npg

Variation in Nicotinic Acetylcholine Receptor Genes is Associated with Multiple Substance Dependence Phenotypes

Richard Sherva[1], Henry R Kranzler[2], Yi Yu[1], Mark W Logue[1], James Poling[3,4], Albert J Arias[2], Raymond F Anton[5], David Oslin[6], Lindsay A Farrer[1,7] and Joel Gelernter[3,4,8]

[1]Department of Medicine (Genetics Program), Boston University School of Medicine, Boston, MA, USA; [2]Departments of Psychiatry and Genetics and Developmental Biology, University of Connecticut School of Medicine, Farmington, CT, USA; [3]Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA; [4]VA CT Healthcare System, West Haven, CT, USA; [5]Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina, Charleston, SC, USA; [6]Department of Psychiatry, University of Pennsylvania and Philadelphia Veterans Affairs Medical Center, Philadelphia, PA, USA; [7]Departments of Neurology, Genetics & Genomics, Biostatistics, and Epidemiology, Boston University Schools of Medicine and Public Health, Boston, MA, USA; [8]Departments of Neurobiology and Genetics, Yale University School of Medicine, New Haven, CT, USA

There is shared genetic risk for dependence on multiple substances, and the nicotinic receptor gene cluster on chromosome 15 harbors multiple polymorphisms that associate to this risk. Here, we report the results of an association study with 21 SNPs genotyped across the CHRNA5, CHRNA3, and CHRNB4 loci on chromosome 15q25.1. The sample consists of a discovery set ($N = 1858$) of European-American and African-American (AA) families, ascertained on the basis of a sibling pair with cocaine and/or opioid dependence, and a case–control replication sample ($N = 3388$) collected for association studies of alcohol, cocaine, and opioid dependence. We tested the SNPs for association with lifetime cocaine, opioid, nicotine, and alcohol dependence. We replicated several previous findings, including associations between rs16969968 and nicotine dependence ($P = 0.002$) and cocaine dependence ($P = 0.02$), with opposite risk alleles for each substance. We observed these associations in AAs, which is a novel finding. The strongest association signal in either sample was between rs684513 in CHRNA5 and cocaine dependence (OR = 1.43, $P = 0.0004$) in the AA replication set. We also observed two SNPs associated with alcohol dependence, that is, rs615470 in CHRNA5 (OR = 0.77, $P = 0.0006$) and rs578776 (OR = 0.78, $P = 0.001$). The associations between CD and rs684513, AD and rs615470, and AD and rs578776 remained significant after a permutation-based correction for multiple testing. These data reinforce the importance of variation in the chromosome 15 nicotinic receptor subunit gene cluster for risk of dependence on multiple substances, although the direction of the effects may vary across substances.
Neuropsychopharmacology (2010) 35, 1921–1931; doi:10.1038/npp.2010.64; published online 19 May 2010

Keywords: genetic association; substance dependence; nicotinic acetylcholine receptors

INTRODUCTION

Dependence on legal and illicit drugs is a major public health problem. Heavy use of alcohol and nicotine represent the greatest causes of preventable morbidity and mortality in Western societies (Edwards, 2004). Illicit drug use also causes substantial morbidity and mortality. Misusers of cocaine and opioids have six times the mortality rate of an age-matched population, with the highest mortality rate seen in people who combine illicit drugs and alcohol (Gossop et al, 2002).

Correspondence: Dr J Gelernter, Yale University School of Medicine, VA CT Healthcare Center 116A2, 950 Campbell Avenue, West Haven, CT 06516, USA, Tel: + 1 203 932 5711 ext 3599, Fax: + 1 203 937 3897, E-mail: joel.gelernter@yale.edu

Susceptibility to abuse and/or dependence on alcohol, nicotine, and illicit drugs all have genetic components (Williams et al, 1999; Long et al, 1998; Carmelli et al, 1992; Heath and Martin 1993; Tsuang et al, 1996; Kendler et al, 2000). Although there are substance-specific factors, much of the genetic vulnerability to abuse of different substances is shared (Kendler and Prescott, 1998; Tsuang et al, 1998; Kendler et al, 1999; Karkowski et al, 2000; Bierut et al, 1998). Substantial research on the genetics of addiction has focused on reward pathways in the brain, especially those mediated by the biogenic amines acetylcholine (ACh), dopamine (DA), and serotonin, including their transporters and receptors. Here, we report on associations with nicotine, alcohol, cocaine, and opioid dependence (ND, AD, CD, and OD, respectively) in large samples of European-American (EA) and African-American (AA) families genotyped for 21 SNPs in three nicotinic ACh receptor genes (CHRNA5, CHRNA3, and CHRNB4) located

in tandem on chromosome 15. The receptor subunits encoded by these genes mediate some of the physiological effects of nicotine (Benowitz, 1996). We also tested these SNPs for association with the same phenotypes in a replication sample of unrelated EA and AA cases and controls.

Most of the SNPs genotyped here have been previously tested for association with substance abuse or dependence phenotypes. One study identified a SNP in *CHRNA5* (rs16969968), whose minor alleles seem to protect against ND in subjects who, during their lives, had smoked at least 100 cigarettes (Bierut *et al*, 2007; Saccone *et al*, 2007). This finding was replicated in independent samples (Sherva *et al*, 2008; Bierut *et al*, 2008), and was associated with experiencing a pleasurable 'buzz' on smoking initiation (Sherva *et al*, 2008). Another study showed that *CHRNA5* knockout mice had a reduced somatic response to nicotine withdrawal (Jackson *et al*, 2008). A functional study showed that this polymorphism decreases the response of the receptor cluster to a nicotine agonist (Bierut *et al*, 2008). This SNP also seems to be associated with CD (Grucza *et al*, 2008) and AD (Wang *et al*, 2008). SNP rs578776 in the *CHRNA5/CHRNA3/CHRNB4* cluster also seems to affect ND risk (Bierut *et al*, 2008). In a different study, two other SNPs, rs1948 and rs8023462, were associated with the age of initiation for both tobacco and alcohol use (Schlaepfer *et al*, 2008). A SNP (rs1051730) in *CHRNA3* was associated with smoking quantity (Thorgeirsson *et al*, 2008) in a European population. This SNP was recently shown to be associated with risk of continued smoking during pregnancy (Freathy *et al*, 2009). Another study recently identified variants in *CHRNB3* that were associated with 'dizziness' after smoking (Ehringer *et al*, 2009). Subsequently, these and other nicotinic receptors were shown to affect cancer risk through a process independent of their effect on smoking behavior (Hung *et al*, 2008; Thorgeirsson *et al*, 2008; Amos *et al*, 2008).

The fact that the effect of many variants differs across population groups further complicates the search for the genetic influences on addiction. Although several regions common to AA and EA groups have been identified through linkage (Li *et al*, 2008), evidence suggests that certain variants exert a stronger effect in certain populations (Saccone *et al*, 2009). This is consistent with expectations that different populations may have different risk alleles, different LD relationships between markers and functional variants, and different effects of alleles that are directly functional based on differences in epistasis.

In addition, smoking topography differs across population groups, with AAs generally consuming more nicotine per cigarette than EAs (Perez-Stable *et al*, 1998). The strong LD in the region and the different LD patterns observed across populations further complicate the search for functional variants. Despite these obstacles, the combination of genetic association studies and molecular genetic experiments has provided insight into the functional aspects of how nicotinic receptors contribute to individual responses to substances and to the larger process of addiction.

