Human Molecular Genetics, 2006, Vol. 15, No. 24 ***3498–3507***
doi:10.1093/hmg/ddl426
Advance Access published on November 3, 2006

Haplotype spanning *TTC12* and *ANKK1*, flanked by the *DRD2* and *NCAM1* loci, is strongly associated to nicotine dependence in two distinct American populations

Joel Gelernter[1,2,3,4,*], Yi Yu[5], Roger Weiss[10,11], Kathleen Brady[12], Carolien Panhuysen[5,6], Bao-zhu Yang[1,4], Henry R. Kranzler[13] and Lindsay Farrer[5,6,7,8,9]

[1]Division of Human Genetics, Department of Psychiatry, [2]Department of Neurobiology, [3]Department of Genetics and [4]VA CT Healthcare Center, Yale University School of Medicine, VA CT 116A2, 950 Campbell Avenue, West Haven, CT 06516, USA, [5]Department of Medicine (Genetics Program), [6]Department of Biostatistics, [7]Department of Neurology, [8]Department of Genetics & Genomics, and [9]Department of Epidemiology, Boston University Schools of Medicine and Public Health, Boston, MA, USA, [10]Alcohol and Drug Abuse Treatment Program, McLean Hospital, Belmont, MA, USA, [11]Department of Psychiatry, Harvard Medical School, Boston, MA, USA, [12]Department of Psychiatry, Medical University of South Carolina, Charleston, SC, USA and [13]Department of Psychiatry, University of Connecticut School of Medicine, Farmington, CT, USA

Received August 24, 2006; Revised and Accepted October 26, 2006

Nicotine dependence (ND) is a moderately heritable trait. We ascertained a set of 1615 subjects in 632 families [319 African-American (AA) and 313 European-American (EA)] based on affected sibling pairs with cocaine or opioid dependence. Subjects were interviewed with the Semi-Structured Assessment for Drug Dependence and Alcoholism (SSADDA). Previously, we identified a modest linkage peak (LOD score =1.97) for ND in the EA part of the sample on chromosome 11q23, a region that includes the *NCAM1–TTC12–ANKK1–DRD2* gene cluster. *DRD2* and *NCAM1* are functional candidate genes for substance dependence; the *TTC12* and *ANKK1* loci are not well characterized. We genotyped a set of 43 single nucleotide polymorphisms (SNPs) spanning this region, and performed family-based association and haplotype analysis. There was relatively weak evidence for association of the flanking *DRD2* and *NCAM1* markers to ND, but very strong evidence of association of multiple SNPs at *TTC12* and *ANKK1* in both populations (minimal $P = 0.0007$ in AAs and minimal $P = 0.00009$ in EAs), and in the pooled sample, as well as strong evidence for highly significant association of a single haplotype spanning *TTC12* and *ANKK1* to ND in the pooled sample ($P = 0.0000001$). We conclude that a risk locus for ND, important both in AAs and EAs, maps to a region that spans *TTC12* and *ANKK1*. Functional studies of these loci are warranted. These results provide additional information useful in evaluating the many earlier discrepant findings regarding association of *DRD2* with substance dependence.

INTRODUCTION

Nicotine dependence (ND) is highly prevalent and highly destructive in the US and throughout the world. In 2004, it was reported that 67.3% of the US population (aged 12 and over) had ever smoked cigarettes, and 29.1% had smoked cigarettes in the past year (1). ND and many behaviors related to it, like other forms of substance dependence, have been demonstrated to be heritable. Li *et al*. (2) conducted a meta-analysis of twin studies for phenotypes related to smoking, and found that smoking initiation heritability could be estimated as 0.37 for males and 0.55 for females;

*To whom correspondence should be addressed. Tel: +1 2039325711, ext 3590; Fax: +1 2039374741; Email: joel.gelernter@yale.edu

smoking persistence heritability was estimated as 0.59 for males and 0.46 for females. Increased understanding of genetic risk factors for ND would be of great social and economic value. Identification of risk loci has been addressed so far through linkage methodology [2, review of studies up to 2004; (3–9)], and through association approaches. Several candidate gene associations have been replicated [e.g. *DDC*, dopa-decarboxylase (10,11)].

