

# Genetic Association of the *CHRNA6* and *CHRN3* Genes with Tobacco Dependence in a Nationally Representative Sample

Nicole R Hoft<sup>1</sup>, Robin P Corley<sup>1</sup>, Matthew B McQueen<sup>1,2</sup>, Isabel R Schlaepfer<sup>1,3</sup>, David Huizinga<sup>4</sup> and Marissa A Ehringer<sup>\*1,3</sup>

<sup>1</sup>Institute for Behavioral Genetics, University of Colorado, Boulder, CO, USA; <sup>2</sup>Department of Psychology, University of Colorado, Boulder, CO, USA;

<sup>3</sup>Department of Integrative Physiology, University of Colorado, Boulder, CO, USA; <sup>4</sup>Institute for Behavioral Science, University of Colorado, Boulder, CO, USA

Neuronal nicotinic acetylcholine receptors are activated by both endogenous acetylcholine and exogenous nicotine, making sequence variations in these receptors likely candidates for association with tobacco phenotypes. Previous studies have found evidence for significant association between single nucleotide polymorphisms (SNPs) in the genomic region containing the *CHRNA6* and *CHRN3* genes and tobacco behaviors. In this study, we provide support for an association between these genes and tobacco dependence in the National Youth Survey Family Study wave 10, a nationally representative sample of households. Eight SNPs in the *CHRNA6* and *CHRN3* genomic region were genotyped in 1051 subjects, approximately half of whom are members of sibling pairs. Genetic association with DSM-IV dependence was assessed using a family-based approach as implemented in the statistical package PBAT. Individual SNPs were tested for association with quit attempts and overall dependence. Variation in *CHRNA6* was found to be associated with tobacco dependence ( $p = 0.007$  in Caucasians). SNPs in *CHRN3* were found to be associated with the number of quit attempts ( $p = 0.0024$ ). Together these results further implicate the region downstream of *CHRNA6* and the region upstream of *CHRN3* in risk of nicotine dependence.

*Neuropsychopharmacology* (2009) **34**, 698–706; doi:10.1038/npp.2008.122; published online 13 August 2008

**Keywords:** nicotinic receptors; SNP; genetic association; tobacco use; nicotine dependence

## INTRODUCTION

Tobacco use is a major health problem in the United States and worldwide. Despite the increasing awareness of health-related consequences, about 25% of American adults continue to smoke (Center for Disease Control, 2000). Some individuals continue despite knowledge of the risks, and others do not understand the risks, but the majority of individuals continue smoking because they are nicotine dependent (Hymowitz *et al*, 1997; Jarvis, 2004). That is, despite having the desire to quit, a large percentage of smokers are unsuccessful in stopping. Studies have shown that genes influence level of nicotine dependence with heritability estimates indicating genetic factors account for between 40 and 70% of the phenotypic variance (Li *et al*, 2003; Maes *et al*, 2004; Rhee *et al*, 2003; Swan *et al*, 1990). Determining which genes are contributing to this heritability is an area of active research.

Neuronal nicotinic acetylcholine receptors (nAChRs) are likely candidate genes to be involved in mediating nicotine dependence. These pentameric ion channels are present in the peripheral and central nervous systems, including dopaminergic neurons, and are activated by endogenous acetylcholine and exogenous nicotine (Mineur and Picciotto, 2008). The *CHRNA6* and *CHRN3* genes coding for the  $\alpha 6$  and  $\beta 3$  receptor subunits are located contiguously in a tail-to-tail configuration on chromosome 8. Their protein products are colocalized in nicotinic receptors in the substantia nigra, ventral tegmental area, striatum, and locus coeruleus (Gotti *et al*, 2006a, b). Located on chromosome 8 at position 8p11.21, *CHRNA6* spans 16.01 kb (NCBI Build 35) and has at least three alternatively spliced transcripts. *CHRN3* is located on chromosome 8 at position 8p11.2 and spans 39.99 kb. Six exons can be alternatively spliced into at least three different transcripts (GeneCards: <http://www.genecards.org>; Rebhan *et al*, 1998).

Both  $\alpha 6$ - and  $\beta 3$ -containing receptors have been found to have a significant role in dopaminergic neurotransmission, thus contributing to reward and reinforcement of behavior (Cui *et al*, 2003). Receptors containing  $\alpha 6\beta 2\beta 3$ - as well as  $\alpha 6\alpha 4\beta 2\beta 3$ -containing receptors in the striatum mediate  $\alpha$ -conotoxin MII-sensitive dopamine release. In contrast,  $\alpha 6\beta 2$ -containing receptors in the superior colliculus seem

\*Correspondence: Dr MA Ehringer, Institute for Behavioral Genetics, University of Colorado, 447 UCB, Boulder, CO 80309, USA, Tel: +1 303 492 1463, Fax: +1 303 492 8063,

E-mail: Marissa.Ehringer@colorado.edu

Received 18 January 2008; revised 2 June 2008; accepted 30 June 2008

to be involved in GABA release (Champtiaux *et al*, 2003; Gotti *et al*, 2006a, b; Salminen *et al*, 2004). The  $\beta 3$  subunit is believed to be important for assembly and stability of  $\alpha 6$ -containing nAChRs (Cui *et al*, 2003).

Recent genetic association studies have provided evidence that single nucleotide polymorphisms (SNPs) in *CHRNA6* and *CHRNA3* are associated with nicotine dependence and subjective response to nicotine. Bierut *et al* (2007) and Saccone *et al* (2007) found SNPs upstream of *CHRNA3* as well as in exon 6 (3' UTR) of *CHRNA6* to be associated with nicotine dependence in a case-control sample (Bierut *et al*, 2007; Saccone *et al*, 2007). Greenbaum *et al* (2006) also found association of a SNP in  $\alpha 6$  (intron 2) with nicotine dependence in a sample of female Israeli students (Greenbaum *et al*, 2006). In addition, Zeiger *et al* (2008) have found significant associations between SNPs located in the upstream region of *CHRNA3* and in exon 6 of the *CHRNA6* gene and subjective responses to nicotine, in separate samples of selected and community young adults. To the best of our knowledge, association of *CHRNA6* and *CHRNA3* with dependence measured by DSM-IV has not been reported, nor has quit attempts been examined. In an effort to strengthen these recent findings as relevant to the general US population, we present our results from testing a subset of SNPs in these genes in a population-representative sample using the DSM-IV definition of dependence.