METHODS

Recruitment and Ascertainment

The sample consisted of a discovery set of families and a replication set of unrelated cases and controls. Related individuals were recruited at four US clinical sites: the Yale University School of Medicine (APT Foundation; New Haven, CT; $N = 1473$), the University of Connecticut Health Center (Farmington, CT; $N = 1430$), McLean Hospital (Harvard Medical School, Belmont, MA; $N = 505$), the Medical University of South Carolina (Charleston, SC; $N = 409$), and the University of Pennsylvania (Philadelphia, PA; $N = 5$). The family sample contained 2129 AAs (including 132 self-reported Hispanics) and 1706 EAs (including 310 self-reported Hispanics). Of these, 1858 subjects from 893 families had genotype and phenotype data. Families were ascertained from treatment centers and advertisements that recruited affected sibling pairs (ASPs) meeting Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) criteria for CD or OD. Probands were excluded from further study if they had ever received a clinical diagnosis of a major psychotic illness (for example, schizophrenia or schizoaffective disorder). Other family members of the ASPs were recruited when available, regardless of affection status. The case–control subjects were recruited from substance abuse treatment centers and through advertisements at the University of Connecticut Health Center ($N = 1690$), Yale University School of Medicine ($N = 1471$), the Medical University of South Carolina ($N = 383$), the University of Pennsylvania ($N = 260$), and McLean Hospital ($N = 3$). The replication sample of cases and controls contained 1912 AAs (including 76 Hispanics) and 1476 EAs, (including 176 Hispanics). Table 1 shows the number of cases and controls for each substance by recruitment site. Subjects gave informed consent as approved by the institutional review board at each clinical site, and certificates of confidentiality were obtained from the National Institute on Drug Abuse and the National Institute on Alcohol Abuse and Alcoholism.

Table 1 Number of Cases and Controls Recruited by Study Site

Site	AD cases	AD controls	CD cases	CD controls	OD cases	OD controls	ND cases	ND controls
MUSC	216	228	230	310	72	469	208	344
YALE	1021	627	1444	543	740	1278	1309	745
UCONN	1136	664	1407	616	839	1217	1238	847
UPENN	148	77	255	4	24	230	166	93
MCLEAN	106	60	197	10	84	119	140	71

Abbreviations: AD, alcohol dependence; CD, cocaine dependence; ND, nicotine dependence; OD, opioid dependence.

Phenotypic Measures

All subjects were interviewed using an electronic version of the Semi-Structured Assessment for Drug Dependence and Alcoholism (Pierucci-Lagha et al, 2005) to derive diagnoses for lifetime ND, CD, OD, and AD according to DSM-IV criteria. The inter-rater reliability (κ (95% confidence interval)) for these diagnoses was 0.77 (0.67, 0.87), 0.83 (0.75, 0.91), 0.91 (0.79, 1.00), and 0.66 (0.54, 0.78), respectively, and the test–retest reliability was 0.97 (0.93, 1.00), 0.92 (0.84, 1.00), 0.94 (0.84, 1.00), and 0.87 (0.77, 0.97), respectively (Pierucci-Lagha et al, 2005).

Genotyping

SNP genotyping was performed at Yale University School of Medicine using a closed-tube fluorescent TaqMan 5′-nuclease allelic discrimination assay ordered as 'assays-on-demand' (Applied Biosystems, Foster City, CA). Fluorescence plate reads and genotype calls were made using an ABI 7900 Sequence Detection Systems. Two nanograms of genomic DNA was PCR amplified in 384-well plates using a 2 μl reaction volume. The insertion/deletion marker was genotyped by PCR amplification followed by agarose gel size fractionation. We designed two primer pairs such that two reactions could be multiplexed in each gel lane: rs3841324-f (5′-gagacaaaacgagggcagac-3′) and rs3841324-r (5′-aaaggaacaaggcgaggatt-3′) (product sizes 166 bp and 188 bp), and rs3841324-F-Long (5′-GGGAACGCGAACTCTTGG-3′) and rs3841324-R-Long (5′-GCTAGGAGCAGACAGGGTTG-3′) (300 and 322 bp). For both genotyping methods, 8% of genotypes were repeated for quality control; any mismatches triggered repeats of all genotypes on a given plate. Such duplicate genotyping was carried out in at least some plates for markers rs1051730 and rs16969968 and the mismatches were discarded. Our analysis included 21 markers in the CHRNA5/CHRNA3/CHRNB4 gene cluster on chromosome 15. A total of 13 SNPs were nominally associated with one or more of the DSM-IV dependence traits ($P \leqslant 0.05$) in the discovery set and were subsequently typed in the replication sample. All SNPs were in Hardy–Weinberg equilibrium ($P > 0.01$).

Statistical Analysis

Analysis was conducted separately in AAs and EAs based on self-reported race. To verify self-reported race, we used a Bayesian model-based clustering method implemented in the program STRUCTURE (Pritchard et al, 2000) to divide a subset ($N = 1619$) of the family sample into ancestry groups based on ~400 microsatellite markers spaced across the genome. This clustering produced two distinct groups that were highly concordant with self-reported AA and EA race. As genome-wide marker data was not available in the case–control sample, we used self-reported race to adjust for ancestry in order to maintain consistency in the analysis of each sample. Within the AA and EA groups, we included a covariate for self-reported Hispanic ancestry. Asians, Native Americans, and Pacific Islanders were excluded from all analyses owing to low numbers.

We analyzed association of the SNPs with the substance dependence traits in the family-based discovery data set using generalized estimating equation (GEE) models implemented in SAS v.9.1, which employ a standard error correction to account for the genotypic and phenotypic correlation among family members. These models also allow for the inclusion of covariates and produce effect size estimates in the form of yield odds ratios calculated by exponentiation of the β-coefficients. The case–control data were evaluated using logistic regression models implemented in PLINK v.1.05 (http://pngu.mgh.harvard.edu/purcell/plink/) (Purcell et al, 2007). We first analyzed each of the four dependence phenotypes without considering dependence on other substances, meaning that controls for one substance might meet the criteria for dependence on another substance. As this strategy could mask the effects of SNPs increasing the risk of addiction in general (although it could increase our ability to identify variants with a specific effect on the substance in question but not a general effect on risk), we also tested cases for each dependence phenotype against a common reference group that did not meet the criteria for dependence on any of the substances.

All models were adjusted for age, sex, and Hispanic ancestry. To address the potential for residual confounding due to population group, we also compared the primary results from models (1) that excluded Hispanic subjects, and (2) that included a covariate for study site to correct for potential geographic differences in subjects reporting the same race. We tested for the additive effect of minor alleles in all analyses. To explore the relationship and pairwise correlation between phenotypes and to determine whether SNPs associated with multiple phenotypes had independent effects on those phenotypes (as opposed to substance dependence in general), we tested models in which we adjusted the phenotype of interest for the other phenotypes with which a given SNP was nominally associated.

As an additional test to confirm findings from regression-based approaches in the family data set, we used a family-based association test implemented in the FBAT statistical package (Rabinowitz and Laird, 2000). These transmission tests are performed within families and remove potential bias due to population stratification. FBAT analyses were considered confirmatory rather than primary because these statistics underutilize the data when transmission cannot be determined with the available genotype data.

To account for the multiple tests performed and to obtain an empirical null distribution of association test P-values, we conducted 1000 null simulations in both the family and case–control samples. In the simulations, we preserved the correlation structure of both the genotypes and the four phenotypes by randomly permuting them as blocks. Individuals' covariates were preserved and adjusted for in each simulation as well, and separate simulations were performed in AAs and EAs. The simulated data sets were then tested for association using logistic regression in the case–control sample and GEE models for the families. We took the smallest overall P-value for each replicate (that is, the smallest value obtained from any SNP–phenotype combination) to account for the multiple SNPs and phenotypes analyzed. In the case–control sample, we randomly assigned genotypes to phenotypes across the entire data set. In the families, we randomly permuted genotypes and phenotypes for sibling pairs across families.