There is also a long history of reports of association of the D2 dopamine receptor (*DRD2*) locus with substance dependence. Dopamine has been implicated in the reinforcing effects of a variety of drugs of abuse, including nicotine (12). In concentrations achieved by smokers, nicotine activates and desensitizes nicotinic receptors and regulates the activity of mesolimbic dopamine neurons (13). D2 dopamine receptors are largely distributed throughout limbic structures (14), consistent with a prominent role of this receptor type in motivated behavior, including addiction. Association studies of variants that map at or near *DRD2* with ND have yielded variable findings, with most failing to show association. A meta-analysis of 12 studies (2) showed that overall there was a significant association (pooled OR = 1.47). Another meta-analysis of 13 studies of this association (15) showed significant heterogeneity among the studies, and no association when a random-effects model (which accounts for between-study differences) was applied. Overall, the possibility of an association at this locus has not been resolved definitively, and deserves to be addressed in an adequately powered and rigorously assessed sample.

We collected a set of small families suitable for linkage and family-based association analyses of cocaine and opioid dependence (16,17). These subjects were evaluated with the Semi-Structured Assessment for Drug Dependence and Alcoholism (SSADDA) (16,18), a polydiagnostic instrument that provides for DSM-IV ND, and in which the Fagerstrom Test for ND (FTND) (19,20) is embedded. This family set proved to be highly informative for studies based on ND and the FTND as well, although ascertainment was not conditioned on smoking (9,11).

The neural cell adhesion molecule (*NCAM1*) locus maps molecularly close to *DRD2* (21). *NCAM1* is also an attractive candidate locus; its protein product regulates the secretion of neurotransmitter as a membrane-bound glycoprotein that mediates cell–cell contact by homophilic interactions (22). *NCAM1* has been associated genetically with neural tube defects, bipolar affective disorder and schizophrenia (23–25), and with adverse effects of nicotine on brain plasticity (26). On the basis of the hypothesis that *NCAM1* was also a gene that could influence risk of addiction, we genotyped sets of single nucleotide polymorphism (SNP) markers at *DRD2* and *NCAM1*. In the course of our study, it was reported that two additional genes, tetratricopeptide repeat domain 12 (*TTC12*) and ankyrin repeat and kinase domain containing 1 (*ANKK1*), map between *DRD2* and *NCAM1*, so we extended our work to provide coverage of these loci also. Thus, to test the hypothesis that variants in *NCAM1*, *DRD2* or genes located between them, namely *TTC12* and *ANKK1*, are associated with smoking behavior, we evaluated 43 SNP markers spanning this region in our family sample.

Several noteworthy findings emerged from our genomewide linkage study of ND and FTND in this cohort, including a possible locus on chromosome 11 (at D11S908), where we observed a LOD score of 1.97 in the European-American (EA) part of the sample for ND (9). Under this LOD score peak is a 600 kb region including the *NCAM1*, *TTC12*, *ANKK* and *DRD2* genes, which we will refer to as a 'cluster' (without intending any implication beyond their close physical localization). This finding (noted after SNP genotyping for this project had been completed) greatly increased our already-strong interest in this candidate region, in our specific clinical sample. Initially, interest was focused on *NCAM1* and *DRD2* as candidate loci, as discussed earlier. The protein encoded by *ANKK1* is one of a family of proteins involved in signal transduction pathways. The so-called '*DRD2 Taq*I A' system causes an amino acid substitution within the 11th ankyrin repeat of this gene (Glu713Lys), which may affect substrate-binding specificity (27). *TTC12* is still undercharacterized.

Thus, the three of the four genes mapped to this cluster that are characterized (at least to some extent) functionally, may all be considered candidate loci for regulating risk for substance dependence. Prior, although ambiguous, genetic association data, added to this *a priori* support; and our own linkage data were also consistent with a risk locus in this region. Our analysis, however, led us to an unexpected conclusion; we discovered *weak* evidence of association of ND with SNPs located within *DRD2* and *NCAM1*, but *strong* and *consistent* evidence of association of ND with markers within *TTC12* and *ANKK1*. A specific haplotype comprising SNPs that span parts of both of these genes was strongly associated with ND in sub-samples of African-American (AA) and EA families.