**MATERIALS AND METHODS**

**Subjects**

*NYS-FS participants and phenotypic assessments.* Polymorphisms in the *CHRNA6* and *CHRNA3* subunit genes were examined in a genetic supplemental sample participating in the National Youth Survey Family Study (NYS-FS; Institute for Behavioral Science, and Institute for Behavioral Genetics, University of Colorado) wave 10. The NYS is a nationally representative probability sample of adolescents aged 11–17 in 1976 and living in the United States in 1977. A follow-up interview was conducted in 2002 (age range 35–44 years), where behavioral data and buccal cell swab samples were obtained on a strictly voluntary basis (Elliott *et al*, 1989). Of the 1051 individuals who agreed to follow-up interviews and provided buccal swab DNA samples, 856 of them had tried cigarettes. Although there has been some selective/disproportionate loss in some waves by sex, or age, or race/ethnicity, or social class across the first 10 waves of the NYS, these differences are quite small and often do not significantly affect the original distribution of these variables as established at the first wave of the NYS, which has been shown to be representative of individuals born in the

continental US in 1959–1965. One of the few significant and the largest of these is a change from 53.2 to 49.2% male individuals (46.8–50.8% female individuals) from the 1st to 10th survey. There has been no significant disproportionate loss over the years by initial alcohol use, marijuana use, or other illicit drug use or by general delinquency or more serious uniform crime rates offending. The distributions of these variables are not significantly different between respondents who completed interviews in a given year and those who did not. In addition, the distributions of sex, age, race/ethnicity, and social class of waves 7–10 are usually not significantly different from those observed in wave 1 (1976) or wave 6 (1983). Thus, overall, the NYS sample appears to remain reasonably representative of the US population born in 1959–1965. The sample consisted of 227 families with sibships ranging from 2 to 5 offspring per family (592 individuals) and 479 individuals without siblings in the sample. A total of 48.1% of subjects were male and 51.9% female, frequencies of minorities were reflective of the US population in 1976 with 80.3% white, 12.3% African American, and 5.4% of Hispanic or other ancestry. Additional summary statistics are given in Table 1.

Alcohol and tobacco use behaviors were assessed during a face-to-face structured interview including an adaptation of the Composite International Diagnostic Interview—Substance Abuse Module (Cottler and Keating, 1990). Nicotine dependence was assessed using DSM-IV criteria. The seven symptoms with their respective endorsement rates are given in Table 2. An individual was considered dependent if they endorsed three or more symptoms. Two phenotypes were analyzed: nicotine dependence as defined by three or more DSM-IV dependence symptoms, and reported number of unsuccessful quit attempts. The number of unsuccessful quit attempts was used because it is a quantitative measure and was hypothesized to be a closer biological measure of addiction (than overall dependence). Phenotypes were examined for ethnic and sex differences using ANOVA in SPSS.

Three subsets of the sample were used, which are described in Table 3. The initial analysis was performed only with ‘current frequent smokers’, which included individuals reporting having used tobacco at least 180 days (almost every day) in the 6 months before the interview. This is more stringent (restricting the sample to current smokers) but similar to the inclusion criteria used by Bierut *et al* (2007) of having smoked at least 100 cigarettes. Secondary analyses were performed with (1) sets of ‘smokers’ who were individuals who reported having used tobacco almost every day for at least a month at some point in their life and (2) the entire sample of 1051 individuals.

**Table 1** Sample Characteristics for the NYS-FS Sample

Sample	Male (%)	Female (%)	Ever tobacco (%)	Ever alcohol (%)
Pooled (1051)	506 (49)	545 (51)	856 (81)	990 (94)
Caucasian (860)	410 (48)	450 (52)	727 (85)	824 (96)
Hispanic (28)	11 (39)	17 (61)	23 (82)	27 (96)
African American (132)	67 (51)	65 (49)	82 (62)	109 (83)

**Table 2** Phenotype Descriptions

DSM-IV-dependence symptoms		Endorsement rate % (n)
Tolerance	Increased amounts needed for same effect, or decreased effect	39.8 (426)
Withdrawal	Use to avoid symptoms, or symptoms when discontinue use	25.2 (270)
Increased use	Increasing either quantity or duration	26.4 (283)
Quit failure	Loss of control inability to quit or cut down despite desire to	18.3 (196)
Important activities	Activities given up due to substance use	1.9 (20)
Time	A great deal of time is spent obtaining, using, or recovering from effects of the substance	8.7 (94)
Continued use	Continued use despite substance causing, or exacerbating physical, or psychological problems	18.8 (201)

**Table 3** Inclusion Criteria

Subsample	Description
All	All subjects interviewed and genotyped
Smokers	Individuals reporting having ever used tobacco almost everyday for at least a month
Current frequent smokers	Individuals who reported using tobacco 180 days in the 6 months before interview (almost every day)

### Selection of SNPs and Genotyping

Candidate polymorphisms for the *CHRNA6* and *CHRN3* genes were identified using the SNPbrowser software version 3.5 from Applied Biosystems (<http://www.appliedbiosystems.com>), the public database, dbSNP (<http://www.ncbi.nlm.nih.gov/SNP/>), and other published research. SNPs were chosen to span the two genes, have moderate minor allele frequencies, and assay reliably. The majority of SNPs chosen were identical to those used in Zeiger *et al* (2008) except for a few additional SNPs added to examine regions where association had recently been reported. The structures of the *CHRNA6* and *CHRN3* genes, and the SNPs selected, are shown in Figure 1, with SNPs examined in this study shown in bold. Other recently published SNPs are also included in gray to show proximity and spacing.