As there were larger families and some parents with genotype information, we created a number of pairs equal to half the total number of genotyped individuals in the sample ($N = 499$ in AA; $N = 383$ in EA) to approximate the information content in the original data. The resulting null distribution is approximate, as larger sibships were not explicitly modeled and parents were not included, but the method used prevents larger families and parents from being oversampled. To exceed an experiment-wide P-value significance threshold of $P < 0.05$, raw P-values had to be less than 0.0006 in AA families, 0.001 in EA families, 0.001 in AA case–controls, 0.002 in EA case–controls.

Haplotype Generation and Association Tests

We constructed haplotypes within the two main population groups (AAs and EAs), which included persons reporting Hispanic descent. As more SNPs were genotyped in the family samples, we derived the haplotype structure for each population group using family data and the program Haploview v.3.2 (Barrett *et al*, 2005). Haplotype associations were evaluated in the family sample with FBAT (Rabinowitz and Laird, 2000) and in the case–control sample with PLINK (Purcell *et al*, 2007). For all tests of haplotypic association, we defined the haplotypes based on the SNP combinations identified by Haploview.

RESULTS

Table 2 shows several characteristics of both samples by population. Among the four substance dependence diagnoses tested in both samples, ND and CD were most strongly correlated ($r = 0.36$). AD and OD were not significantly correlated ($r = 0.02$, $P > 0.05$). All of the other pairwise correlations were between 0.22 and 0.31.

In the total sample of families and case–control individuals, the significance of population group as a covariate differed among the phenotypes. There were significantly more AAs with CD and ND; OD and AD rates did not differ significantly by population. Among EAs, participants with Hispanic descent had significantly higher rates of CD, OD, and ND. Among AAs, those with Hispanic descent had higher rates of OD. These differences may be due in part to sociological and environmental variation across the groups studied, as well as to study design and ascertainment factors, highlighting the need to adjust for ancestry.

There was also substantial correlation between the SNPs across the three genes. In EAs, there were four haplotype blocks (Figure 1) and in AAs there were five haplotype blocks (Figure 2). In EAs, one block spanned the *CHRNA5* and *CHRNA3* genes and another spanned *CHRNA3* and *CHRNB4*. In AAs, only one block contained SNPs in multiple genes (*CHRNA5* and *CHRNA3*). Linkage disequilibrium decreased more rapidly with genetic distance in AAs.

Alcohol Dependence

Four *CHRNA5* SNPs were nominally associated with lifetime AD in the EA family-based discovery sample (Table 3). Three of these SNPs (rs3841324, rs684513, and rs621849) were tested in the replication sample and all showed

Table 2 Baseline Characteristics of the Family and Case–Control Samples

Variable	African-American		European-American	
	Family (N = 2116)	CC (N = 1912)	Family (N = 1705)	CC (N = 1476)
Mean age (SD)	40.7 (7.5)	40.7 (9.9)	38.0 (11.2)	38.2 (10.8)
Male (%)	52	54	53	57
OD/control (N)	311/721	349/1519	480/306	624/807
CD/control (N)	874/147	1273/587	595/176	790/615
ND/control (N)	644/401	1014/879	584/229	834/622
AD/control (N)	462/375	1061/611	394/217	740/463
Two SD diagnoses (N)	355	439	194	310
Three SD diagnoses (N)	311	431	267	275
Four SD diagnoses (N)	110	117	167	69

Abbreviations: AD, alcohol dependence; CC, case control; CD, cocaine dependence; ND, nicotine dependence; OD, opioid dependence; SD, substance dependence
Overall numbers reflect the number of family members recruited to the study; case and control numbers reflect the number of participants with non-missing genotype and phenotype data.

nominally significant association in the AA case–control group. For each SNP, however, the direction of the effect was opposite to that in the EA families (Table 3), which does not indicate a replicable association. Nominally significant associations were observed with five other SNPs in the AA case–control group, and two of these, rs615470 (OR = 0.77, $P = 0.0006$, $P_{empirical} = 0.02$) located in the 3′ UTR of *CHRNA5* and rs578776 (OR = 0.78, $P = 0.001$, $P_{empirical} = 0.03$) in the 3′ UTR of *CHRNA3*, were significant at the experiment-wide 0.05 threshold. The trend for rs615470 was the same in the EA case–control sample, although it was not significant (OR = 0.85, $P = 0.13$). Although only one of the eight SNPs that were associated with AD in the AA case–control sample was also nominally significant in the EA case–control sample (rs2869546), the patterns of association and odds ratios were similar for six others (Table 4).

Nicotine Dependence

Rs16969968, a widely studied nonsynonymous coding SNP in *CHRNA5*, was nominally significantly associated with lifetime ND in the AA family (OR = 1.78, $P = 0.002$, $P_{empirical} = 0.12$) and case–control (OR = 1.43, $P = 0.01$, $P_{empirical} = 0.30$) samples. This SNP was not associated with ND in either of the EA samples. Although there is some evidence that this SNP acts recessively with respect to ND risk (Saccone *et al*, 2007), assuming a recessive model for the minor allele did not yield evidence for association in any of the samples. An adjacent *CHRNA5* SNP, rs615470, was nominally associated with ND in the EA families (OR = 0.74, $P = 0.03$).

Cocaine Dependence

Rs16969968 was nominally associated with lifetime CD in the combined group of EA and AA families (OR = 0.78, $P = 0.04$), thus replicating a previous finding (Grucza *et al*,
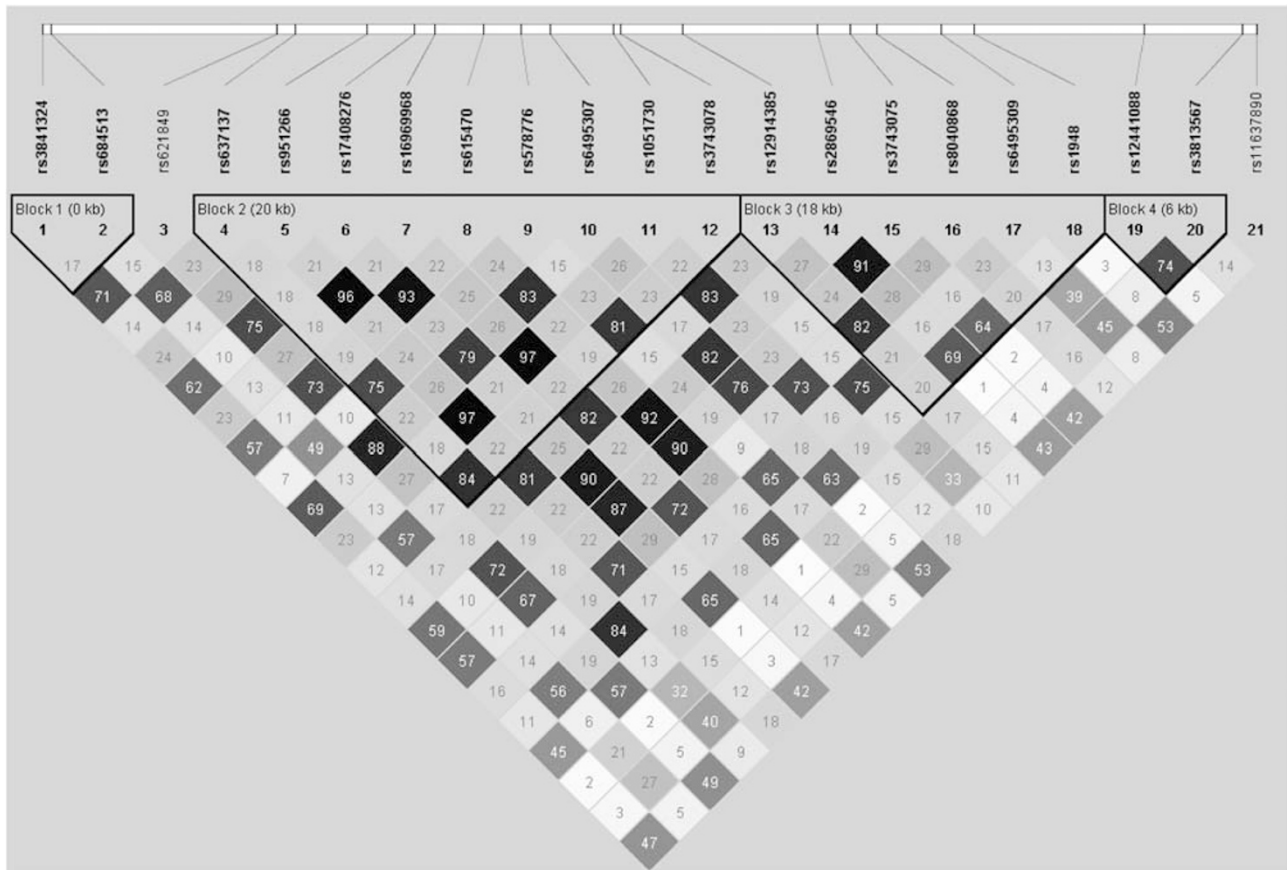
Figure 1 Linkage disequilibrium in the CHRN cluster in European-American families.

