An Association Study of DRD5 With Smoking Initiation and Progression to Nicotine Dependence

Department of Psychiatry, Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, Virginia

A large body of genetic epidemiological data strongly implicate genetic factors in the etiology of smoking behavior. Polymorphisms of genes in the dopaminergic system are plausible functional candidate genes and a linkage and an association study suggested that the type 5 dopamine receptor gene (DRD5) may be etiologically involved. We investigated the association of four DRD5 polymorphisms with smoking initiation and progression to nicotine dependence in a population-based sample of over 900 subjects. For smoking initiation, there was no significant association with the four DRD5 markers we studied; however, maximum likelihood analyses suggested the presence of a haplotype protective against smoking initiation. For progression to nicotine dependence, there were no strongly significant associations with the four DRD5 markers or for the estimated haplotypes. These data are not consistent with a strong etiological role for DRD5 in the etiology of these complex smoking behaviors.

KEY WORDS: DRD5; Smoking; nicotine dependence

INTRODUCTION

Tobacco smoking is an enormous public health problem that is associated with considerable morbidity, mortality, and personal and public cost [U.S. Department of Health and Human Services, 1989; World Health Organization, 1997]. For example, in the United States, cigarettes are responsible for 30% of deaths due to cancer and 21% of deaths from cardiovascular disease [U.S. Department of Health and Human Services, 1989]. Half of those beginning to smoke in adolescence will die from a cigarette-related cause [World Health Organization, 1997]. Costs of medical care attributable to smoking in the United States were conservatively estimated to be $50 billion in 1993 and the true total may approach $100 billion [Centers for Disease Control, 1994].

Smoking is a familial trait [Green, 1979; U.S. Department of Health and Human Services, 1989; Conrad et al., 1992], and one adoption [Eaves and Eysenck, 1980] and numerous twin studies [Sullivan and Kendler, 1999] have demonstrated that genetic factors account for most of the observed familiality. Genetic factors may be particularly important for nicotine dependence and account for approximately 67% of the variance in liability [Sullivan and Kendler, 1999].

A large body of research implicates the mesolimbic dopaminergic pathway projecting from the ventral tegmental area to the nucleus accumbens in reward-related incentive learning. A major hypothesis is that commandeering [Dani and Heinemann, 1996] of this pathway by certain drugs of abuse results in the positively and negatively reinforcing effects characteristic of addictive drugs [Koob, 1996]. There is direct evidence that nicotinic acetylcholine receptors are expressed on dopaminergic neurons in the ventral tegmental area [Pidoplichko et al., 1997], leading to the hypothesis that genetic polymorphisms in components of the dopaminergic system are candidate genes for nicotine dependence.

There are at least nine proteins in the dopaminergic system that are plausible candidate genes for nicotine dependence. These include the five G protein–coupled dopamine receptors (DRD1–5), the presynaptic dopamine reuptake transporter, dopamine B-hydroxylase (which converts dopamine to norepinephrine), and the catabolic enzymes catechol-O-methyl transferase and monoamine oxidase B.

In this report, we investigate the association of DRD5 with smoking initiation and progression to nicotine dependence in the context of a large, population-based case control study. Several lines of evidence suggest that the D1-like dopaminergic receptors (DRD1 and
DRD5 [Missale et al., 1998] are particularly important components of the mesolimbic reward-related incentive learning pathway [Beninger and Miller, 1998] and, moreover, DRD5 binds dopamine 10 times more avidly than DRD1 [Sunahara et al., 1991]. It is important to note, however, that DRD5 is apparently not expressed at high levels in the brain regions thought to be critical to reward-related incentive learning [Jaber et al., 1996; Missale et al., 1998]. Most importantly, genetic linkage and association data suggest that DRD5 may be a positional candidate gene. Vanyukov et al. [1998] reported a significant association between the microsatellite polymorphism D5(CT/GT/GA)$_n$ located about 55 kb from the DRD5 gene [Sherrington et al., 1993] with a heterogeneous set of psychoactive substance dependencies. In our linkage study of nicotine dependence [Straub et al., 1999], there was a trend toward statistical significance for nonparametric linkage scores for D5(CT/GT/GA)$_n$ in the main sample (pointwise $P = 0.096$), although not in the replication sample ($P = 0.24$).