RESULTS

Single SNP analysis

In general, there were many nominally-significant results for DSM-IV ND, and fewer for the FTND trait. We found a moderate correlation between the diagnosis of ND and FTND score (AAs: Spearman's $\rho = 0.521$; EAs: $\rho = 0.591$; combined: $\rho = 0.575$), indicating that most of the variance in these measures is not shared. These findings are consistent with the differential association findings for these traits. For AAs, eight markers showed at least nominally-significant association to ND and two to FTND. For EAs, 15 markers showed at least nominally-significant association to ND and seven to FTND. For the entire (pooled) sample, 13 markers showed at least nominally-significant association to ND and four to FTND. All of these findings, but one in the pooled sample ($P = 0.04$ for SNP3 with ND), were observed between SNP17 (in *NCAM1*) and SNP40 (in *DRD2*). Of these 24 SNPs, spanning 201 179 bp, all but six showed at least 'trend-level' (i.e. $P \leq 0.10$) significance for association to ND in the EAs, and all but eight did so in the pooled sample.

Nearly all variants that were statistically significant after correction cluster at *TTC12* and *ANKK1*; one SNP meeting this threshold (rs6277) is located in *DRD2* and is the most proximal marker to *ANKK1* studied at that locus.

This variant, C957T, has been shown previously to be functional (28); moreover, it has been reported to regulate the effect of nicotine on working memory (29). However, the association in our study is seen only in EAs, where rs6277 is in strong LD with the markers in *TTC12* and *ANKK1*, which are significantly associated with ND. There is no association of this marker with phenotype in AAs, nor is the marker in LD with markers in the other genes in this population. Thus, it is possible, and more consistent with our other observations, that this association reflects LD with markers in TTC12 or ANKK1.

In EAs, the most statistically significantly associated markers (besides rs6277) map to *ANKK1*: rs4938012 ($P = 0.0003$), rs4938013 ($P = 0.0004$), rs4938015 ($P = 0.00009$) and rs11604671 ($P = 0.0007$). In AAs, only the first two of these reach even nominal significance, and the most statistically significantly associated marker is rs2303380 located in *TTC12* ($P = 0.0007$ in AAs and $P = 0.007$ in EAs).

In the combined (EA plus AA) sample, numerous markers that tend to cluster at the *TTC12/ANKK1* border show increased evidence of association. The most strongly associated markers include two in TTC12, rs2303380 ($P = 0.00002$) and rs2282511 ($P = 0.00009$); and the next three markers in ANKK1, rs4938012 ($P = 0.000008$), rs4938013 ($P = 0.00003$) and rs4938015 ($P = 0.0006$). One possible explanation for this pattern, which is consistent with LD as observed in this sample, is that the findings in AAs and EAs are driven by the same functional variant or variants, and the difference in single SNP associations is due to population-based LD differences with that variant or set of variants. If this explanation is correct, the functional variant would be expected to be localized near the 3′ region of *TTC12* and/or the 5′ region of *ANKK1*.

Haplotype association analysis

Analysis of SNPs individually indicated that the region most likely to contain an ND susceptibility locus is bounded by *TTC12* rs2303380 centromerically and extends into *ANKK1* but not into *DRD2*. Haplotype analysis was conducted to help narrow the location of a susceptibility locus, and determine whether a single functional variant could explain the pattern of association findings with individual SNPs in each population group. As a first step, we evaluated LD among the 43 SNPs in the cluster of four genes in order to reduce the number of markers that would be potentially informative for haplotype analysis. This analysis, shown in Figure 1, revealed slightly more extensive LD in EAs than in AAs, and these population-specific patterns are consistent with the LD structure reported in the HapMap database for this genetic region (30). There was a considerable amount of intergenic LD, particularly in the EAs.

We observed a total of 10 LD blocks in EAs and nine LD blocks in AAs. The block structures are remarkably similar in these two population groups; in fact, three blocks accounting for nine SNPs are identical. Taking into account the LD block structure (Fig. 1) and the association findings with individual SNPs (Table 1), four SNPs (rs2303380, rs4938012, rs4938015 and rs11604671) were selected for haplotype analysis. These markers include four of the six responsible for the most significant results in the combined sample, and account for the potentially uniquely important information from each LD block spanning the distal portion of *TTC12* and all of *ANKK1* in both EAs and AAs.

The haplotype including this SNP combination was significantly associated with ND in AAs (global $P = 0.0003$), EAs (global $P = 0.0008$) and the combined sample of families (global $P = 0.000005$) (Table 2). The specific haplotype G-A-T-C was associated with ND in EAs ($P = 0.00006$), AAs ($P = 0.0008$) and in the pooled sample ($P = 0.0000001$). Thus, these results showing strong evidence for association of the same haplotype to ND in two distinct populations suggest the possibility of a single causative variant. Haplotypes A-G-C-T and A-G-T-C were associated with decreased risk of ND in EAs ($P = 0.001$) and AAs ($P = 0.0009$), respectively. The divergent frequencies of these protective haplotypes in these two population groups may explain the observation of distinct association peaks, at rs2303380 in AAs and near rs4938013 in EAs. This SNP combination was also modestly associated with FTND score, but in EAs only (global $P = 0.04$). This result is attributable primarily to the haplotype G-A-T-C ($P = 0.007$), which was associated with increase in FTND score (data not shown).