Genomic DNA was isolated from buccal cell swabs and preamplified using the method of Zhang *et al* (1992). Data obtained using this DNA are of high quality; these methods have been shown to be reliable for genotyping (Anchordoquy *et al*, 2003; Haberstick and Smolen, 2004). A Biomek<sup>®</sup> 3000 Laboratory Automation Workstation (Beckman Coulter) was used to automate DNA genotyping assay preparation in a 384-well plate format. TaqMan<sup>®</sup> assays for allelic discrimination (Applied Biosystems) were used to determine SNP genotypes, per instructions of the manufacturer under standard conditions using an ABI PRISM<sup>®</sup> 7900 instrument. In total, eight SNPs located in the *CHRNA6*–*CHRN3* region of chromosome 8 were genotyped in 1051 subjects.

### Analytic Methods

Linkage disequilibrium (LD) and genotype quality were analyzed using Haploview (Barrett *et al*, 2005). This freely available software package allows for estimation of several

LD statistics, including pairwise LD correlations or  $r^2$ . All SNPs were checked for Hardy–Weinberg equilibrium (HWE) within ethnic groups using Haploview and further verified in SPSS version 15.0.

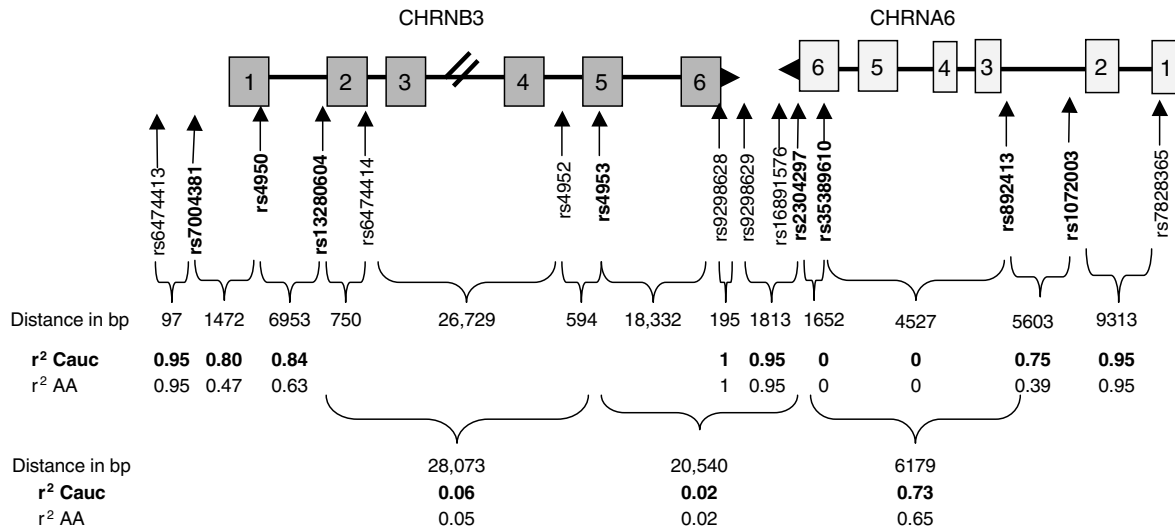
Analyses were conducted separately on the subset of genetically related individuals (227 families totaling 529 individuals). Single marker analyses were performed using PBAT (Lange *et al*, 2004). This statistical genetics program takes full advantage of the family structure, making use of both the between- and within-family information (Fulker *et al*, 1999). All analyses were first conducted on the subset of ‘current frequent smokers’ using the ethnic group information as covariates, and then conducted separately for the Caucasians. Three pairs of monozygotic twins were removed and 10 suspected half siblings were removed from the data set before genetic analysis. Secondly, because subject inclusion and recruitment criteria vary widely among genetic association studies, following the initial analysis, *post hoc* exploratory analyses were performed to examine the effects of the inclusion criteria on the significance of the observed genetic associations. The inclusion criteria are described in Table 3.

Following the analysis of the siblings, subsequent analyses were conducted using the singleton individuals (i.e. those with no genetic data from other family members). The analysis of the singletons (which could not be included in the PBAT analysis due to lack of family genotypic information) was performed in SPSS version 15.0 and SAS version 9.1 using  $\chi^2$ -tests for binary traits and ANOVA/regression models for continuous traits. In SPSS and SAS, ethnicity was included as a covariate when possible, and also the subset of Caucasians was examined separately. Thus, two subsamples were generated from the genetic sample: the group of individuals with sibships who were analyzed with PBAT, and the singletons group, analyzed using SAS and SPSS. Significance levels from the two samples were combined using Fisher’s method where appropriate.

## RESULTS

### Phenotype Analysis in the NYS-FS

The mean number of DSM-IV symptoms for each represented ethnic group, as well as each phenotypic subgroup is given in Table 4, along with the significance of group differences tested by ANOVA. Overall responses to the DSM-IV items were available for 1051 individuals with a



**Figure 1** Representation of the *CHRNA6-CHRN3* locus. Boxes represent exons separated by intronic regions (not drawn to scale). Eight single nucleotide polymorphisms (SNPs) were genotyped shown in bold, with their reference sequence numbers and gene locations indicated. The number of nucleotide base pairs (bp) between each SNP, and the linkage disequilibrium measured by  $r^2$ , is also indicated.

