01 CA089392), and the Family Study of Cocaine Dependence (FSCD; R01 DA013423). Funding support for genotyping, which was performed at the Johns Hopkins University Center for Inherited Disease Research (CIDR), was provided by the NIH GEI (U01HG004438), the National Institute on Alcohol Abuse and Alcoholism, NIDA, and the NIH contract “High throughput genotyping for studying the genetic contributions to human disease” (HHSN268200782096C). The datasets used for the analyses described in this manuscript were obtained via dbGaP accession number phs000092.v1.p1.

(2) The authors acknowledge the contribution of data from “Genetic Architecture of Smoking and Smoking Cessation” accessed through dbGAP accession number phs000404.v1.p1. Funding support for genotyping, which was performed at CIDR, was provided by 1 ×01 HG005274–01. CIDR is fully funded through a federal contract from the NIH to The Johns Hopkins University, contract number HHSN268200782096C. Assistance with genotype cleaning, as well as with general study coordination, was provided by the GENEVA Coordinating Center (U01 HG004446). Funding support for collection of datasets and samples was provided by COGEND (P01 CA089392) and the University of Wisconsin Transdisciplinary Tobacco Use Research Center (P50 DA019706, P50 CA084724).

(3) Funding support for the GWAS of Ischemic Stroke study was provided through the NIH GEI (U01HG004436). The GWAS of Ischemic Stroke study is one of the GWAS funded as part of GENEVA under GEI. Assistance with phenotype harmonization and genotype cleaning, as well as with general study coordination, was provided by the GENEVA Coordinating Center (U01 HG004446). Assistance with data cleaning was provided by the NCBI. Funding support for genotyping, which was performed at The Johns Hopkins University CIDR, was provided by the NIH GEI (U01HG004438) and the NIH contract “High throughput genotyping for studying the genetic contributions to human disease” (HHSN268200782096C). Field work for this project was supported by a Cooperative Agreement with the Division of Adult and Community Health, Centers for Disease Control and Prevention; the National Institute of Neurological Disorders and Stroke (NINDS) and the NIH Office of Research on Women’s Health (ORWH) (R01 NS45012); Office of Research and Development, Medical Research Service, Department of Veterans Affairs; and the University of Maryland General Clinical Research Center (M01 RR 165001), National Center for Research Resources, NIH. This study used samples from the NINDS Human Genetics Resource Center DNA and Cell Line Repository (<http://ccr.coriell.org/ninds>). The datasets used for the analyses described in this manuscript were obtained via dbGaP accession number phs000292.v1.p1.

(4) Funding support for the GENEVA Prostate Cancer study was provided through the National Cancer Institute (R37CA54281, R01CA63464, P01CA33619, U01CA136792, U01CA98758, and RC2 CA148085) and the National Human Genome Research Institute (U01HG004726). Assistance with phenotype harmonization, SNP selection, data cleaning, meta-analyses, data management and dissemination, and general study coordination, was provided by the GENEVA Coordinating Center (U01HG004789–01). The datasets used for the analyses described in this manuscript were obtained from dbGaP at phs000306.v2.p1.

(5) This work utilized in part data from the NINDS dbGaP database from the CIDR NeuroGenetics Research Consortium (NGRC) Parkinson's disease study (accession number phs000196.v2.p1).

(6) For the "High Density SNP Association Analysis of Melanoma: Case-Control and Outcomes Investigation," research support to collect data and develop an application to support this project was provided by P50CA093459, 5P50CA097007, 5R01ES011740, and 5R01CA133996 (dbGaP accession number phs000187.v1.p1).

The replication cohort with drug abuse cases and controls was also obtained from dbGaP. For the first replication cohort, funding support for the Genome-Wide Association of Schizophrenia Study was provided by the National Institute of Mental Health (NIMH) (R01 MH67257, R01 MH59588, R01 MH59571, R01 MH59565, R01 MH59587, R01 MH60870, R01 MH59566, R01 MH59586, R01 MH61675, R01 MH60879, R01 MH81800, U01 MH46276, U01 MH46289, U01 MH46318, U01 MH79469, and U01 MH79470) and the genotyping of samples was provided through the Genetic Association Information Network (GAIN). The datasets used for the analyses described in this manuscript were obtained from dbGaP through accession number phs000021.v3.p2. Samples and associated phenotype data for the Genome-Wide Association of Schizophrenia Study were provided by the Molecular Genetics of Schizophrenia Collaboration (PI: Pablo V. Gejman, Evanston Northwestern Healthcare (ENH) and Northwestern University, Evanston, IL, USA).