2008). However, the pattern of association is opposite that observed with ND in the same groups (Table 3), which was also the case in the previous study (Grucza *et al*, 2008). Nominally significant associations were also observed for *CHRNA5* SNPs, rs621849 (OR = 0.82, P = 0.01) and rs615470 (OR = 0.84, P = 0.02), in the AA case–control group. One SNP in the gene, rs684513 (OR = 1.43, P = 0.0004, $P_{empirical}$ = 0.01), was significant at the experiment-wide 0.05 level, and was also the smallest P-value observed for any phenotype in any sample. In the EA families, CD was nominally associated with two *CHRNA3* SNPs, rs12914385 (OR = 0.64, P = 0.004, $P_{empirical}$ = 0.19) and rs8040868 (OR = 0.66, P = 0.008, $P_{empirical}$ = 0.31).

Opioid Dependence

We observed a nominally significant association between rs16969968 and lifetime OD in the AA case–control sample (OR = 1.45, P = 0.03) and in the combined sample of AA and EA families (OR = 1.25, P = 0.03). In addition, OD was significantly associated with rs615470 and rs2869546 in the EA families (Table 3) and with rs684513 and rs3743078 in the AA case–controls (Table 4). The FBAT analyses yielded two nominally significant results for OD in white families: one SNP in *CHRNA5* (rs686513) and one SNP in *CHRNA3* (rs578776), both at P = 0.02.

Haplotype Association

Three haplotypes were more strongly associated with one of the four dependence traits than the individual SNPs within the haplotype. In the EA family sample, a haplotype comprised of rs3841324 and rs684513 (block 1 in Figure 1) was associated with OD using a global test (P = 0.03), whereas the individual SNPs in the haplotype were not associated with the trait. In the AA case–control sample, there were two haplotype associations with global P-values <0.01. A haplotype comprised of rs615470 and rs578776 was associated with AD (P = 0.005), and another haplotype containing rs621849 and rs16969968 was associated with OD (P = 0.008). The single SNP association between rs16969968 and OD was nominally significant (P = 0.03), whereas rs621849 was not associated with OD. Table 5 shows the haplotypes that were significantly associated with one of the dependence phenotypes.

Independence of Genetic Risk Factors

As many subjects in our sample were dependent on multiple substances, we tested whether the SNPs associated with multiple outcomes were independent risk factors or whether the multiple association signals were the result of correlated phenotypes. Several markers were associated with multiple addictions in at least one sample, including rs648513 (AD and CD), rs16969968 (ND, OD, and AD), and rs615470 (ND,
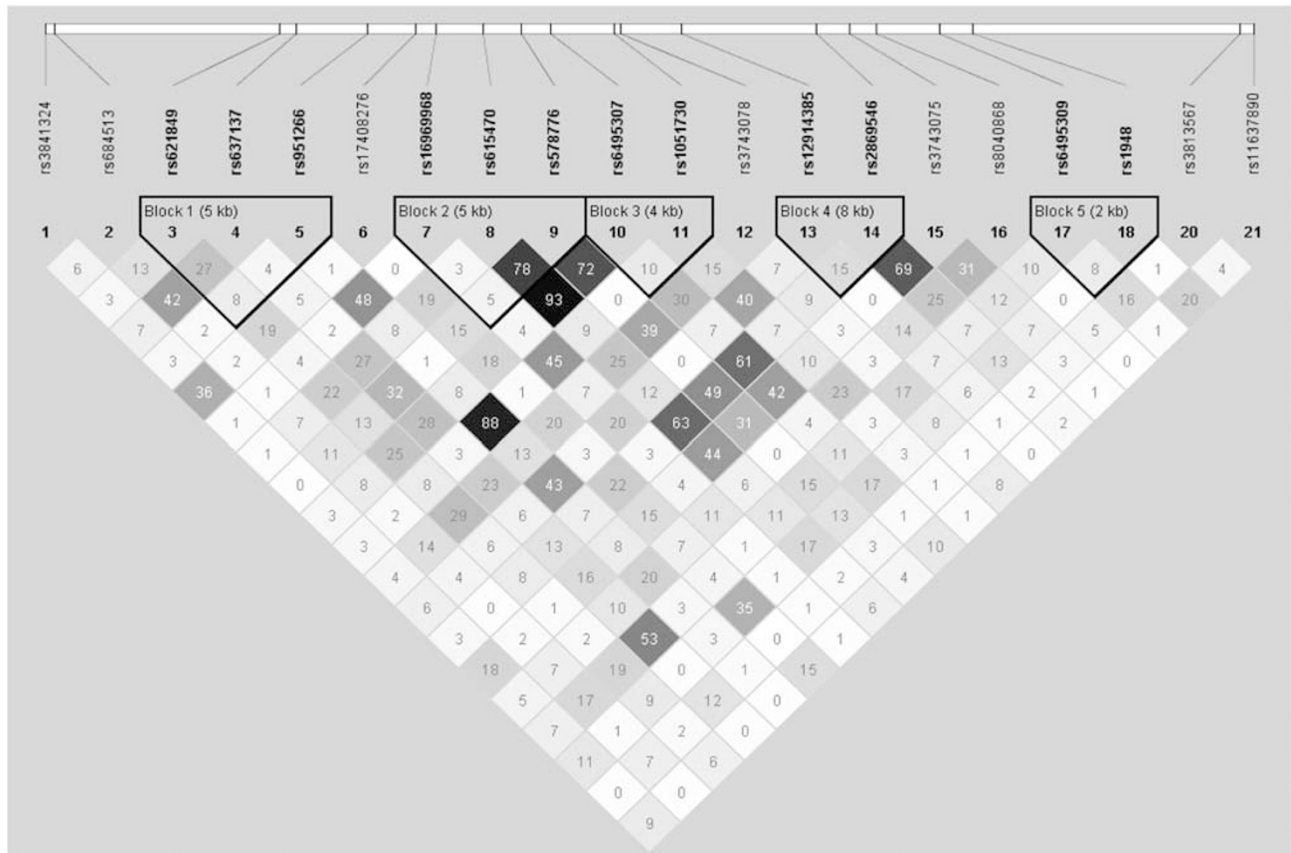
Figure 2 Linkage disequilibrium in the CHRN cluster in African-American families.