**Table 4** Phenotype Summary by Ethnicity

Sample	Mean DSM-IV-dependence symptom counts (n)				
	All	CEU	AA	Hispanic	American Indian
All	1.4 ± 1.8 (1051)	1.5 ± 1.9 (860)	1.2 ± 1.7 (132)	0.9 ± 1.5 (28)	1.3 ± 1.7 (12)
Smokers	2.9 ± 1.7 (488)	3.0 ± 1.7 (412)	2.7 ± 1.7 (56)	2.1 ± 1.8 (12)	2.3 ± 1.7 (7)
Current frequent smokers	3.3 ± 1.6 (295)	3.4 ± 1.6 (245)	2.9 ± 1.7 (40)	1.8 ± 0.8 (5)	2.8 ± 1.8 (5)

Abbreviations: CEU, Caucasian; AA, African American.

mean symptom count of  $1.4 \pm 1.8$ . Within ‘current frequent smokers’, the mean symptom count was  $3.3 \pm 1.6$ . No significant differences in DSM-IV symptoms were seen between the sexes, nor between ethnicities. Hispanic individuals tended to endorse fewer symptoms than Caucasians and African Americans, but this was not significant. As expected, number of quit attempts and nicotine dependence were correlated with an  $r^2 = 0.35$  ( $p < 0.001$ ) in the whole sample and  $r^2 = 0.30$  ( $p < 0.001$ ) within current frequent smokers.

**Individual SNP Frequencies in Different Ethnic Groups**

The allele frequencies in the NYS were similar to those reported in the literature (Table 5). There was evidence for differences in SNP minor allele frequencies among ethnic groups, to the extent that the minor allele was different in Caucasians and African Americans. This dramatic difference in allele frequencies was seen in rs7004381, rs4950, rs13280604, rs35389610, rs892413, and rs1072003 ( $\chi^2$ , all  $p < 0.01$ ). Only rs4953 showed no evidence of differences between allele frequencies in African Americans vs Caucasians ( $p = 0.30$ ). The representation of other ethnic groups, such as Hispanics, is too small in this sample to obtain accurate allele frequencies and test for frequency differences. One SNP, rs35389610, was too rare to be informative

in a sample of this size (MAF = 0.006) and was excluded from all genetic association analyses.

Block structure and inter-SNP correlations in the NYS were similar to those available from HapMap. Owing to the large differences in allele frequencies, LD between SNPs was examined separately in African Americans and Caucasians. The patterns of LD and correlations are shown in Figure 1. In *CHRNA6*, rs892413 and rs1072003 are in high LD with an  $r^2$  of 0.75 in Caucasians and  $r^2$  of 0.39 in African Americans. These SNPs also form a weak block with rs2304297. Although rs35389610 is located between rs2304297 and rs892413, it is only polymorphic in Caucasians and so rare that the estimated  $r^2$  between it and any other SNP is effectively 0. SNPs rs7004381 and rs4950 upstream of *CHRN3* also showed high LD with  $r^2$  of 0.80 in Caucasians and 0.47 in African Americans. All SNPs were in HWE within self-reported ethnic groups.

**Genetic Association**

*Primary analysis.* Association tests focusing on overall nicotine dependence and number of quit attempts were performed using PBAT on the group of ‘current frequent smokers’. Significant associations in the NYS-FS are shown in Table 6. In cases where the  $p$ -value within Caucasians is different than across ethnic groups it is given. In examining

**Table 5** Allele Frequencies and Ethnic Differences in the NYS-FS

Reference sequence (rs)	Location	Gene	Position chromosome 8	Relative position	MAF sample	MAF CEU	MAF AA	Minor allele CEU	Minor allele AA	p-value ethnicity difference
rs7004381	Upstream	CHRNA6	42 670 318	0	0.285	0.227	0.399	A	G	0.0084
rs4950	3' UTR	CHRNA6	42 671 790	1472	0.291	0.215	0.258	G	A	<0.0001
rs13280604	Intron 1	CHRNA6	42 678 743	8425	0.298	0.218	0.299	G	A	<0.0001
rs4953	Exon 5	CHRNA6	42 706 816	36 498	0.041	0.03	0.06	C	C	0.3
rs2304297	Exon 6	CHRNA6	42 725 148	57 038	0.295	0.237	0.271	C	G	<0.0001
rs35389610	Intron 5	CHRNA6	42 726 954	58 690	0.006	0.007	0	T	C	—
rs892413	Intron 2	CHRNA6	42 727 356	63 217	0.267	0.212	0.308	A	C	<0.0001
rs1072003	Intron 2	CHRNA6	42 729 008	68 840	0.236	0.186	0.345	G	C	<0.0001

Abbreviations: CEU, Caucasian; AA, African American.

**Table 6** Association Results for Significant SNP–Phenotype Associations ‘Current Frequent Smokers’<sup>a</sup>

Gene	SNP	No. of informative families	Risk allele	Three or more dependence symptoms <sup>b</sup>		Quit attempts p-value <sup>b</sup>	
				Additive	Recessive	Additive	Recessive
CHRNA6	rs7004381	58	G	0.60	0.39	<b>0.011</b>	<b>0.0024</b>
CHRNA6	rs4950	47	A	0.85	0.98	<b>0.030</b>	<b>0.021</b>
CHRNA6	rs13280604	49	A	0.65 ( <b>0.037</b> )	0.72 (0.09)	0.074 ( <b>0.036</b> )	<b>0.024</b>
CHRNA6	rs4953	13	G	0.24	0.25	<b>0.030</b>	<b>0.031</b>
CHRNA6	rs2304297	43	G	0.91 ( <b>0.0066</b> )	0.56 ( <b>0.018</b> )	0.55 ( <b>0.0044</b> )	0.43 ( <b>0.013</b> )
CHRNA6	rs892413	46	—	0.72	0.68	0.96	0.64

p-values less than 0.05 shown in bold.

<sup>a</sup>p-values for association within current frequent smokers reporting at time of interview at least 180 days of smoking in past 6 months.