Lastly, the BrainCloud cohort was used to test SNP associations with mRNA expression levels. The BrainCloud genotype data were obtained from the database of Genotypes and Phenotypes (dbGaP, <http://www.ncbi.nlm.nih.gov/gap>) via accession number phs000417.v1.p1. Submission of the data to dbGaP was provided by Drs. Barbara Lipska and Joel Kleinman. Collection of the data was through a collaborative study sponsored by the NIMH Intramural Research Program. The gene expression data are available in the BrainCloud software application (downloadable at <http://braincloud.jhmi.edu/>) as well as the NCBI Gene Expression Omnibus (series number GSE30272).

References

- Agrawal A, Lynskey MT, Bucholz KK, Kapoor M, Almasy L, Dick DM, Edenberg HJ, Foroud T, Goate A, Hancock DB, Hartz S, Johnson EO, Hesselbrock V, Kramer JR, Kuperman S, Nurnberger JI Jr, Schuckit M, Bierut LJ. DSM-5 cannabis use disorder: a phenotypic and genomic perspective. Drug and alcohol dependence. 2014; 134:362–369. [PubMed: 24315570]
- Agrawal A, Verweij KJ, Gillespie NA, Heath AC, Lessov-Schlaggar CN, Martin NG, Nelson EC, Slutske WS, Whitfield JB, Lynskey MT. The genetics of addiction-a translational perspective. Transl Psychiatry. 2012; 2:e140. [PubMed: 22806211]
- Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. GenABEL: an R library for genome-wide association analysis. Bioinformatics. 2007; 23:1294–1296. [PubMed: 17384015]
- Colantuoni C, Lipska BK, Ye T, Hyde TM, Tao R, Leek JT, Colantuoni EA, Elkahoul AG, Herman MM, Weinberger DR, Kleinman JE. Temporal dynamics and genetic control of transcription in the human prefrontal cortex. Nature. 2011; 478:519–523. [PubMed: 22031444]
- Cross-Disorder Group of the Psychiatric Genomics C. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. Lancet. 2013; 381:1371–1379. [PubMed: 23453885]
- Durbin RM, Abecasis GR, Altshuler DL, Auton A, Brooks LD, Gibbs RA, Hurles ME, McVean GA. A map of human genome variation from population-scale sequencing. Nature. 2010; 467:1061–1073. [PubMed: 20981092]
- Ferrer-Alcon M, La Harpe R, Garcia-Sevilla JA. Decreased immunodensities of micro-opioid receptors, receptor kinases GRK 2/6 and beta-arrestin-2 in postmortem brains of opiate addicts. Brain Res Mol Brain Res. 2004; 121:114–122. [PubMed: 14969742]
- Frank J, Cichon S, Treutlein J, Ridinger M, Mattheisen M, Hoffmann P, Herms S, Wodarz N, Soyka M, Zill P, Maier W, Mossner R, Gaebel W, Dahmen N, Scherbaum N, Schmal C, Steffens M, Lucae S, Ising M, Muller-Myhsok B, Nothen MM, Mann K, Kiefer F, Rietschel M. Genome-wide significant association between alcohol dependence and a variant in the ADH gene cluster. Addiction biology. 2012; 17:171–180. [PubMed: 22004471]
- Garner C, Ahn R, Ding YC, Steele L, Stoven S, Green PH, Fasano A, Murray JA, Neuhausen SL. Genome-wide association study of celiac disease in North America confirms FRMD4B as new celiac locus. PloS one. 2014; 9:e101428. [PubMed: 24999842]
- Gelernter J, Kranzler HR, Sherva R, Almasy L, Koesterer R, Smith AH, Anton R, Preuss UW, Ridinger M, Rujescu D, Wodarz N, Zill P, Zhao H, Farrer LA. Genome-wide association study of