**MATERIALS AND METHODS**

**Subjects**

The sample reported here is entirely distinct from our prior linkage study of nicotine dependence [Straub et al., 1999]. The subjects for this study were chosen from two large, population-based twin investigations whose structure and sampling procedures are described elsewhere [Kendler and Prescott, 1999]. We studied a subset of these twins who were of European ancestry; a randomly selected member of a dizygotic pair; and had returned a buccal epithelial cell sample for DNA extraction. Ancestry was determined via self-report; any twin who reported any non-European ancestry (e.g., Amerindian, Asian, African) was excluded. All of these individuals were unrelated, and all provided informed consent.

Smoking data were collected in a similar manner in each of the twin studies [Kendler et al., 1999b] and included questions about basic smoking history, the Fagerstrom tolerance questionnaire (FTQ) [Fagerstrom, 1978], and withdrawal symptoms. The FTQ is an eight-item scale (range, 0–11) that is widely used in the smoking literature and that assesses the degree of dependence on nicotine. Scores of seven or more are consistent with nicotine dependence [Fagerstrom and Schneider, 1989]. The time frame for the FTQ was the subject’s lifetime period of maximum cigarette use.

We defined three groups of individuals who were lifetime nonsmokers who reported never having smoked an entire cigarette; regular smokers who evidenced a low degree of nicotine dependence (defined as an FTQ score of 0–3 in the lifetime period of maximum use) despite having smoked regularly for at least 5 years; and lifetime regular smokers with a high degree of nicotine dependence (FTQ score of 7–11 in the lifetime period of maximum use). The current conceptualization of the process of becoming dependent on nicotine includes at least two critical steps [U.S. Department of Health and Human Services, 1989; Kendler et al., 1999b]: beginning to smoke regularly and the development of nicotine dependence contingent upon beginning to smoke regularly. Therefore, we defined two dichotomous comparisons: smoking initiation (never-smoked vs. regular smokers with low and high FTQ scores) and progression to nicotine dependence (regular smokers with low vs. high FTQ scores).

**DNA Extraction**

DNA was obtained during direct interview by having the subjects collect buccal epithelial cells using standard cytology brushes (Fisher Scientific) [Richards et al., 1993]. DNA preparation was done from four brushes per subject using the Instagene Purification Matrix (Biorad Laboratories). The procedure simultaneously lyses the buccal cells, lyses the nuclei, and the majority of the non-DNA contaminants are adsorbed to the matrix, which is pelleted, leaving the chromosomal DNA free in solution. We prepare two tubes of DNA per subject (600 ml each) and made working stocks at a dilution of 1:50.

**DRD5 Markers**

The DRD5 gene has been localized to chromosome 4p15.1–p15.3 [Eubanks et al., 1992; Sherrington et al., 1993] and encodes a 477 amino acid product [Sunahara et al., 1991]. Although there is a small intron in the 5’ untranslated region, the translated portion of the gene lacks introns [Beischlag et al., 1995]. Numerous sequence variants have been described [Sherrington et al., 1993; Sommer et al., 1993; Sobell et al., 1995; Beischlag et al., 1996; Feng et al., 1998] and the functional consequences of many of these have been investigated [Cravchik and Gejman, 1999]. There are no known DRD5 sequence variants with altered expression or pharmacological properties that have allele frequencies > 0.01. Therefore, we chose to test four polymorphisms in or near the DRD5 gene with the highest heterozygosities (even though none is known to result in altered DRD5 expression or function) to increase our chances of detecting linkage disequilibrium between one of these markers and a mutation of etiological importance to smoking initiation and progression to nicotine dependence.