CD, OD, and AD). Although the associations of rs684513 with CD ($P = 0.004$) and OD ($P = 0.02$) remained significant after adjustment for the other two outcomes in the AA case–control sample, the association with AD was not significant after adjustment for OD and CD, indicating that the observed effects were not substance specific. The associations of rs16969968 with each of the dependence diagnoses in the AA families seemed to be independent because the *P*-values were unchanged or increased in significance for each of the three traits after adjustment for the other two (ND = 0.001, CD = 0.04, and OD = 0.02). Finally, for the analysis of rs615470, we combined the family and case–control samples, as the association signals for the individual diagnoses were observed in various data sets (see Tables 2 and 3). In the total sample with adjustment for race, this SNP was associated with AD after adjustment for all other traits ($P = 0.01$), though it was not associated with any of the other traits after adjusting for comorbid dependence diagnoses.

DISCUSSION

Consensus is emerging within the field of addiction genetics that at least two (Saccone *et al*, 2009) and possibly three (Wang *et al*, 2008; Wang *et al*, 2009) distinct loci in the *CHRNA5/A3/B4* cluster contribute to risk, although smoking phenotypes and lung cancer have received the most thorough analysis. Two of the SNPs that tag these loci,

rs16969968 and rs578776, were genotyped in this sample, and rs621849 is a proxy for a SNP (rs588765) which is also associated with AD (Wang *et al*, 2008). Our study samples provided a unique opportunity to test these SNPs for association with multiple dependence phenotypes and in people with multiple dependence diagnoses. We found associations between SNPs in the *CHRNA5/A3/B4* gene cluster on chromosome 15 and four substance dependence phenotypes in large samples from two different populations. The two most significant results were observed in the AA case–control sample: rs684513, with the minor allele more common in CD cases (OR = 1.43, $P = 0.0004$, $P_{empirical} = 0.01$), and rs615470, with the minor allele less common in AD cases (OR = 0.77, $P = 0.0006$, $P_{empirical} = 0.02$). Other associations that we report replicate previous findings. First, rs16969968 was nominally associated with CD, OD, and ND in various samples, and the directions of the association are consistent with those observed in other studies (Bierut *et al*, 2007; Saccone *et al*, 2007; Sherva *et al*, 2008; Bierut *et al*, 2008; Grucza *et al*, 2008; Beutler *et al*, 2001). The MAF in our AA sample is slightly higher than in previous studies, which may explain why the association with ND was evident here. We did not observe an association between ND and this SNP in EAs. The primary differences between this sample and previous EA samples in which this SNP was associated with ND is the prevalence of multiple substance dependence diagnoses and not nicotine dependence *per se*, and the makeup of the comparison group, which in some previous studies was defined as

Table 3 Associations Between CHRN SNPs and Dependence Phenotypes in Families

SNP	GENE	POS (BP)	Min (AA/EA)	MAF AA	MAF EA	Class	Alcohol dependence				Cocaine dependence				Opioid dependence				Nicotine dependence			
							AA (N = 757)		EA (N = 706)		AA (N = 929)		EA (N = 667)		AA (N = 939)		EA (N = 667)		AA (N = 951)		EA (N = 700)	
							OR	P	OR	P	OR	P	OR	P	OR	P	OR	P	OR	P	OR	P
rs3841324	CHRNA5	76644868	A	0.23	0.35	In/Del	0.85	0.19	1.56	0.004	1.08	0.65	1.06	0.67	1.15	0.22	0.88	0.32	1.06	0.62	0.99	0.94
rs684513	CHRNA5	76645455	G	0.18	0.25	Intron	1.04	0.79	0.64	0.004	1.02	0.91	1.09	0.64	0.89	0.38	1.26	0.11	0.84	0.16	1.06	0.71
rs621849	CHRNA5	76659916	G	0.44	0.38	Intron	0.93	0.49	1.33	0.05	1.10	0.51	1.05	0.74	0.90	0.32	0.83	0.14	1.00	0.96	0.88	0.32
rs637137	CHRNA5	76661031	A	0.28	0.28	Intron	1.03	0.83	0.82	0.21	0.86	0.37	1.30	0.12	0.95	0.61	1.24	0.12	0.91	0.36	1.13	0.43
rs951266	CHRNA5	76665596	A	0.10	0.34	Intron	1.45	0.02	0.97	0.87	0.68	0.08	0.79	0.13	1.11	0.49	1.06	0.67	1.09	0.57	1.10	0.52
rs17408276	CHRNA5	76668673	C	0.13	0.32	Intron	0.99	0.97	1.43	0.03	1.18	0.47	1.18	0.32	1.14	0.39	0.78	0.08	0.87	0.35	0.84	0.22
rs16969968	CHRNA5	76669980	A	0.06	0.33	Non-syn	1.17	0.42	0.86	0.31	0.72	0.21	0.76	0.06	1.28	0.18	1.04	0.77	1.78	0.002	1.12	0.39
rs615470	CHRNA5	76673043	T	0.43	0.33	3' UTR	1.01	0.90	1.17	0.32	1.10	0.50	1.24	0.18	1.06	0.61	0.75	0.03	0.89	0.24	0.74	0.03
rs578776	CHRNA3	76675455	G/A	0.48	0.33	3' UTR	0.97	0.76	0.93	0.63	1.06	0.72	1.09	0.61	0.97	0.76	1.22	0.13	0.99	0.95	1.11	0.46
rs6495307	CHRNA3	76677376	T	0.44	0.36	Intron	0.98	0.84	1.24	0.14	1.14	0.35	0.98	0.85	1.02	0.89	0.80	0.08	0.91	0.31	0.81	0.10
rs1051730	CHRNA3	76681394	A	0.10	0.32	Syn-coding	1.28	0.13	0.92	0.54	0.75	0.18	0.79	0.10	1.21	0.21	1.10	0.48	1.20	0.22	1.12	0.41
rs3743078	CHRNA3	76681814	G/C	0.43	0.31	Intron	1.19	0.11	0.82	0.16	0.91	0.53	1.32	0.10	0.98	0.87	1.35	0.03	1.01	0.91	1.14	0.37
rs12914385	CHRNA3	76685778	T	0.21	0.37	Intron	1.10	0.49	0.92	0.57	0.85	0.33	0.64	0.004	1.07	0.59	1.05	0.71	1.21	0.11	1.13	0.37
rs2869546	CHRNA3	76694400	C	0.39	0.34	Intron	0.92	0.45	1.22	0.19	1.24	0.15	1.10	0.53	1.02	0.89	0.77	0.05	0.93	0.45	0.82	0.15
rs3743075	CHRNA3	76696507	T	0.47	0.33	Syn-coding	0.90	0.34	1.32	0.08	1.30	0.06	1.16	0.37	0.98	0.84	0.78	0.07	0.95	0.61	0.85	0.25
rs8040868	CHRNA3	76698236	C	0.32	0.39	Syn-coding	1.08	0.49	0.91	0.52	0.92	0.57	0.66	0.008	1.15	0.18	1.05	0.73	1.11	0.27	1.07	0.63
rs6495309	CHRNA3	76702300	T	0.25	0.26	3' UTR	0.95	0.69	0.87	0.38	0.93	0.68	1.39	0.06	0.98	0.82	1.14	0.34	0.93	0.49	0.98	0.88
rs1948	CHRNB4	76704454	A	0.23	0.29	3' UTR	1.00	1.00	1.27	0.13	1.09	0.65	0.91	0.51	1.11	0.42	0.84	0.19	0.89	0.33	0.93	0.61
rs12441088	CHRNB4	76715319	G	0.32	0.24	Intron	0.92	0.55	0.74	0.12	0.91	0.67	1.12	0.56	1.25	0.12	1.03	0.86	0.90	0.38	0.79	0.19
rs3813567	CHRNB4	76721606	G	0.21	0.23	Upstream	0.89	0.41	0.85	0.31	0.76	0.16	1.09	0.61	1.18	0.22	1.05	0.72	1.05	0.66	0.91	0.56
rs11637890	CHRNB4	76722474	G	0.15	0.34	Upstream	1.21	0.23	1.40	0.03	1.14	0.53	1.10	0.51	1.30	0.04	0.95	0.67	1.06	0.67	0.96	0.79