- alcohol dependence: significant findings in African- and European-Americans including novel risk loci. *Molecular psychiatry*. 2014a; 19:41–49. [PubMed: 24166409]
- Gelernter J, Kranzler HR, Sherva R, Koesterer R, Almasy L, Zhao H, Farrer LA. Genome-Wide Association Study of Opioid Dependence: Multiple Associations Mapped to Calcium and Potassium Pathways. *Biological psychiatry*. 2013
- Gelernter J, Sherva R, Koesterer R, Almasy L, Zhao H, Kranzler HR, Farrer L. Genome-wide association study of cocaine dependence and related traits: FAM53B identified as a risk gene. *Molecular psychiatry*. 2014b; 19:717–723. [PubMed: 23958962]
- Greenwood TA, Akiskal HS, Akiskal KK, Kelsoe JR. Genome-wide association study of temperament in bipolar disorder reveals significant associations with three novel Loci. *Biological psychiatry*. 2012; 72:303–310. [PubMed: 22365631]
- Gris P, Cheng P, Pierson J, Maixner W, Diatchenko L. Molecular assays for characterization of alternatively spliced isoforms of the u opioid receptor (MOR). *Methods Mol Biol*. 2010; 617:421–435. [PubMed: 20336438]
- Gruzca RA, Wang JC, Stitzel JA, Hinrichs AL, Saccone SF, Saccone NL, Bucholz KK, Cloninger CR, Neuman RJ, Budde JP, Fox L, Bertelsen S, Kramer J, Hesselbrock V, Tischfield J, Nurnberger JI Jr, Almasy L, Porjesz B, Kuperman S, Schuckit MA, Edenberg HJ, Rice JP, Goate AM, Bierut LJ. A risk allele for nicotine dependence in CHRNA5 is a protective allele for cocaine dependence. *Biological psychiatry*. 2008; 64:922–929. [PubMed: 18519132]
- Guo Y, Lanktree MB, Taylor KC, Hakonarson H, Lange LA, Keating BJ. Gene-centric meta-analyses of 108 912 individuals confirm known body mass index loci and reveal three novel signals. *Hum Mol Genet*. 2013; 22:184–201. [PubMed: 23001569]
- Ho LA, Lange EM. Using public control genotype data to increase power and decrease cost of case-control genetic association studies. *Hum Genet*. 2010; 128:597–608. [PubMed: 20821337]
- Howie B, Marchini J, Stephens M. Genotype imputation with thousands of genomes. *G3 (Bethesda)*. 2011; 1:457–470. [PubMed: 22384356]
- Johnson EO, Hancock DB, Levy JL, Gaddis NC, Saccone NL, Bierut LJ, Page GP. Imputation across genotyping arrays for genome-wide association studies: assessment of bias and a correction strategy. *Hum Genet*. 2013; 132:509–522. [PubMed: 23334152]
- Kaun KR, Azanchi R, Maung Z, Hirsh J, Heberlein U. A Drosophila model for alcohol reward. *Nature neuroscience*. 2011; 14:612–619. [PubMed: 21499254]
- Keebler ME, Sanders CL, Surti A, Guiducci C, Burt NP, Kathiresan S. Association of blood lipids with common DNA sequence variants at 19 genetic loci in the multiethnic United States National Health and Nutrition Examination Survey III. *Circ Cardiovasc Genet*. 2009; 2:238–243. [PubMed: 20031591]
- Kendler KS, Jacobson KC, Prescott CA, Neale MC. Specificity of genetic and environmental risk factors for use and abuse/dependence of cannabis, cocaine, hallucinogens, sedatives, stimulants, and opiates in male twins. *Am J Psychiatry*. 2003; 160:687–695. [PubMed: 12668357]
- Kral AH, Lorvick J, Gee L, Bacchetti P, Rawal B, Busch M, Edlin BR. Trends in human immunodeficiency virus seroincidence among street-recruited injection drug users in San Francisco, 1987–1998. *Am J Epidemiol*. 2003; 157:915–922. [PubMed: 12746244]
- Lansky A, Finlayson T, Johnson C, Holtzman D, Wejnert C, Mitsch A, Gust D, Chen R, Mizuno Y, Crepaz N. Estimating the number of persons who inject drugs in the united states by meta-analysis to calculate national rates of HIV and hepatitis C virus infections. *PloS one*. 2014; 9:e97596. [PubMed: 24840662]
- Lipska BK, Deep-Soboslay A, Weickert CS, Hyde TM, Martin CE, Herman MM, Kleinman JE. Critical factors in gene expression in postmortem human brain: Focus on studies in schizophrenia. *Biological psychiatry*. 2006; 60:650–658. [PubMed: 16997002]
- Liu Z, Guo X, Jiang Y, Zhang H. NCK2 is significantly associated with opiates addiction in African-origin men. *ScientificWorldJournal*. 2013; 2013:748979. [PubMed: 23533358]
- Luca D, Ringquist S, Klei L, Lee AB, Gieger C, Wichmann HE, Schreiber S, Krawczak M, Lu Y, Styche A, Devlin B, Roeder K, Trucco M. On the use of general control samples for genome-wide association studies: genetic matching highlights causal variants. *Am J Hum Genet*. 2008; 82:453–463. [PubMed: 18252225]