Except as noted below, touchdown thermal cycling conditions [Hecker and Roux, 1996] were used for the PCR with initial annealing temperatures specific for each primer pair. For example, touchdown 65 (TD65) cycling conditions were incubation for 10 min at 95°C, and then cycles of 1 min–2 min–1 min at temperatures (°C) of 94-75-72, 94-74-72, 94-73-72, 94-72-72, 94-71-72, 94-70-72, 94-69-72, 94-68-72, 94-67-72, 94-66-72; 30 cycles at 94-65-72; and then 6 min at 72°C.

For the (TC) repeat polymorphism (promoter; positions 1393–1418 in Genbank U21164), PCR analysis was performed using primers VCU-P626 (5’-ATC-CACCACCTGGGCTCCAAA-3’) and 32P-radiolabeled VCU-P627 (5’-GTCCCCATATGTTTGTACCC-3’) under thermal cycling conditions described by Beischlag et al. [1996]. Heterozygosity was 0.354. The available data do not suggest that this polymorphism
has major functional consequences for DRD5 transactivation [Beischlag et al., 1996]. For the T978C polymorphism [Sobell et al., 1995] (exon 1; position 1125 in Genbank M67439; Pro326Pro), PCR was performed using primers VCU-P240 (5'-CCA-GCAGCTGGGCAACACCTTCTGA-3') and 32P-radiolabeled primer VCU-P239 (5'-GCCCATGTCATGGTCGCGCTTGGCA-3') under TD65 cycling conditions. The PCR product was 739 bp. After digestion with Eco57I, allele 1 (C, frequency ~40%) was 652 bp and allele 2 (T) was 530 bp. Heterozygosity was 0.473. T978C is a silent mutation within the coding region [Sobell et al., 1995; Feng et al., 1998].

For the C1481T polymorphism (3' UTR; position 1628 in Genbank M67439), PCR was performed using primer VCU-P249 (5'-GTCCCTTTTCGACTGGAACCCCTTG-3') and 32P-radiolabeled VCU-P250 (5'-CACACCATATCTCCCTCTTCATAGGT-3') under TD65 cycling conditions. The PCR product was 739 bp. After digestion with AlwN I, allele 1 (T, frequency ~31%) was 579 bp and allele 2 (C) was 216 bp. Heterozygosity was 0.438.

For the D5(CT/GT/GA)n repeat polymorphism, primers and PCR conditions were as in Sherrington et al. [1993]. Based on the location of the repeats in the BAC clone containing the pseudogene DRD5P1 (positions 24172-24234 in Genbank AC006453) located in 2p11.2-p11.1, this polymorphism is probably about 55 kb downstream from the C1481T polymorphism in DRD5. Heterozygosity was 0.801.

**Stratification Analyses**

Population stratification is a well-known threat to the validity of case control association studies. Recently, Pritchard and Rosenberg [1999] described a test for population stratification based on marker data and showed that a minimum of 15 unlinked microsatellite markers was generally sufficient to detect population stratification. The population stratification test is computed by summing the chi-square statistic from each of the case control by marker tables (with df equal to the sum of the df from the individual tables) and determining the associated $P$ value. Therefore, we generated genotypes at 16 unlinked microsatellite markers (heterozygosity mean, 0.73; range, 0.62–0.85) in order to assess the presence of population stratification.

**Statistical Analysis**

We used the Pearson chi-square test to compare the allele distributions for each of the four markers for groups defined by smoking initiation and progression to nicotine dependence. Because D5(CT/GT/GA)n has multiple rare alleles, the expected cell frequencies were often less than five, rendering the Pearson chi-square test suspect. We used the chi-square permutation test implemented in CLUMP [Sham and Curtis, 1995] to generate empirical $P$ values.

We estimated statistical power [Schlesselman, 1982] for allelic comparisons assuming a type I error rate of 0.05 and genetic homogeneity. The allele frequencies of the four DRD5 markers in the control groups for smoking initiation and progression to nicotine dependence ranged from approximately 0.2 to 0.4. Given the sample size for smoking initiation, there was 80% power to detect the effect of a predisposing allele if its frequency in cases was on the order of 11% greater than controls (relative risk ~1.6) and to detect the effect of a protective allele if its frequency in cases was ~10% less than controls (relative risk ~0.60). For progression to nicotine dependence, there was 80% power to detect the effect of a predisposing allele if its frequency in cases was ~13% greater than controls (relative risk ~1.7) and to detect the effect of a protective allele if its frequency in cases was ~12% less than controls (relative risk ~0.45).