Correlation of trait with cocaine and opioid dependence

Since the sample was recruited for cocaine and opioid dependence, we examined the correlations between these diagnoses and ND. The correlations were low (ND–CD: Spearman's $\rho = 0.190$; ND–OD, $\rho = 0.166$). The modest association between ND and these other diagnoses argues against the possibility that the allelic and haplotypic associations reported here for ND can be explained on the basis of confounding by comorbid CD or OD. Further, we did not observe noteworthy evidence of association with the DSM-IV diagnoses of CD or OD in the specific region implicated for ND. However, our ability to detect association with these traits in this sample is less than that for ND, since the sample was ascertained for linkage studies [i.e. affected sibling pairs (ASPs)] with CD and OD (but not ND), whereas discordant sibling pairs are more informative for family-based association studies.

DISCUSSION

We present here strong evidence for association of markers that span portions of *TTC12* and *ANKK1* to ND in two distinct populations, EAs and AAs. Although the single SNP results for the two populations favor somewhat different localizations for a risk locus, with the strongest results in *ANKK1* for EAs and in *TTC12* for AAs, both populations show multiple at least nominally significant associations at both loci; and pooled results show multiple significant (when corrected) associations at both loci. One haplotype that comprised SNPs from both loci is significantly associated in each population group, and is highly significantly associated in the pooled population sample ($P = 0.0000001$). All of these results derive from family-based analyses, and thus cannot easily be attributed to population stratification artifact. We therefore conclude