- Malvaez M, Mhillaj E, Matheos DP, Palmery M, Wood MA. CBP in the nucleus accumbens regulates cocaine-induced histone acetylation and is critical for cocaine-associated behaviors. *J Neurosci*. 2011; 31:16941–16948. [PubMed: 22114264]
- Morrall AR, McCaffrey D, Iguchi MY. Hardcore drug users claim to be occasional users: drug use frequency underreporting. *Drug and alcohol dependence*. 2000; 57:193–202. [PubMed: 10661670]
- Mukherjee S, Simon J, Bayuga S, Ludwig E, Yoo S, Orlov I, Viale A, Offit K, Kurtz RC, Olson SH, Klein RJ. Including additional controls from public databases improves the power of a genome-wide association study. *Hum Hered*. 2011; 72:21–34. [PubMed: 21849791]
- Namjou B, Kim-Howard X, Sun C, Adler A, Chung SA, Kaufman KM, Kelly JA, Glenn SB, Guthridge JM, Scofield RH, Kimberly RP, Brown EE, Alarcon GS, Edberg JC, Kim JH, Choi J, Ramsey-Goldman R, Petri MA, Reveille JD, Vila LM, Boackle SA, Freedman BI, Tsao BP, Langefeld CD, Vyse TJ, Jacob CO, Pons-Estel B, Niewold TB, Moser Sivils KL, Merrill JT, Anaya JM, Gilkeson GS, Gaffney PM, Bae SC, Alarcon-Riquelme ME, Harley JB, Criswell LA, James JA, Nath SK. PTPN22 Association in Systemic Lupus Erythematosus (SLE) with Respect to Individual Ancestry and Clinical Sub-Phenotypes. *PloS one*. 2013; 8:e69404. [PubMed: 23950893]
- Nestler EJ. Molecular neurobiology of addiction. *The American journal on addictions / American Academy of Psychiatrists in Alcoholism and Addictions*. 2001; 10:201–217. [PubMed: 11579619]
- Nielsen DA, Ji F, Yuferov V, Ho A, He C, Ott J, Kreek MJ. Genome-wide association study identifies genes that may contribute to risk for developing heroin addiction. *Psychiatr Genet*. 2010; 20:207–214. [PubMed: 20520587]
- Oertel BG, Doebling A, Roskam B, Kettner M, Hackmann N, Ferreiros N, Schmidt PH, Lotsch J. Genetic-epigenetic interaction modulates mu-opioid receptor regulation. *Hum Mol Genet*. 2012; 21:4751–4760. [PubMed: 22875838]
- Park BL, Kim JW, Cheong HS, Kim LH, Lee BC, Seo CH, Kang TC, Nam YW, Kim GB, Shin HD, Choi IG. Extended genetic effects of ADH cluster genes on the risk of alcohol dependence: from GWAS to replication. *Hum Genet*. 2013; 132:657–668. [PubMed: 23456092]
- Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics*. 2006; 38:904–909. [PubMed: 16862161]
- Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer CJ. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics*. 2010; 26:2336–2337. [PubMed: 20634204]
- Quillen EE, Chen XD, Almasy L, Yang F, He H, Li X, Wang XY, Liu TQ, Hao W, Deng HW, Kranzler HR, Gelernter J. ALDH2 is associated to alcohol dependence and is the major genetic determinant of “daily maximum drinks” in a GWAS study of an isolated rural chinese sample. *American journal of medical genetics Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics*. 2014; 165B:103–110.
- Ravnskjaer K, Hogan MF, Lackey D, Tora L, Dent SY, Olefsky J, Montminy M. Glucagon regulates gluconeogenesis through KAT2B- and WDR5-mediated epigenetic effects. *J Clin Invest*. 2013
- Rhodes W, Layne M, Johnston P, Hozik L. What America's Users Spend on Illegal Drugs. 2000:1988–1998. Policy OoNDC (ed).
- Ruderfer DM, Fanous AH, Ripke S, McQuillin A, Amdur RL, Schizophrenia Working Group of the Psychiatric Genomics C; Bipolar Disorder Working Group of the Psychiatric Genomics C; Cross-Disorder Working Group of the Psychiatric Genomics C. Gejman PV, O'Donovan MC, Andreassen OA, Djurovic S, Hultman CM, Kelsoe JR, Jamain S, Landen M, Leboyer M, Nimgaonkar V, Nurnberger J, Smoller JW, Craddock N, Corvin A, Sullivan PF, Holmans P, Sklar P, Kendler KS. Polygenic dissection of diagnosis and clinical dimensions of bipolar disorder and schizophrenia. *Molecular psychiatry*. 2013
- Saccone SF, Hinrichs AL, Saccone NL, Chase GA, Konvicka K, Madden PA, Breslau N, Johnson EO, Hatsukami D, Pomerleau O, Swan GE, Goate AM, Rutter J, Bertelsen S, Fox L, Fugman D, Martin NG, Montgomery GW, Wang JC, Ballinger DG, Rice JP, Bierut LJ. Cholinergic nicotinic receptor genes implicated in a nicotine dependence association study targeting 348 candidate genes with 3713 SNPs. *Hum Mol Genet*. 2007; 16:36–49. [PubMed: 17135278]