Abbreviations: AA, African American; EA, European American; MAF, minor allele frequency; Min, minor allele; OR, odds ratio for each additional copy of the minor allele.
P indicates raw P-value, no associations were significant at the 5% level based empirical simulations and N indicates the total number of subjects contributing to each analysis.

Table 4 Associations Between CHRN SNPs and Dependence Phenotypes in the Case–Control Sample

SNP	GENE	POS (BP)	Min (AA/EA)	MAF AA	MAF EA	Alcohol dependence						Cocaine dependence						Opioid dependence						Nicotine dependence					
						AA (N = 1537)			EA (N = 1016)			AA (N = 1708)			EA (N = 1186)			AA (N = 1721)			EA (N = 1210)			AA (N = 1745)			EA (N = 1232)		
						OR	P	P_{emp}	OR	P	P_{emp}	OR	P	P_{emp}	OR	P	P_{emp}	OR	P	P_{emp}	OR	P	P_{emp}	OR	P	P_{emp}	OR	P	P_{emp}
rs3841324	CHRNA5	76644868	A	0.21	0.37	0.78	0.03	0.28	0.87	0.34	0.94	0.87	0.25	0.99	0.96	0.88	1	0.93	0.43	1	1.03	0.41	1	0.95	0.70	1	1.02	0.55	1
rs684513	CHRNA5	76645455	G	0.20	0.24	1.24	0.04	0.63	1.15	0.32	0.99	1.43	0.002	**0.01**	1.03	0.47	1	0.79	0.04	0.83	1.08	0.16	1	1.07	0.65	1	1.00	0.72	1
rs621849	CHRNA5	76659916	G	0.43	0.39	0.80	0.007	0.14	0.92	0.46	0.99	0.82	0.02	0.30	1.00	0.86	1	1.10	0.25	1	1.07	0.39	0.99	0.88	0.08	0.83	1.03	0.59	1
rs16969968	CHRNA5	76669980	A	0.06	0.34	0.99	0.763	1	0.98	0.74	1	1.14	0.54	1	0.98	0.61	1	1.45	0.23	0.64	0.90	0.08	0.98	1.43	0.02	0.30	1.06	0.93	1
rs615470	CHRNA5	76673043	T	0.39	0.35	0.77	0.0004	**0.02**	0.85	0.22	0.93	0.84	0.04	0.51	0.90	0.37	0.99	1.00	0.69	1	0.99	0.98	1	0.89	0.11	0.96	0.95	0.96	1
rs578776	CHRNA3	76675455	G/A	0.45	0.31	0.78	0.002	**0.03**	1.17	0.13	0.96	0.89	0.20	0.99	1.09	0.28	0.99	1.13	0.24	0.99	1.17	0.02	0.81	0.98	0.80	1	0.98	0.78	1
rs6495307	CHRNA3	76677376	T	0.41	0.39	0.82	0.01	0.30	0.82	0.11	0.65	0.87	0.19	0.93	0.87	0.23	0.94	1.01	0.69	1	0.97	0.94	1	0.94	0.53	1	0.93	0.80	1
rs1051730	CHRNA3	76681394	A	0.11	0.35	1.05	0.49	1	1.02	0.82	1	0.98	0.66	1	0.97	0.41	1	1.17	0.64	1	0.86	0.02	0.65	1.15	0.20	1	1.07	0.78	1
rs3743078	CHRNA3	76681814	G/C	0.41	0.27	0.79	0.006	0.07	1.13	0.11	0.99	0.86	0.07	0.88	1.02	0.64	1	1.25	0.04	0.30	1.12	0.11	0.99	0.96	0.70	1	0.92	0.65	1
rs12914385	CHRNA3	76685778	T	0.18	0.38	0.98	0.9631	1	1.02	0.83	1	0.99	0.84	1	0.98	0.58	1	1.00	0.56	1	0.91	0.13	0.99	1.22	0.05	0.51	1.13	0.37	0.93
rs2869546	CHRNA3	76694400	C	0.38	0.36	0.82	0.009	0.30	0.80	0.09	0.58	0.96	0.60	1	0.91	0.45	0.99	0.89	0.31	1	0.98	0.89	1	0.94	0.35	1	0.95	0.78	1
rs8040868	CHRNA3	76698236	C	0.34	0.41	1.07	0.33	1	1.09	0.76	0.99	0.94	0.56	1	1.01	0.67	1	1.06	0.97	1	0.91	0.09	0.99	1.00	0.94	1	1.09	0.54	0.99
rs11637890	CHRNB4	76722474	G	0.15	0.35	0.85	0.24	0.99	0.92	0.25	1	0.91	0.93	1	0.96	0.61	1	1.20	0.10	0.99	1.06	0.41	1	0.91	0.91	1	1.03	0.66	1

Abbreviations: AA, African American; EA, European American; MAF, minor allele frequency; Min, minor allele; P, Raw P-value; P_{emp}, Empirical P-value based on 1000 null simulations; OR, odds ratio for each additional copy of the minor allele.
Values with an experiment-wide significance less than 0.05 are in bold.
N indicates the total number of subjects contributing to each analysis.

Table 5 Nicotinic Receptor Haplotypes Associated with Substance Dependence Traits

SNPs	Alleles (Freq case/control)	Sample	Phenotype	P-value (global/haplotype)
rs3841324, rs684513	A–C (0.22/0.26)	EA families	OD	0.03/0.01
rs615470, rs578776	C–A (0.51/0.57)	AA case–control	AD	0.005/0.001
rs621849, rs16969968	A–A (0.05/0.08)	AA case–control	OD	0.008/0.01

Alleles shown represent the haplotype with the strongest effect on the trait of interest and the haplotype frequency in cases and controls.

nicotine-exposed but not dependent. There are a large number of subjects dependent on both nicotine and cocaine ($N = 449$ in the families and $N = 634$ in the case–controls), and ND and CD are highly correlated ($r = 0.5$) in the EA case–controls. It is possible that the number of EA subjects dependent on both substances limited our ability to detect association with either. For example, the protective effect of a low-risk allele for ND (which is a high-risk allele for CD) might be obscured by the reinforcing effect of nicotine on cocaine-induced DA overflow. The SNP was not associated with ND in EA families or case–controls without CD, although only 125 related and 639 unrelated individuals contributed to the results, yielding most power to detect an association. Finally, this variant had a relatively small effect on risk in earlier reports, and therefore chance can influence whether or not effects are detected in individual samples.