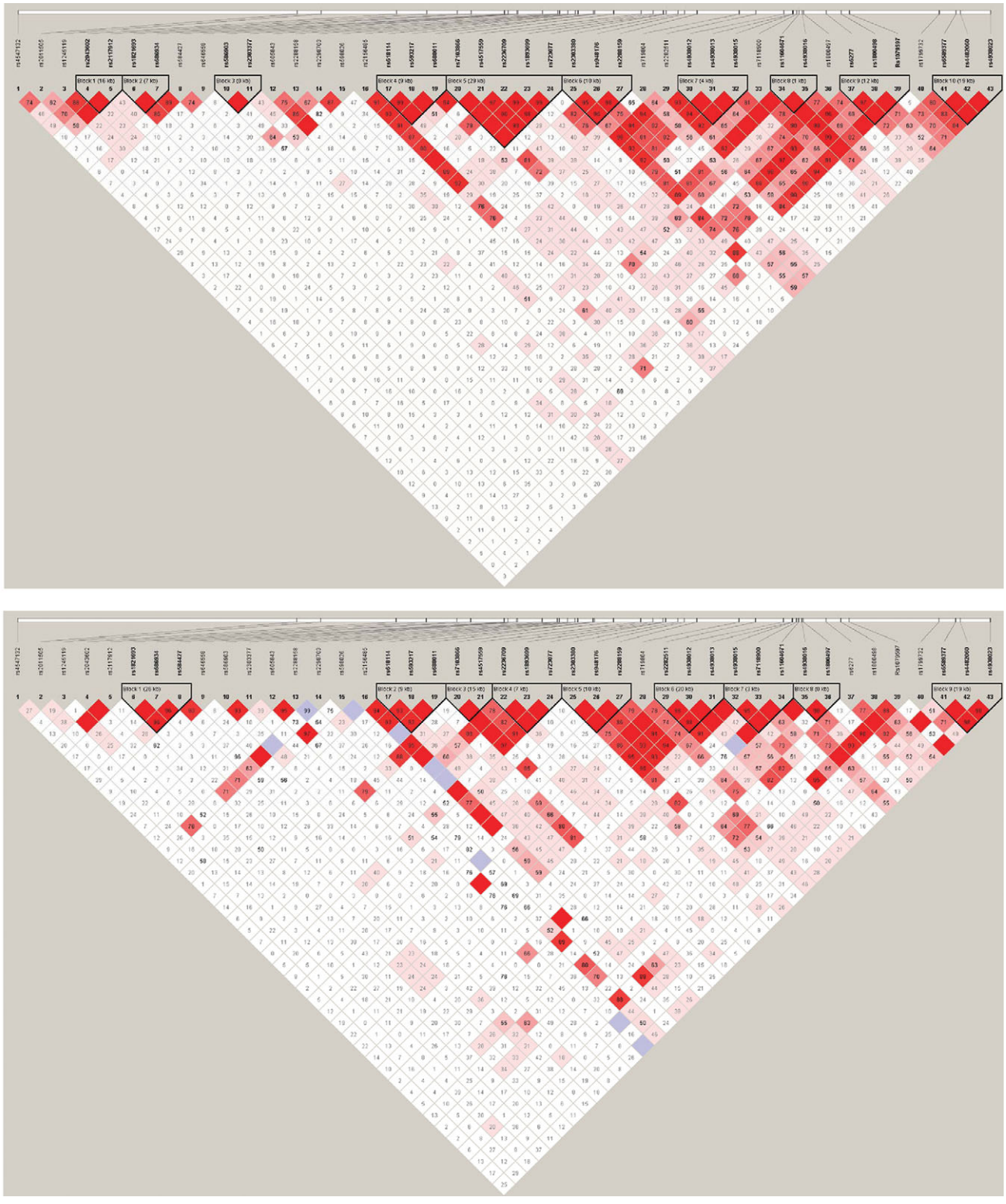

Figure 1. LD patterns in (**A**) EA and (**B**) AA subjects.

that a variant influencing risk for ND maps to this region, probably within *ANKK1* or *TTC12*, and probably not within *DRD2* or *NCAM1*. This does not exclude the presence of risk loci of smaller effect in either of these flanking loci, but in this data set, there are much stronger results supporting a risk locus in *ANKK1* or *TTC12*.

DRD2 is well known in psychiatric genetics for many early claims of association (mostly to alcohol dependence)

Table 1. Results of family based analysis for the single SNPs

Locus	Marker number	dbSNP rs number	AA		EA		Pooled	
			TobDep	Fag_res	TobDep	Fag_res	TobDep	Fag_res
NCAM1	1	rs4547132	0.99	0.87	0.95	**0.04**	0.95	0.11
	2	rs2011505	0.38	1.00	0.31	0.10	0.18	0.27
	3	rs1245119	0.10	0.41	0.16	0.23	**0.03**	0.80
	4	rs2043602	0.77	0.46	0.37	0.63	0.40	0.83
	5	rs2117912	0.17	0.21	0.35	0.16	0.12	0.53
	6	rs1821693	0.18	0.85	**0.02**	0.09	0.50	0.22
	7	rs686934	0.15	0.81	**0.05**	0.13	0.74	0.18
	8	rs584427	0.81	0.98	0.17	0.21	0.35	0.26
	9	rs646558	0.52	0.66	0.61	0.79	0.83	0.61
	10	rs586903	0.67	0.91	0.14	0.86	0.57	0.94
	11	rs2303377	0.15	0.29	0.44	0.26	0.58	0.95
	12	rs605843	0.38	0.52	0.12	0.39	0.72	0.28
	13	rs2288158	0.61	0.57	0.53	0.88	0.42	0.82
	14	rs2298703	**0.03**	0.64	0.20	0.84	0.60	0.73
	15	rs598026	0.71	0.38	0.17	0.38	0.66	0.21
	16	rs2156485	0.25	**0.03**	0.81	0.96	0.69	0.52
	17	rs618114	**0.03**	0.45	0.09	0.40	**0.008**	0.27
	18	rs593217	0.18	0.71	0.08	0.36	**0.03**	0.36
	19	rs688011	0.50	0.25	0.33	0.12	0.24	0.67
	20	rs7103866	0.87	0.24	0.17	0.23	0.35	0.73
TTC12	21	rs4517559	0.15	0.28	**0.02**	**0.04**	0.58	0.48
	22	rs2236709	0.08	0.42	0.48	**0.01**	0.08	0.36
	23	rs1893699	0.32	0.51	0.07	0.11	**0.05**	0.50
	24	rs723077	0.78	0.10	**0.02**	0.08	0.13	0.67
	25	rs2303380	**0.0007**	0.62	**0.007**	**0.04**	**0.00002**	0.06
	26	rs948176	0.06	0.37	0.42	0.58	0.30	0.28
	27	rs2288159	**0.008**	0.22	0.30	0.41	0.15	0.13
	28	rs719804	0.72	0.53	**0.03**	0.15	0.07	0.47
	29	rs2282511	**0.02**	0.46	**0.002**	0.09	**0.00009**	0.07
ANKK1	30	rs4938012	**0.01**	0.46	**0.0003**	**0.03**	**0.000008**	**0.03**
	31	rs4938013	**0.02**	0.77	**0.0004**	**0.02**	**0.00003**	**0.05**
	32	rs4938015	0.48	0.54	**0.00009**	**0.02**	**0.0006**	**0.04**
	33	rs7118900	0.78	0.24	0.07	0.71	0.29	0.31
	34	rs11604671	0.93	0.95	**0.0007**	0.18	**0.004**	0.24
	35	rs4938016	0.58	0.12	**0.002**	0.48	0.09	0.55
	36	rs1800497	0.95	0.14	0.06	0.38	0.22	0.10
DRD2	37	rs6277	0.49	0.53	**0.001**	0.16	**0.014**	0.14
	38	rs1800498	0.41	0.76	**0.004**	0.17	0.05	0.20
	39	rs1079597	0.50	0.25	**0.03**	0.32	**0.04**	0.15
	40	rs1799732	**0.002**	**0.02**	0.90	0.20	**0.02**	**0.008**
	41	rs6589377	0.85	0.33	0.71	0.46	0.68	0.24
	42	rs4482060	0.50	0.87	0.81	0.25	0.81	0.45
	43	rs4938023	0.91	0.31	0.75	0.29	0.85	0.15