- Schizophrenia Working Group of the Psychiatric Genomics C. Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. 2014; 511:421–427. [PubMed: 25056061]
- Sherva R, Kranzler HR, Yu Y, Logue MW, Poling J, Arias AJ, Anton RF, Oslin D, Farrer LA, Gelernter J. Variation in nicotinic acetylcholine receptor genes is associated with multiple substance dependence phenotypes. *Neuropsychopharmacology*. 2010; 35:1921–1931. [PubMed: 20485328]
- Shi J, Levinson DF, Duan J, Sanders AR, Zheng Y, Pe'er I, Dudbridge F, Holmans PA, Whittemore AS, Mowry BJ, Olincy A, Amin F, Cloninger CR, Silverman JM, Buccola NG, Byerley WF, Black DW, Crowe RR, Oksenberg JR, Mirel DB, Kendler KS, Freedman R, Gejman PV. Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature*. 2009; 460:753–757. [PubMed: 19571809]
- Sullivan SE, Whittard JD, Jacobs MM, Ren Y, Mazloom AR, Caputi FF, Horvath M, Keller E, Ma'ayan A, Pan YX, Chiang LW, Hurd YL. ELK1 transcription factor linked to dysregulated striatal mu opioid receptor signaling network and OPRM1 polymorphism in human heroin abusers. *Biological psychiatry*. 2013; 74:511–519. [PubMed: 23702428]
- Silverberg MS, Cho JH, Rioux JD, McGovern DP, Wu J, Annese V, Achkar JP, Goyette P, Scott R, Xu W, Barmada MM, Klei L, Daly MJ, Abraham C, Bayless TM, Bossa F, Griffiths AM, Ippoliti AF, Lahaie RG, Latiano A, Pare P, Proctor DD, Regueiro MD, Steinhart AH, Targan SR, Schumm LP, Kistner EO, Lee AT, Gregersen PK, Rotter JI, Brant SR, Taylor KD, Roeder K, Duerr RH. Ulcerative colitis-risk loci on chromosomes 1p36 and 12q15 found by genome-wide association study. *Nature genetics*. 2009; 41:216–220. [PubMed: 19122664]
- Steinberg S, de Jong S, Mattheisen M, Costas J, Demontis D, Jamain S, Pietilainen OPH, Lin K, Papiol S, Huttenlocher J, Sigurdsson E, Vassos E, Giegling I, Breuer R, Fraser G, Walker N, Melle I, Djurovic S, Agartz I, Tuulio-Henriksson A, Suvisaari J, Lonnqvist J, Paunio T, Olsen L, Hansen T, Ingason A, Pirinen M, Strengman E, Hougaard DM, Orntoft T, Didriksen M, Hollegaard MV, Nordentoft M, Abramova L, Kaleda V, Arrojo M, Sanjuan J, Arango C, Etain B, Bellivier F, Meary A, Schurhoff F, Szoke A, Ribolsi M, Magni V, Siracusano A, Sperling S, Rossner M, Christiansen C, Kiemeny LA, Franke B, van den Berg LH, Veldink J, Curran S, Bolton P, Poot M, Staal W, Rehnstrom K, Kilpinen H, Freitag CM, Meyer J, Magnusson P, Saemundsen E, Martsenkovsky I, Bikshaieva I, Martsenkovska I, Vashchenko O, Raleva M, Paketchieva K, Stefanovski B, Durmishi N, Milovancevic MP, Tosevski DL, Silagadze T, Naneishvili N, Mikeladze N, Surguladze S, Vincent JB, Farmer A, Mitchell PB, Wright A, Schofield PR, Fullerton JM, Montgomery GW, Martin NG, Rubino IA, van Winkel R, Kenis G, De Hert M, Rethelyi JM, Bitter I, Terenius L, Jonsson EG, Bakker S, van Os J, Jablensky A, Leboyer M, Bramon E, Powell J, Murray R, Corvin A, Gill M, Morris D, O'Neill FA, Kendler K, Riley B, Craddock N, Owen MJ, O'Donovan MC, Thorsteinsdottir U, Kong A, Ehrenreich H, Carracedo A, Golimbet V, Andreassen OA, Borglum AD, Mors O, Mortensen PB, Werge T, Ophoff RA, Nothen MM, Rietschel M, Cichon S, Ruggeri M, Tosato S, Palotie A, St Clair D, Rujescu D, Collier DA, Stefansson H, Stefansson K, GROUP, Consor WTCC. Common variant at 16p11.2 conferring risk of psychosis. *Molecular psychiatry*. 2014; 19:108–114. [PubMed: 23164818]
- Stranger BE, Montgomery SB, Dimas AS, Parts L, Stegle O, Ingle CE, Sekowska M, Smith GD, Evans D, Gutierrez-Arcelus M, Price A, Raj T, Nisbett J, Nica AC, Beazley C, Durbin R, Deloukas P, Dermitzakis ET. Patterns of cis regulatory variation in diverse human populations. *PLoS Genet*. 2012; 8:e1002639. [PubMed: 22532805]
- Thorgeirsson TE, Gudbjartsson DF, Surakka I, Vink JM, Amin N, Geller F, Sulem P, Rafnar T, Esko T, Walter S, Gieger C, Rawal R, Mangino M, Prokopenko I, Magi R, Keskitalo K, Gudjonsdottir IH, Gretarsdottir S, Stefansson H, Thompson JR, Aulchenko YS, Nelis M, Aben KK, den Heijer M, Dirksen A, Ashraf H, Soranzo N, Valdes AM, Steves C, Uitterlinden AG, Hofman A, Tonjes A, Kovacs P, Hottenga JJ, Willemsen G, Vogelzangs N, Doring A, Dahmen N, Nitz B, Pergadia ML, Saez B, De Diego V, Lezcano V, Garcia-Prats MD, Ripatti S, Perola M, Kettunen J, Hartikainen AL, Pouta A, Laitinen J, Isohanni M, Huei-Yi S, Allen M, Krestyaninova M, Hall AS, Jones GT, van Rij AM, Mueller T, Dieplinger B, Haltmayer M, Jonsson S, Matthiasson SE, Oskarsson H, Tyrifingsson T, Kiemeny LA, Mayordomo JI, Lindholt JS, Pedersen JH, Franklin WA, Wolf H, Montgomery GW, Heath AC, Martin NG, Madden PA, Giegling I, Rujescu D, Jarvelin MR, Salomaa V, Stumvoll M, Spector TD, Wichmann HE, Metspalu A, Samani NJ, Penninx BW, Oostra BA, Boomsma DI, Tiemeier H, van Duijn CM, Kaprio J, Gulcher JR, McCarthy MI,