An increasing body of evidence suggests that rs16969968 is associated with multiple substance dependence phenotypes (Bierut et al, 2007; Saccone et al, 2007; Sherva et al, 2008; Bierut et al, 2008; Grucza et al, 2008; Wang et al, 2008), though the mechanism by which minor alleles can protect against ND while increasing risk of CD and AD remains to be determined. One potential biological explanation for these effects is based on reduced receptor function due to the polymorphism differentially impacting the mechanisms through which nicotine and cocaine alter the mesolimbic dopaminergic system (Grucza et al, 2008). We also observed a weak association between ND and rs1051730, which was associated with smoking quantity in an Icelandic sample (Thorgeirsson et al, 2008). We did not find an association between this SNP and daily cigarette consumption in either data set. Although we failed to replicate the association between the minor allele of rs578776 and reduced ND risk (Bierut et al, 2008), this SNP allele was associated with reduced risk of AD in EAs. The observed associations between AD and both rs3841324 and rs621849 have also been reported previously (Wang et al, 2008).

The differences in the results observed across population groups, coupled with the fact that rates of use and dependence on each substance studied here vary by population (http://archives.drugabuse.gov/pdf/minorities03.pdf), are consistent with what is already known about the importance of both genetic and environmental factors in substance dependence. The allele frequencies of all but three of the 21 SNPs differ significantly by population, and 12 of these were nominally associated with at least one of the phenotypes. The different patterns of association and allele frequencies suggest that at least part of the population differences in susceptibility to substance dependence is

genetic. The haplotypic association results do not support the idea that the same CHRN risk alleles existing on different genetic backgrounds explain the population differences observed in the single SNP analyses. The inclusion of participants reporting Hispanic identification should be noted, as the allele frequencies of SNPs in this cluster have been shown to differ in Hispanics compared with AAs and EAs (Schlaepfer et al, 2008), which we also observed in these data. Although we adjusted for Hispanic descent within the AA and EA analysis groups, we also performed the association tests within larger population groups after excluding Hispanics to test for any residual confounding. In general, the strength of the association signals improved after excluding Hispanics. In these analyses, we observed association between rs16969968 and ND in AA families (OR = 1.62, $P = 0.01$) and also with CD in EA families (OR = 0.66, $P = 0.01$). None of the association findings discussed disappeared after excluding Hispanics, indicating that their inclusion did not result in positive confounding. As a second test to limit the potential for confounding, we performed the association tests with study site as a covariate. Although study site was a significant predictor of the dependence phenotypes (to varying degrees in population and sample subgroups) because of the intended recruitment at each site, the association results did not change substantially (data not shown).

Our results should be interpreted in light of several study limitations. The substance-dependent cases studied here were ascertained from treatment settings and through advertisements, and had high rates of comorbid dependence diagnoses. Although the controls from the case–control sample were largely recruited through advertisements, the family-based controls, while not dependent on the substance of interest, came from families in which addiction was prevalent. This is reflected in the fact that for every trait except OD, cases outnumbered controls, often by a substantial margin. In addition, the control group for each individual dependence trait contained individuals who were dependent on other substances, which may have masked some associations; on the other hand, positive associations observed in this context likely reflect specific associations with the substance considered, as opposed to general SD risk alleles. Thus, we expect substantial enrichment for addiction-related alleles, especially in the families. To address this issue, we conducted analyses in which cases for each substance were compared with a control group that was not dependent on any of the substances of interest ($N = 309$ AA, 287 EA). This was only practical in the case–control sample because of the small number of 'clean' controls in the families. These analyses showed many of the same associations seen in the primary analyses (data not

shown), although the significance of the associations was diminished, probably because of the smaller informative sample size. These analyses suggest that our use of control groups dependent on different substances than the one of interest did not mask association signals, with the caveat that they had lower power than the primary analyses.

Other caveats to the interpretation of these results include the paucity of significant *P*-values on the basis of our simulation-based multiple test correction and the reduction in power to detect significant effects when compared with subjects with no substance dependence. The strengths of this study include the inclusion of discovery and replication samples from two populations, standardized phenotyping of dependence traits and constituent symptoms, excellent SNP coverage of the nicotinic receptor cluster on chromosome 15, and a robust, empirical measurement of the significance of the results.

CONCLUSIONS

These data support the importance of variants in the *CHRNA5/A3/B4* gene cluster as mediators of the genetic risk for substance dependence. In addition to replicating previous associations of rs16969968 with ND and CD, these data provide additional evidence that SNPs in the cluster are associated with AD, and extend these findings to AAs. As a whole, more associations were observed in AAs, indicating that these genes may be more important risk factors in this population, at least in persons ascertained through treatment facilities and advertisements that target substance-dependent individuals. Alternatively, this observation could be due to the larger effective sample size of AAs than EAs in this study. Numerous findings of SNPs associated with multiple phenotypes suggest that the gene cluster is partially responsible for the shared genetic risk across multiple substances.

DISCLOSURE

Dr Gelernter reports that he has received compensation for professional services in the previous 3 years from the following entities: the Yale University School of Medicine, Veterans Affairs Healthcare System (VA) and the National Institutes of Health (NIAAA, NIDA, and NIMH) and related to academic lectures and editorial functions in various scientific venues (including the ACNP). Dr Kranzler has received compensation for professional services from the National Institutes of Health (NIAAA and NIDA) and for academic lectures and editorial functions in various scientific venues (including the ACNP). Dr Kranzler also reports consulting arrangements with Alkermes, Elbion

Pharmaceuticals, Ortho-McNeil Pharmaceuticals, Sanofi-Aventis Pharmaceuticals, and Solvay Pharmaceuticals, and research support from Bristol-Myers Squibb Company, Merck & Co., and Ortho-McNeil Pharmaceuticals. Dr Anton reports the following: being a consultant for Sanofi-Aventis, Eli Lilly, Merck, Organon, Hythiam, Novartis, Johnson & Johnson, and GlaxoSmithKline; serving on an Advisory Board for Solvay, Hythiam, Novartis, and Johnson & Johnson, and receiving grant support from Eli Lilly, Merck, Hythiam, and Johnson & Johnson. Dr Anton and Dr Kranzler also report having associations with the following pharmaceutical companies: Eli Lilly, Janssen, Schering Plough, Lundbeck, Alkermes, GlaxoSmithKline, Abbott, and Johnson & Johnson, as these companies provide support to the Alcohol Clinical Trials Initiative (**ACTIVE**) and both Dr Kranzler and Dr Anton receive support from **ACTIVE**. These disclosures do not hold any influence on the content of this paper.

REFERENCES

Amos CI, Wu X, Broderick P, Gorlov IP, Gu J, Eisen T *et al.* (2008). Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat Genet* **40**: 616–622.

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

Benowitz NL (1996). Pharmacology of nicotine: addiction and therapeutics. *Annu Rev Pharmacol Toxicol* **36**: 597–613.

Beutler E, Lichtman M, Coller BS, Kipps TJ, Seligsohn U (2001). The sickle cell diseases and related disorders. In: Beutler E, Lichtman MA, Coller BS (eds). *Williams Hematology*, 6th edn. McGraw Hill: New York. pp 581–605.

Bierut LJ, Dinwiddie SH, Begleiter H, Crowe RR, Hesselbrock V, Nurnberger JI *et al.* (1998). Familial transmission of substance dependence: alcohol, marijuana, cocaine, and habitual smoking: a report from the Collaborative Study on the Genetics of Alcoholism. *Arch Gen Psychiatry* **55**: 982–988.