<sup>b</sup>p-value within Caucasians given in parentheses.

overall nicotine dependence, rs2304297 showed association in Caucasians ( $p=0.0066$ ). In the case of number of quit attempts, three SNPs at the upstream end of *CHRNA6* showed significant association under a recessive model (rs7004381,  $p=0.0024$ ; rs4950,  $p=0.021$ ; rs13280604,  $p=0.024$ ). Marker rs2304297 in the 3' region of *CHRNA6* was associated also with number of quit attempts in Caucasians ( $p=0.0044$ ). Using a Bonferroni correction for multiple testing (8 SNPs and 2 phenotypes = 16 tests), the cutoff for a 5% significance level would be a  $p$ -value <0.003. Only the association of rs7004381 with quit attempts is significant at this level. However, it is worth noting that this cutoff is overly conservative due to both correlations between the SNPs (LD) and between the two phenotypes.

**Examination of inclusion criteria.** As subject inclusion and recruitment criteria vary widely among genetic association studies, we carried out further analyses to examine the effects of the inclusion criteria on the significance of the observed genetic associations. Figure 2 clearly shows that  $p$ -values are different within different subsets of the sample. In some cases, the more stringent inclusion criteria yielded stronger genetic associations, whereas in other cases association was most significant in the whole sample. This set of analyses was performed last and conducted separately

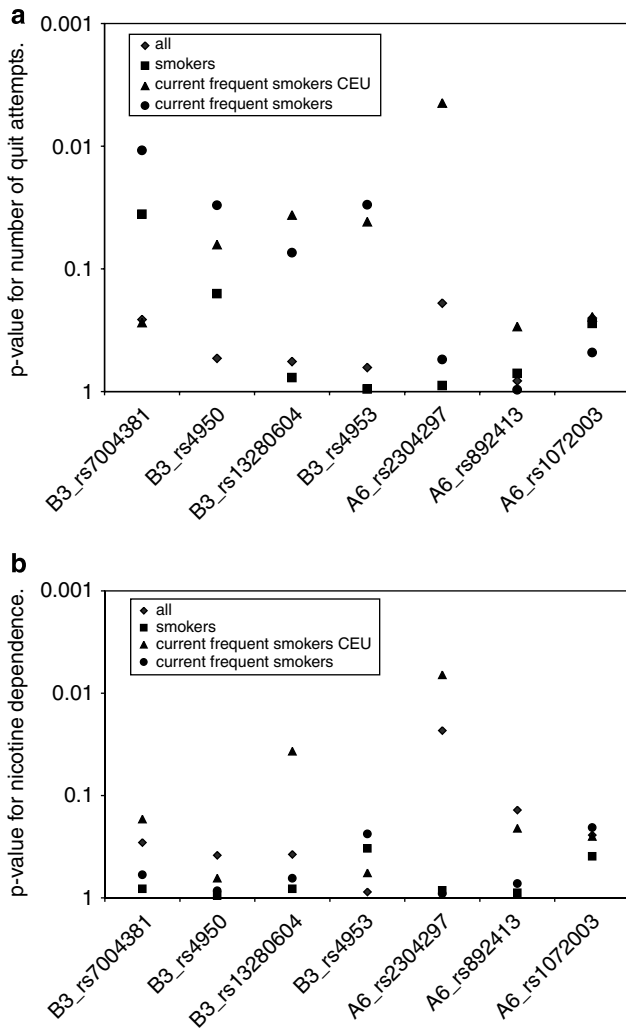
to examine the patterns of  $p$ -values across inclusion criteria. Individual  $p$ -values from these plots are not adjusted for multiple testing, because the goal was to examine the pattern, not to extract genetic significance of a single SNP.

**Singletons.** Unrelated individuals who could not be included in the PBAT analyses were used to confirm the results. These singletons were examined for genetic association of SNPs with quit attempts and with overall dependence, using simple  $\chi^2$  and ANOVA analysis in SPSS. The left-most two columns of Table 7 show the results of these analyses in the current frequent smokers. rs7004381, rs4950, and rs13280604 in *CHRNA6* were found to be associated with overall nicotine dependence using a recessive model ( $p=0.039$ , 0.007, and 0.006), with corresponding odds ratios of 3.36 (1.01–11.11) for rs7004381, 4.62 (1.44–14.84) for rs4950, and 4.67 (1.46–14.95) for rs13280604. Fisher's method was used to combine the  $p$ -values from the above-described PBAT sibpairs analysis with the results from the singletons, shown in the second set of columns in Table 7 (combined). In addition to boosting the significance of the *CHRNA6* SNP rs4950 to  $p=0.022$ , rs2304297 in *CHRNA6* was also nominally associated with tobacco dependence ( $p=0.011$ ). Results of analysis of all subjects are shown in the right half of Table 7. In the singletons, there was modest evidence for rs2304297 using

an additive model ( $p=0.028$ ), and modest evidence for rs13280604 and rs892413 with a recessive model ( $p=0.044$  and  $0.024$ ). After combining  $p$ -values with the PBAT sibpair

results, there is some evidence for association of rs2304297 in *CHRNA6* ( $p=0.005$ ).

Number of quit attempts was not normally distributed so the data were log-transformed in an attempt to achieve normality and satisfy homoscedasticity assumptions necessary for ANOVA, which were then sufficiently, but not perfectly, satisfied. rs7004381, rs4950, and rs13280604 showed a trend toward association with number of quit attempts in the singletons. For example, for rs7004381 individuals with AA genotype averaged  $1.38 \pm 1.16$  quit attempts vs  $5.77 \pm 2.22$  in GG individuals; similarly at rs4950 individuals with GG genotype averaged  $0.47 \pm 0.27$  quit attempts vs  $4.38 \pm 1.74$  in AA individuals, and at rs13280604 individuals with CC genotype averaged  $1.2 \pm 0.82$  quit attempts vs  $5.81 \pm 2.81$  in AA genotype individuals. However, these failed to be significant due to a low response rate for this question, which greatly reduced power (fewer than 100 individuals within current frequent smokers).