- Peltonen L, Thorsteinsdottir U, Stefansson K. Sequence variants at CHRNA3-CHRNA6 and CYP2A6 affect smoking behavior. *Nature genetics*. 2010; 42:448–453. [PubMed: 20418888]
- Torgerson DG, Ampleford EJ, Chiu GY, Gauderman WJ, Gignoux CR, Graves PE, Himes BE, Levin AM, Mathias RA, Hancock DB, Baurley J, Eng C, Stern DA, Celedon JC, Rafaels N, Capurso D, Conti DV, Roth LA, Soto-Quiros M, Torgias A, Li X, Myers RA, Romieu I, van den Berg DJ, Hu D, Hansel NN, Abel RA, Israel E, Salam MT, Galanter J, Avila PC, Avila L, Rodriguez-Santana JR, Chapela R, Rodriguez-Cintron W, Diette GB, Adkinson NF, Ross KD, Shi M, Faruque MU, Dunston GM, Watson HR, Mantese VJ, Ezurum SC, Liang L, Ruczinski I, Ford JG, Huntsman S, Chung KF, Vora H, Li X, Calhoun WJ, Castro M, Sierra-Monge JJ, del Rio-Navarro B, Deichmann KA, Heinzmann A, Wenzel SE, Busse WW, Gern JE, Lemanske RF, Beaty TH, Bleecker ER, Raby BA, Meyers DA, London SJ, Gilliland FD, Burchard EG, Martinez FD, Weiss ST, Williams LK, Barnes KC, Ober C, Nicolae DL. Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations. *Nature genetics*. 2011; 43:887–892. [PubMed: 21804549]
- Tsuang MT, Bar JL, Harley RM, Lyons MJ. The Harvard Twin Study of Substance Abuse: what we have learned. *Harv Rev Psychiatry*. 2001; 9:267–279. [PubMed: 11600486]
- Uhl GR, Drgon T, Liu QR, Johnson C, Walther D, Komiyama T, Harano M, Sekine Y, Inada T, Ozaki N, Iyo M, Iwata N, Yamada M, Sora I, Chen CK, Liu HC, Ujike H, Lin SK. Genome-wide association for methamphetamine dependence: convergent results from 2 samples. *Archives of general psychiatry*. 2008; 65:345–355. [PubMed: 18316681]
- Verweij KJ, Zietsch BP, Lynskey MT, Medland SE, Neale MC, Martin NG, Boomsma DI, Vink JM. Genetic and environmental influences on cannabis use initiation and problematic use: a meta-analysis of twin studies. *Addiction*. 2010; 105:417–430. [PubMed: 20402985]
- Wang JC, Kapoor M, Goate AM. The genetics of substance dependence. *Annu Rev Genomics Hum Genet*. 2012; 13:241–261. [PubMed: 22703173]
- Wang KS, Liu XF, Aragam N. A genome-wide meta-analysis identifies novel loci associated with schizophrenia and bipolar disorder. *Schizophr Res*. 2010; 124:192–199. [PubMed: 20889312]
- Wellcome Trust Case Control C. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007; 447:661–678. [PubMed: 17554300]
- Wetherill L, Agrawal A, Kapoor M, Bertelsen S, Bierut LJ, Brooks A, Dick D, Hesselbrock M, Hesselbrock V, Koller DL, Le N, Nurnberger JI Jr, Salvatore JE, Schuckit M, Tischfield JA, Wang JC, Xuei X, Edenberg HJ, Porjesz B, Bucholz K, Goate AM, Foroud T. Association of substance dependence phenotypes in the COGA sample. *Addiction biology*. 2014
- Waller CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010; 26:2190–2191. [PubMed: 20616382]
- Wu L, Aster JC, Blacklow SC, Lake R, Artavanis-Tsakonas S, Griffin JD. MAML1, a human homologue of *Drosophila* mastermind, is a transcriptional co-activator for NOTCH receptors. *Nature genetics*. 2000; 26:484–489. [PubMed: 11101851]
- Yu D, Mathews CA, Scharf JM, Neale BM, Davis LK, Gamazon ER, Derks EM, Evans P, Edlund CK, Crane J, Fageress JA, Osiecki L, Gallagher P, Gerber G, Haddad S, Illmann C, McGrath LM, Mayerfeld C, Arepalli S, Barlassina C, Barr CL, Bellodi L, Benarroch F, Berrio GB, Bienvenu OJ, Black DW, Bloch MH, Brentani H, Bruun RD, Budman CL, Camarena B, Campbell DD, Cappi C, Silgado JC, Cavallini MC, Chavira DA, Chouinard S, Cook EH, Cookson MR, Coric V, Cullen B, Cusi D, Delorme R, Denys D, Dion Y, Eapen V, Egberts K, Falkai P, Fernandez T, Fournier E, Garrido H, Geller D, Gilbert D, Girard SL, Grabe HJ, Grados MA, Greenberg BD, Gross-Tsur V, Grunblatt E, Hardy J, Heiman GA, Hemmings SM, Herrera LD, Hezel DM, Hoekstra PJ, Jankovic J, Kennedy JL, King RA, Konkashbaev AI, Kremeyer B, Kurlan R, Lanzagorta N, Leboyer M, Leckman JF, Lennertz L, Liu C, Lochner C, Lowe TL, Lupoli S, Macciardi F, Maier W, Manunta P, Marconi M, McCracken JT, Mesa Restrepo SC, Moessner R, Moorjani P, Morgan J, Muller H, Murphy DL, Naarden AL, Nurmi E, Ochoa WC, Ophoff RA, Pakstis AJ, Pato MT, Pato CN, Piacentini J, Pittenger C, Pollak Y, Rauch SL, Renner T, Reus VI, Richter MA, Riddle MA, Robertson MM, Romero R, Rosario MC, Rosenberg D, Ruhrmann S, Sabatti C, Salvi E, Sampaio AS, Samuels J, Sandor P, Service SK, Sheppard B, Singer HS, Smit JH, Stein DJ, Strengman E, Tischfield JA, Turiel M, Valencia Duarte AV, Vallada H, Veenstra-VanderWeele J, Walitza S, Wang Y, Weale M, Weiss R, Wendland JR, Westenberg HG, Shugart YY, Hounie AG, Miguel EC,