Haplotype frequencies in the three groups of subjects were estimated by maximum likelihood using Mx [Neale et al., 1999]. The method was based on the approach described by Sham [1998] for two linked biallelic loci to be estimated in a single group. Using Mx, the method was extended to cover haplotypes based on four biallelic loci and to perform planned comparisons to test two hypotheses: that the estimated DRD5 haplotypes of regular smokers with low and high FTQ scores did not differ (a test of progression to nicotine dependence), and that lifetime nonsmokers did not differ from the two groups of regular smokers (a test of smoking initiation). Initially, a saturated model was fit with all three groups having separate haplotype frequencies. The first hypothesis was tested by equating haplotype frequencies for the low- and high-nicotine-dependent groups while allowing the nonsmokers to have separate haplotype frequencies (model A). The difference of fit between this model and the saturated model is asymptotically distributed as chi-square with degrees of freedom equal to the difference in the number of parameters equaled. The second hypothesis was tested by constraining all three groups to have the same haplotype frequencies and comparing the fit of this model to that of model A. The script for these analyses can be obtained from the Mx web page (views.vcu.edu/mx).

**RESULTS**

**Subject Characteristics**

Table I shows the characteristics of subjects in the three basic groups that were used to test the key comparisons suggested by the literature. For smoking initiation, we contrast lifetime never-smokers with regular smokers with low and high FTQ scores. For progression to nicotine dependence, we contrast regular smokers with low and high FTQ scores. It is evident from Table I that the nicotine-dependent regular smokers manifest the key elements of nicotine dependence and that the nondependent regular smokers do not.

**Stratification Analyses**

Based on data from 16 unlinked microsatellite markers, there was no significant evidence of population
stratification across groups defined by smoking initiation (stratification chi-square $= 134.2, df = 117, P = 0.13$) and for groups defined by progression to nicotine dependence (stratification chi-square $= 110.1, df = 107, P = 0.40$).

**Linkage Disequilibrium**

Using an analytic approach described elsewhere [Kendler et al., 1999a], we determined that all four markers were in significant linkage disequilibrium ($P < 0.001$ for all pairwise comparisons).

**DRD5 Marker Comparisons**

For smoking initiation, as shown in Table II, there were small differences in T978C genotype frequencies that approached statistical significance ($P = 0.08$), although the allelic distributions were similar. For D5(CT/GT/GA)$_n$, the allelic distributions for smoking initiation approached statistical significance (CLUMP $P = 0.05$). However, because 31% of the overall chi-square came from cells with counts less than 10, this result is unlikely to hold substantive importance. Finally, to allow direct comparison of our data with those of Vanyukov et al. [1998], we compared the modal 148 bp 5 D5(CT/GT/GA)$_n$ allele vs. all other alleles and found no differences in allele frequencies across the smoking initiation groups. For progression to nicotine dependence, only the genotypic comparison for the dichotomized D5(CT/GT/GA)$_n$ marker reached nominal statistical significance ($P = 0.04$).

When all four markers [promoter (TC)*2 vs. all other alleles, T978C, C1481T, and D5(CT/GT/GA)$_n$*5 vs. all other alleles] were entered into multiple logistic regression analyses, there were no significant associations with smoking initiation or progression to nicotine dependence (both with and without control for subject gender; data not shown).

CART (classification and regression trees) [Breiman et al., 1984; MathSoft, 1998] is a recursive partitioning technique that might be particularly useful in detecting epistatic interactions [Rao, 1998]. When we used CART with fivefold cross-validation [Stone, 1974; Efron and Tibshirani, 1993] for this purpose, the actual smoking initiation phenotype agreed poorly with that predicted from the four DRD5 markers ($k = 0.13$). Similar results were found for progression to nicotine dependence ($k = 0.15$). These multivariate results are consistent with the univariate results in Table II.