DISCUSSION

Here we provide additional support for a role of the *CHRNA6* and *CHRNA3* genes in mediating nicotine dependence. This study has used a population-representative sample to show that three SNPs, rs7004381, rs4950, and rs2304297, are associated with tobacco dependence. To our knowledge, this is the first study to show evidence for association of variation in *CHRNA6* and *CHRNA3* with dependence measured by DSM-IV symptom endorsement. Therefore, results presented here suggest that association of tobacco dependence with these genes is present using different diagnostic criteria across samples, thus strengthening the evidence that they are likely to contribute to nicotine-related behaviors.

In the context of this study, it is useful to compare results for specific SNPs, phenotypes, and samples to those that have been previously reported. To facilitate this process, Supplementary Table 1 summarizes results from the four other publications to compare with new data presented here. The fourth column gives the allele of each SNP present in the most common haplotype, using data from the caucasian HapMap sample. It is important to recognize that some of the studies included subjects from other ethnic backgrounds, including hispanics in the CADD sample and

**Figure 2** Association results for number of quit attempts (a) and nicotine dependence (b) and across the *CHRNA6-CHRNA3* region. The y axis shows  $p$ -values for each SNP across the x axis for each of the three different sample subsets.

**Table 7** Association Results for Tobacco Dependence Within Singletons and Combined

Gene	SNP	Current frequent smokers					All				
		Singletons		N	Combined		Singletons		N	Combined	
		Additive	Recessive		Additive	Recessive	Additive	Recessive		Additive	Recessive
<i>CHRNA3</i>	rs7004381	0.059	<b>0.039</b>	109	0.055	0.060	0.30	0.056	410	0.299	0.118
<i>CHRNA3</i>	rs4950	0.071	<b>0.007</b>	101	0.185	<b>0.022</b>	0.34	0.45	375	0.393	0.307
<i>CHRNA3</i>	rs13280604	0.661	<b>0.006</b>	105	0.114	<b>0.004</b>	0.50	<b>0.044</b>	388	0.500	<b>0.044</b>
<i>CHRNA3</i>	rs4953	—	—	—	—	—	—	—	—	—	—
<i>CHRNA6</i>	rs2304297	0.232	0.188	107	<b>0.011</b>	<b>0.023</b>	<b>0.028</b>	0.89	392	<b>0.005</b>	0.107
<i>CHRNA6</i>	rs892413	0.095	0.091	110	0.097	0.099	0.051	<b>0.024</b>	398	<b>0.042</b>	<b>0.015</b>

Abbreviation: NS, not significant.  
 $p$ -values less than 0.05 shown in bold.

African Americans in the Add Health sample, where allele frequencies were found to differ. Therefore, although the risk alleles for individual SNPs across studies are in agreement for much of this region, there are some instances where they differ. It is possible that certain alleles of specific SNPs exert different effects among different ethnic groups, consistent with the fact that in animal models, the effect of a gene deletion may differ depending on genetic background (Crabbe *et al*, 2006). In this NYS sample of adults, results from the DSM-IV analyses in the sibpairs sample suggest that SNP rs2304297 in *CHRNA6* is associated with nicotine dependence (in Caucasians) and number of unsuccessful quit attempts. This SNP has been examined for its involvement in nicotine dependence in all of the other studies. Greenbaum *et al* (2006) did not find any evidence for association, as measured by Fagerstrom test for nicotine dependence (FTND). However, results from a genome-wide association (Bierut *et al*, 2007), as well as a candidate gene study (Saccone *et al*, 2007) in the same sample, revealed a highly significant association between FTND-defined nicotine dependence and the same risk allele associated in this paper (G). Although Zeiger *et al* (2008) did not find evidence for association with nicotine dependence, there was suggestive evidence for association between rs2304297 and early subjective response to tobacco in both samples studied. Each of these samples consisted of adolescents/young adults (mean ages of  $18.21 \pm 1.5$  and  $22.4 \pm 1.7$ ), as was the relatively small sample in the Greenbaum *et al* (2006) paper (mean age  $23.3 \pm 2.3$ ), which may include individuals who are still in an 'experimental' phase of smoking and have not yet tried to quit. The NYS sample studied here, as well as the NICSNP sample used in Saccone *et al* (2007) and Bierut *et al* (2007), includes older adults (mean ages  $39 \pm 3$  and  $45 \pm 8$ , approximated from Table 6, p 45 in the Saccone *et al* (2007) paper). Therefore, of the five samples where rs2304297 was investigated, two have found evidence for early subjective response to tobacco (Zeiger *et al*, 2008), two have found evidence for nicotine dependence (Saccone *et al*, 2007; present study), and one did not find any evidence of association (Greenbaum *et al*, 2006).

In the *CHRNA6* gene, SNP rs7004381 was found to be associated with number of unsuccessful quit attempts, which is correlated with, and a component of, nicotine dependence measured by DSM-IV. rs7004381 is in close proximity to rs6474413 and rs6474412 (97 and 663 bp downstream, respectively), rs6474413 was one of the top associations in the Saccone *et al* (2007) candidate genes study and rs6474412 was within the top 40 associations out of 2 400 000 markers in the Bierut *et al* (2007) study. This SNP of *CHRNA6* is further upstream of the gene than any of those included the Greenbaum *et al* (2006) or the Zeiger *et al* (2008) studies. SNPs rs4950 and rs13280604, which are moderately associated with number of unsuccessful quit attempts in this NYS sample, have been examined only in the current study and the Zeiger *et al* (2008) study. Although that study did not find evidence for association with dependence, these two SNPs were significantly associated with early subjective response to tobacco. It is possible that the lack of association with nicotine dependence may be related to the fact that both samples were composed of adolescents and young adults. Number of unsuccessful quit attempts was also modestly associated

with rs4953. Although rare, this SNP was also found to be moderately associated with dependence in the Saccone *et al* (2007) study and the Zeiger *et al* (2008) study.