Nicolini H, Wagner M, Ruiz-Linares A, Cath DC, McMahon W, Posthuma D, Oostra BA, Nestadt G, Rouleau GA, Purcell S, Jenike MA, Heutink P, Hanna GL, Conti DV, Arnold PD, Freimer NB, Stewart SE, Knowles JA, Cox NJ, Pauls DL. Cross-Disorder Genome-Wide Analyses Suggest a Complex Genetic Relationship Between Tourette's Syndrome and OCD. *The American journal of psychiatry*. 2014

Zhang L, Loh HH, Law PY. A Novel Noncanonical Signaling Pathway for the mu-Opioid Receptor. *Mol Pharmacol*. 2013; 84:844–853. [PubMed: 24061856]

Zhuang JJ, Zondervan K, Nyberg F, Harbron C, Jawaid A, Cardon LR, Barratt BJ, Morris AP. Optimizing the power of genome-wide association studies by using publicly available reference samples to expand the control group. *Genet Epidemiol*. 2010; 34:319–326. [PubMed: 20088020]

Zuberi K, Franz M, Rodriguez H, Montojo J, Lopes CT, Bader GD, Morris Q. GeneMANIA prediction server 2013 update. *Nucleic acids research*. 2013; 41:W115–122. [PubMed: 23794635]

Figure 1.
Pipeline for generating a control data set for comparison to drug abuse cases from the Urban Health Study.