**Symptom-Level Analyses**

We also investigated whether any of the four markers were associated with certain smoking traits. In those who had ever regularly smoked, there was no significant association of any of the four DRD5 markers with current smoking status, cigarettes per day smoked at maximum, and minutes from awakening to first cigarette.

**Haplotype Analyses**

We entered the three physically closest DRD5 markers into these analyses (end of promoter TC–3,241 bp–T978C–502 bp–C1481T). D5(CT/GT/GA)$_n$ was not included as it is approximately 55 kb downstream from C1481T. We contrasted the *2 allele with all other alleles for the promoter TC polymorphism. The maximum likelihood haplotype estimates and 95% confidence intervals for the three groups (never-smokers, nondependent regular smokers, and nicotine-dependent regular smokers) are shown in Table III.
The 95% confidence intervals for the three groups appear to overlap with two possible exceptions (haplotypes "other-C-C" and "other-C-T").

Our first hypothesis was that the eight DRD5 haplotype frequencies were equal across the three groups; consistent with inspection of Figure 1, this hypothesis could be rejected (chi-square $\hat{\chi}^2_{25.81}$, df $\hat{\chi}^2_{14}$, P $\hat{P}_{0.027}$). Second, as a test of progression to nicotine dependence, we tested the hypothesis that the haplotype frequencies were equal between nondependent and nicotine-dependent regular smokers; this hypothesis could not be rejected (chi-square $\hat{\chi}^2_{10.12}$, df $\hat{\chi}^2_{7}$, P $\hat{P}_{0.18}$). Finally, as a test of smoking initiation, we tested the hypothesis that the haplotype frequencies of the lifetime nonsmoking group were similar to the other two groups (nondependent and nicotine-dependent regular smokers); this hypothesis could be rejected (chi-square $\hat{\chi}^2_{15.69}$, df $\hat{\chi}^2_{7}$, P $\hat{P}_{0.028}$).

DISCUSSION

The goal of this investigation was to assess the evidence that the DRD5 gene was etiologically involved in the complex behaviors of smoking initiation and progression to nicotine dependence in the context of a...
large case control association study. The central dopamine system clearly contains candidate genes for these two behaviors and the available evidence suggests that DRD5 may be both a functional and positional candidate gene.

For smoking initiation, there was no significant association with the four DRD5 markers we studied. The D5(CT/GT/GA)n allele frequencies approached nominal statistical significance ($P = 0.05$), but a considerable fraction of the overall chi-square came from rare alleles, thus diminishing the likelihood of its importance. However, there was statistical evidence that the haplotype defined by an allele other than the *2 allele at the promoter TC, the C-allele at T978C, and the T-allele at C1481T was protective against smoking initiation.

For progression to nicotine dependence, there were no strongly significant associations with the four DRD5 markers or for the estimated haplotypes. When we compared the D5(CT/GT/GA)n genotypes created by dichotomizing by the presence of the modal 5 allele, there was a statistically significant association ($P = 0.04$).

It is possible that the significant findings represent true positive results. However, the statistical significance was marginal particularly when the number of statistical comparisons we performed is taken into account. Therefore, we think that it is more likely that genetic variation in DRD5 is not substantially related to smoking initiation and progression to nicotine dependence in our sample.

There are several caveats to this conclusion. First, although there was good statistical power for the case control comparisons, our power was poorer for the detection of small effects and for haplotype estimation given the absence of parental genotypes to establish phase with greater precision. Second, while we found no compelling evidence for the involvement of DRD5 as a main effect on the two smoking phenotypes, it remains plausible that DRD5 could be etiologically involved via epistatic interactions. Third, we did not exclude the possibility that the rare functional mutations in DRD5 [Cravchik and Gejman, 1999] are of etiological relevance to smoking initiation and progression to nicotine dependence. Fourth, although the FTQ has been reported to possess excellent test-retest correlations [Pomerleau et al., 1994], its adaptation from a current to a lifetime maximum time frame has not been validated.

REFERENCES