Bierut LJ, Madden PA, Breslau N, Johnson EO, Hatsukami D, Pomerleau OF *et al.* (2007). Novel genes identified in a high-density genome wide association study for nicotine dependence. *Hum Mol Genet* **16**: 24–35.

Bierut LJ, Stitzel JA, Wang JC, Hinrichs AL, Grucza RA, Xuei X *et al.* (2008). Variants in nicotinic receptors and risk for nicotine dependence. *Am J Psychiatry* **165**: 1163–1171.

Carmelli D, Swan GE, Robinette D, Fabsitz R (1992). Genetic influence on smoking–a study of male twins. *N Engl J Med* **327**: 829–833.

Edwards R (2004). The problem of tobacco smoking. *BMJ* **328**: 217–219.

Ehringer MA, McQueen MB, Hoft NR, Saccone NL, Stitzel JA, Wang JC *et al.* (2009). Association of CHRN genes with 'dizziness' to tobacco. *Am J Med Genet B Neuropsychiatr Genet*, 2010 **153B**: 600–609.

Freathy RM, Ring SM, Shields B, Galobardes B, Knight B, Weedon MN *et al.* (2009). A common genetic variant in the 15q24 nicotinic acetylcholine receptor gene cluster (CHRNA5-CHRNA3-CHRNB4) is associated with a reduced ability of women to quit smoking in pregnancy. *Hum Mol Genet* **18**: 2922–2927.

Gossop M, Stewart D, Treacy S, Marsden J (2002). A prospective study of mortality among drug misusers during a 4-year period after seeking treatment. *Addiction* **97**: 39–47.

Grucza RA, Wang JC, Stitzel JA, Hinrichs AL, Saccone SF, Saccone NL *et al.* (2008). A risk allele for nicotine dependence

in CHRNA5 is a protective allele for cocaine dependence. *Biol Psychiatry* **64**: 922–929.

Heath AC, Martin NG (1993). Genetic models for the natural history of smoking: evidence for a genetic influence on smoking persistence. *Addict Behav* **18**: 19–34.

Hung RJ, McKay JD, Gaborieau V, Boffetta P, Hashibe M, Zaridze D et al. (2008). A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature* **452**: 633–637.

Jackson KJ, Martin BR, Changeux JP, Damaj MI (2008). Differential role of nicotinic acetylcholine receptor subunits in physical and affective nicotine withdrawal signs. *J Pharmacol Exp Ther* **325**: 302–312.

Karkowski LM, Prescott CA, Kendler KS (2000). Multivariate assessment of factors influencing illicit substance use in twins from female-female pairs. *Am J Med Genet* **96**: 665–670.

Kendler KS, Karkowski LM, Corey LA, Prescott CA, Neale MC (1999). Genetic and environmental risk factors in the aetiology of illicit drug initiation and subsequent misuse in women. *Br J Psychiatry* **175**: 351–356.

Kendler KS, Karkowski LM, Neale MC, Prescott CA (2000). Illicit psychoactive substance use, heavy use, abuse, and dependence in a US population-based sample of male twins. *Arch Gen Psychiatry* **57**: 261–269.

Kendler KS, Prescott CA (1998). Cocaine use, abuse and dependence in a population-based sample of female twins. *Br J Psychiatry* **173**: 345–350.

Li MD, Ma JZ, Payne TJ, Lou XY, Zhang D, Dupont RT et al. (2008). Genome-wide linkage scan for nicotine dependence in European Americans and its converging results with African Americans in the Mid-South Tobacco Family sample. *Mol Psychiatry* **13**: 407–416.

Long JC, Knowler WC, Hanson RL, Robin RW, Urbanek M, Moore E et al. (1998). Evidence for genetic linkage to alcohol dependence on chromosomes 4 and 11 from an autosome-wide scan in an American Indian population. *Am J Med Genet* **81**: 216–221.

Perez-Stable EJ, Herrera B, Jacob III P, Benowitz NL (1998). Nicotine metabolism and intake in black and white smokers. *JAMA* **280**: 152–156.

Pierucci-Lagha A, Gelernter J, Feinn R, Cubells JF, Pearson D, Pollastri A et al. (2005). Diagnostic reliability of the Semi-structured Assessment for Drug Dependence and Alcoholism (SSADDA). *Drug Alcohol Depend* **80**: 303–312.

Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**: 559–575.

Rabinowitz D, Laird N (2000). A unified approach to adjusting association tests for population admixture with arbitrary pedigree structure and arbitrary missing marker information. *Hum Hered* **50**: 211–223.

Saccone NL, Wang JC, Breslau N, Johnson EO, Hatsukami D, Saccone SF et al. (2009). The CHRNA5-CHRNA3-CHRNB4 nicotinic receptor subunit gene cluster affects risk for nicotine dependence in African-Americans and in European-Americans. *Cancer Res* **69**: 6848–6856.

Saccone SF, Hinrichs AL, Saccone NL, Chase GA, Konvicka K, Madden PA et al. (2007). Cholinergic nicotinic receptor genes implicated in a nicotine dependence association study targeting 348 candidate genes with 3713 SNPs. *Hum Mol Genet* **16**: 36–49.

Schlaepfer IR, Hoft NR, Collins AC, Corley RP, Hewitt JK, Hopfer CJ et al. (2008). The CHRNA5/A3/B4 gene cluster variability as an important determinant of early alcohol and tobacco initiation in young adults. *Biol Psychiatry* **63**: 1039–1046.

Sherva R, Wilhelmsen K, Pomerleau CS, Chasse SA, Rice JP, Snedecor SM et al. (2008). Association of a single nucleotide polymorphism in neuronal acetylcholine receptor subunit alpha 5 (CHRNA5) with smoking status and with 'pleasurable buzz' during early experimentation with smoking. *Addiction* **103**: 1544–1552.

Thorgeirsson TE, Geller F, Sulem P, Rafnar T, Wiste A, Magnusson KP et al. (2008). A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature* **452**: 638–642.

Tsuang MT, Lyons MJ, Eisen SA, Goldberg J, True W, Lin N et al. (1996). Genetic influences on DSM-III-R drug abuse and dependence: a study of 3,372 twin pairs. *Am J Med Genet* **67**: 473–477.

Tsuang MT, Lyons MJ, Meyer JM, Doyle T, Eisen SA, Goldberg J et al. (1998). Co-occurrence of abuse of different drugs in men: the role of drug-specific and shared vulnerabilities. *Arch Gen Psychiatry* **55**: 967–972.

Wang JC, Cruchaga C, Saccone NL, Bertelsen S, Liu P, Budde JP, et al., COGEND collaborators and GELCC collaborators (2009). Risk for nicotine dependence and lung cancer is conferred by mRNA expression levels and amino acid change in CHRNA5. *Hum Mol Genet* **18**: 3125–3135.

Wang JC, Grucza R, Cruchaga C, Hinrichs AL, Bertelsen S, Budde JP et al. (2008). Genetic variation in the CHRNA5 gene affects mRNA levels and is associated with risk for alcohol dependence. *Mol Psychiatry*, 2009 **14**: 501–510.

Williams JT, Begleiter H, Porjesz B, Edenberg HJ, Foroud T, Reich T et al. (1999). Joint multipoint linkage analysis of multivariate qualitative and quantitative traits. II. Alcoholism and event-related potentials. *Am J Hum Genet* **65**: 1148–1160.