Figure 2.

Gene expression in post-mortem human prefrontal cortex samples from BrainCloud, as stratified by rs9829896 genotype. Gene expression profiles for (A) *KAT2B*, (B) *CREBBP*, and (C) *OPRM1* are presented separately for 113 African Americans (AA) and 110 European Americans (denoted CAUC). Genotypes are designated as '1' for the major allele homozygotes (AA), '2' for heterozygotes, and '3' for minor allele homozygotes (CC).

Figure 3.

Drugs reported by cases in the Urban Health Study (UHS) and the two independent replication cohorts. The Genetic Association Information Network (GAIN) was ascertained for schizophrenia, and only the original study controls were included in our replication analyses for drug abuse.

Table 1

Cohort	Case definition	Ancestry group	Total no controls	Total no cases	No. (%) cases reporting past 30 day abuse of specific drug category ^a					No. (%) Cases Reporting >1 specific drug category
					Heroin and other opioids	Cocaine	Marijuana	Amphetamines and other stimulants	Other	
UHS	Use of drugs 10 or more times in the past 30 days	AAs	1,725 ^a	2,017	1,294 (64.2)	870 (43.1)	35 (1.7)	113 (5.6)	390 (19.3)	645 (32.0)
		EAs	5,703 ^a	1,142	711 (62.3)	285 (25.0)	6 (0.5)	213 (18.7)	233 (20.4)	281 (24.6)
GAIN	Ever use of at least one drug and report of at least one DSM-IV symptom of dependence	EAs	858	273	177 (64.8)	121 (44.3)	230 (84.2)	136 (49.8)	188 (68.9)	202 (74.0)
		AAs	496	259	113 (43.6)	112 (43.2)	212 (81.8)	48 (18.5)	125 (48.2)	179 (69.1)

^a Specific drug categories were not mutually exclusive, so multidrug abuse cases were represented more than once.

Table 2

Associations of the rs9829896 minor allele (C) with drug abuse in African Americans and European Americans from the Urban Health Study (UHS) and the Genetic Association Information Network (GAIN) cohorts. Rs9829896 was genotyped in the UHS cohort and imputed with high quality in the GAIN control.

Ancestry Group	Cohort	N, cases	N, controls	MAF	info ^a	P	OR (95% CI)	Meta-analysis ^b
African Americans	UHS ^b	2,017	1,725	0.12	1.0	4.63×10 ⁻⁸	0.68 (0.59–0.78)	3.93×10 ⁻¹⁰
	GAIN ^c	259	496	0.12	0.99	0.0019	0.53 (0.35–0.79)	
European Americans	UHS ^b	1,142	5,703	0.36	1.0	0.17	0.93 (0.84–1.04)	0.12
	GAIN ^c	273	858	0.34	1.0	0.43	0.91 (0.73–1.15)	

CI, confidence interval; MAF, minor allele frequency; OR, odds ratio;

^aThe info metric, as generated from IMPUTE2, is a fractional value ranging from 0 and 1 to indicate the degree of imputation certainty. Genotyped SNPs and imputed SNPs with the highest quality are reported with info=1.

^bModels were adjusted for sex and eigenvectors.

^cModels were adjusted for sex, eigenvectors, alcohol abuse, and nicotine dependence.

Table 3

Associations of the rs9829896 minor allele (C) with specific drugs, using separate logistic regression models, in African Americans from the Urban Health Study (UHS) and the Genetic Association Information Network (GAIN).

Specific drug	UHS			GAIN					Meta analysis ^{<i>P</i>}
	N, cases	N, controls	<i>P</i>	OR (95% CI) ^{<i>a</i>}	N, cases	N, controls	<i>P</i>	OR (95% CI) ^{<i>b</i>}	
Heroin and other opioids	1,294	1,725	1.05×10 ⁻⁵	0.70 (0.60–0.82)	113	496	0.16	0.70 (0.43–1.16)	4.48×10 ⁻⁶
Cocaine	870	1,725	1.50×10 ⁻⁴	0.71 (0.59–0.85)	112	496	0.13	0.67 (0.40–1.12)	4.58×10 ⁻⁵
Marijuana	35	1,725	0.78	0.91 (0.46–1.81)	212	496	0.0010	0.47 (0.30–0.74)	0.046
Amphetamines and other stimulants	113	1,725	0.12	0.71 (0.45–1.10)	48	496	0.83	0.83 (0.43–1.63)	1.16×10 ⁻⁵

CI, confidence interval; OR, odds ratio

^aModels were adjusted for sex and eigenvectors.

^bModels were adjusted for sex, eigenvectors, alcohol abuse and nicotine dependence.