Although *CHRNA6* and *CHRNA3* are adjacent to each other, it is biologically plausible that the  $\alpha 6$  and  $\beta 3$  subunits may affect slightly different components of nicotine addiction. The  $\beta 3$  subunit is found in receptors lacking  $\alpha 6$ , and has no acetylcholine-binding site. nAChR receptors with  $\beta 3$  localize with  $\alpha 3$  and  $\beta 4$  in the interpeduncular nucleus and medial habenula, whereas receptors with  $\beta 3$ ,  $\alpha 6$ , and  $\beta 2$  are found in the substantia nigra, VTA, striatum, and locus coeruleus (Gotti *et al*, 2006a, b). Likewise,  $\alpha 6\beta 2$ -containing receptors (lacking  $\beta 3$ ) are found in the superior colliculus and have been shown to be involved in GABA release (Gotti *et al*, 2006a, b). Analysis of quit attempts suggests that this measure may have a strong underlying genetic component that can be explained in part by the *CHRNA6* and *CHRNA3* genes (Table 6).

In the singletons sample, the most significant associations with dependence were found for three SNPs in the *CHRNA3* gene (rs7004381, rs4950, and rs13280604). Evidence for an association with the *CHRNA6* gene was less consistent in the singletons (Table 7). However, although the sample size for the quit attempt variable in singletons was quite small, directional trends emerged that were consistent with findings in the families. The consistency of the results across this region further supports the findings in the larger set of families.

Our initial analysis focused on the subset of 'current frequent smokers', which included individuals who reported using tobacco every day in the 6 months before interview. Our hypothesis for selection criteria seems to be supported because the strongest genetic association was detected in the comparison of dependence symptoms among daily smokers. Specifically, the mean number of symptoms in current frequent smokers with the genotype AA at rs7004381 is  $2.2 \pm 0.37$  vs  $3.4 \pm 0.17$  symptoms in GG individuals.

Reports in the literature support this idea that comparing exposed, regular users may be one way to tease out genes which contribute to dependence. For example, in the recent studies by Bierut *et al* (2007) and Saccone *et al* (2007) who used a minimum of smoking 100 or more times, they found highly significant associations with *CHRNA3* SNPs and nicotine dependence. However, other studies that used more inclusive criteria did not find results that were as significant (Greenbaum *et al*, 2006). Replication of associations between specific SNPs and complex traits has been a challenge for researchers investigating and interpreting these studies. Although lack of replication may be due to the initial finding being a false positive, there are many other sources of inconsistency between findings, including the definition of the phenotype as well as inclusion criteria of the sample. However, it is important to keep in mind that using highly stringent inclusion criteria can greatly reduce power due to reduced sample size and phenotypic heterogeneity. For this reason, we conducted a secondary analysis of our data whereby different subject inclusion criteria were evaluated to determine the effect of sample selection on the ability to detect association.

It is important to comment on the differences in allele frequencies that are present in Caucasians and African

Americans, because ethnicity may be an important factor in the biological relevance of these SNPs. Even though all ethnic groups could be analyzed together when using the family-based PBAT analysis that controls for stratification, the association of rs2304297 became more significant when Caucasians were analyzed separately. This may be due to differences in the size of the LD blocks in Caucasians vs African Americans. There is only a small probability that rs2304297 is 'the functional' allele, so the strength of association will depend on the level of LD between rs2304297 and the actual disease allele. On the basis of data from the HapMap project NCBI Build 35 The International HapMap Project (2003) as well as the results of genotyping presented here, the LD block that contains rs2304297 extends beyond exon 4 of *CHRNA6* in Caucasians but is weaker in African Americans. In addition, as mentioned above, the differences between the results within ethnicities may be affected by genetic background, a well-documented factor in genetically engineered mouse studies (Crabbe et al, 2006), but relatively unexamined in humans.

On the basis of the local sequence around these SNPs, there are several putative functional elements that may be affected by different alleles. The 5' UTR of *CHRNA6* shows evidence for a variety of putative transcription factor binding sites (as predicted by Transcriptional Element Search System: <http://www.cbi.upenn.edu/tess>). In particular rs4950 appears to be in close proximity to a binding site of the ubiquitous transactivator AP1 (ID: T00029). In the case of *CHRNA6*, there are numerous predicted transcription factor-binding sites, particularly near the SNPs in intron 2, but none of note near rs2304297. Polymorphisms such as rs2304297 would be more likely to affect expression through trafficking of mRNA, miRNA binding, and/or rate of degradation as it is the 3' UTR (Brant-Zawadzki et al, 2007; Maziere and Enright, 2007).

We have provided evidence for the involvement of  $\alpha 6$  and  $\beta 3$  in nicotine dependence measured by DSM-IV in a population-representative sample, further supporting the results in the literature for association of dependence defined by the FTND with these genes in selected samples (Bierut et al, 2007; Greenbaum et al, 2006; Saccone et al, 2007). Furthermore, in our secondary analysis using different subsets of the sample based on phenotype definition, we also showed that sample selection can greatly affect the ability to detect association. Given the collection of studies that have now provided evidence for an association between these genes and nicotine behaviors, it will be important for future studies to focus on understanding the molecular functional consequences of different SNPs. This should include work aimed at identifying all of the variation in the region, followed by *in vitro* and *in vivo* approaches to study the effects of these variants. Ultimately, it will take a careful integration of results from human genetic studies with laboratory-based methods to tease apart the complex underlying genetic mechanisms that contribute to risk for nicotine dependence.

#### ACKNOWLEDGEMENTS

This work was supported by NIH grants AA015336 (MAE), AA007464 (NRH), DA017637 (IS, NRH), and AA11949-03 (NYS-FS).

#### DISCLOSURES/CONFLICTS OF INTEREST

Nicole R Hoft, Isabel R Schlaepfer, Robin P Corley, Matthew B McQueen, David Huizinga, and Marissa A Ehringer do not have any potential conflicts of interest, financial or otherwise, relevant to the subject matter of this work.

#### REFERENCES

- Anchordoquy HC, McGeary C, Liu L, Krauter KS, Smolen A (2003). Genotyping of three candidate genes after whole-genome preamplification of DNA collected from buccal cells. *Behav Genet* 33: 73–78.
- Barrett JC, Fry B, Maller J, Daly MJ (2005). Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21: 263–265.
- 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.
- Brant-Zawadzki PB, Schmid DI, Jiang H, Weyrich AS, Zimmerman GA, Kraiss LW (2007). Translational control in endothelial cells. *J Vasc Surg* 45(Suppl A): A8–A14.
- Champtiaux N, Gotti C, Cordero-Erausquin M, David DJ, Przybylski C, Lena C et al (2003). Subunit composition of functional nicotinic receptors in dopaminergic neurons investigated with knock-out mice. *J Neurosci* 23: 7820–7829.
- Cottler LB, Keating SK (1990). Operationalization of alcohol and drug dependence criteria by means of a structured interview. *Recent Dev Alcohol* 8: 69–83.
- Crabbe JC, Phillips TJ, Harris RA, Arends MA, Koob GF (2006). Alcohol-related genes: contributions from studies with genetically engineered mice. *Addict Biol* 11: 195–269.
- Cui C, Booker TK, Allen RS, Grady SR, Whiteaker P, Marks MJ et al (2003). The beta3 nicotinic receptor subunit: a component of alpha-conotoxin MII-binding nicotinic acetylcholine receptors that modulate dopamine release and related behaviors. *J Neurosci* 23: 11045–11053.
- Elliott DS, Huizinga D, Menard SW (1989). *Multiple Problem Youth: Delinquency, Substance Use, and Mental Health Problems*. Springer-Verlag: New York.
- Fulker DW, Cherny SS, Sham PC, Hewitt JK (1999). Combined linkage and association sib-pair analysis for quantitative traits. *Am J Hum Genet* 64: 259–267.
- Gotti C, Riganti L, Vailati S, Clementi F (2006a). Brain neuronal nicotinic receptors as new targets for drug discovery. *Curr Pharm Des* 12: 407–428.
- Gotti C, Zoli M, Clementi F (2006b). Brain nicotinic acetylcholine receptors: native subtypes and their relevance. *Trends Pharmacol Sci* 27: 482–491.
- Greenbaum L, Kanyas K, Karni O, Merbl Y, Olender T, Horowitz A et al (2006). Why do young women smoke? I. Direct and interactive effects of environment, psychological characteristics and nicotinic cholinergic receptor genes. *Mol Psychiatry* 11: 312–322, 223.
- Haberstick BC, Smolen A (2004). Genotyping of three single nucleotide polymorphisms following whole genome preamplification of DNA collected from buccal cells. *Behav Genet* 34: 541–547.
- Hymowitz N, Cummings KM, Hyland A, Lynn WR, Pechacek TF, Hartwell TD (1997). Predictors of smoking cessation in a cohort of adult smokers followed for five years. *Tob Control* 6(Suppl 2): S57–S62.
- Jarvis MJ (2004). Why people smoke. *BMJ* 328: 277–279.
- Lange C, DeMeo D, Silverman EK, Weiss ST, Laird NM (2004). PBAT: tools for family-based association studies. *Am J Hum Genet* 74: 367–369.



- Li MD, Cheng R, Ma JZ, Swan GE (2003). A meta-analysis of estimated genetic and environmental effects on smoking behavior in male and female adult twins. *Addiction* **98**: 23–31.
- Maes HH, Sullivan PF, Bulik CM, Neale MC, Prescott CA, Eaves LJ *et al* (2004). A twin study of genetic and environmental influences on tobacco initiation, regular tobacco use and nicotine dependence. *Psychol Med* **34**: 1251–1261.
- Maziere P, Enright AJ (2007). Prediction of microRNA targets. *Drug Discov Today* **12**: 452–458.
- Mineur YS, Picciotto MR (2008). Genetics of nicotinic acetylcholine receptors: relevance to nicotine addiction. *Biochem Pharmacol* **75**: 323–333.
- Rebhan M, Chalifa-Caspi V, Prilusky J, Lancet D (1998). GeneCards: a novel functional genomics compendium with automated data mining and query reformulation support. *Bioinformatics* **14**: 656–664.
- Rhee SH, Hewitt JK, Young SE, Corley RP, Crowley TJ, Stallings MC (2003). Genetic and environmental influences on substance initiation, use, and problem use in adolescents. *Arch Gen Psychiatry* **60**: 1256–1264.
- 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.
- Salminen O, Murphy KL, McIntosh JM, Drago J, Marks MJ, Collins AC *et al* (2004). Subunit composition and pharmacology of two classes of striatal presynaptic nicotinic acetylcholine receptors mediating dopamine release in mice. *Mol Pharmacol* **65**: 1526–1535.
- Swan GE, Carmelli D, Rosenman RH, Fabsitz RR, Christian JC (1990). Smoking and alcohol consumption in adult male twins: genetic heritability and shared environmental influences. *J Subst Abuse* **2**: 39–50.
- The International HapMap Consortium (2003). The International HapMap Project. *Nature* **426**: 789–796.
- US Department of Health and Human Services. Reducing Tobacco Use: A Report of the Surgeon General. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, 2000.
- Zeiger JS, Haberstick BC, Schlaepfer I, Collins AC, Corley RP, Crowley TJ *et al* (2008). The neuronal nicotinic receptor subunit genes (*CHRNA6* and *CHRNA3*) are associated with subjective responses to tobacco. *Hum Mol Genet* **17**: 724–734.
- Zhang L, Cui X, Schmitt K, Hubert R, Navidi W, Arnheim N (1992). Whole genome amplification from a single cell: implications for genetic analysis. *Proc Natl Acad Sci USA* **89**: 5847–5851.

Supplementary Information accompanies the paper on the Neuropsychopharmacology website (<http://www.nature.com/npp>)