Please see Table 4 for SNP characteristics.

Table 2. Haplotype analysis results for the association of major haplotypes for rs2303380–rs4938012–rs4938015–rs11604671 with TobDep

Haplotype	AA (**0.0003**)*			EA (**0.0008**)*			Pooled (**0.000005**)*		
	Frequency	Number of FAM	(Z) *P*-value	Frequency	Number of FAM	(Z) *P*-value	Frequency	Number of FAM	(Z) *P*-value
A-G-C-C	0.30	52	(−1.4) 0.17	0.17	33	(−0.6) 0.56	0.25	81	(−1.5) 0.14
A-G-C-T	0.07	19	(0.6) 0.58	0.45	48	**(−3.2) 0.001**	0.23	67	**(−2.6) 0.009**
G-A-T-C	0.18	46	**(3.4) 0.0008**	0.28	45	**(4.3) 0.00006**	0.22	89	**(5.3) 0.0000001**
A-G-T-C	0.34	51	**(−3.3) 0.0009**				0.20	59	**(−2.8) 0.005**
G-G-T-C	0.10	23	(1.2) 0.24				0.06	23	(0.9) 0.32

*Global *P*-value.

sometimes based on small study samples. Rs1800497 is a variant that has been called historically '*DRD2 Taq*I A,' although it is now known to map within *ANKK1* and result in a non-synonymous base substitution (27). We found no evidence of association of this variant with ND (or FTND); we have previously studied this variant in other samples, and did not find association with alcohol or drug dependence (31–33) and argued previously that no functional relationship of this

variant to alcohol dependence had been proved (34). However, inspection of Figure 1A reveals that this marker is in LD with other variants we did find to be associated with ND, for example, $D' = 0.73$ with rs4938015, and $D' = 1.0$ with rs1604671. Thus, it is apparent that previous reports of association of substance dependence traits that included this marker only could, potentially, be attributable to LD with a more centromeric marker mapping to *ANKK1* or *TTC12*. Since this marker is also in LD with other markers at *DRD2*, we cannot exclude functional risk-influencing variants at that locus, nor can we exclude a variant 3′ to *DRD2* that has a primary regulatory effect on that gene. There is also a previous report of very strong haplotypic association of *DRD2* variants with heroin dependence (35). That study did not consider other *ANKK1* or *TTC12* variants, so it is unclear to what extent, if at all, variants at those loci could explain the observed association, or if, alternatively, *DRD2* itself is an important locus in some populations and for some substance use disorders.

We found evidence of association to the trait of DSM-IV ND, and only weak evidence of association to the related trait, FTND score. Key elements of the DSM-IV diagnostic criteria are at least three of the following symptoms in a 12-month period: (i) tolerance; (ii) withdrawal; (iii) tobacco taken in larger amounts or over a longer period than intended; (iv) persistent desire to cut down; (v) great deal of time spent acquiring tobacco or dealing with its effects; (vi) important activities given up because of tobacco; (vii) use continued despite subject's knowledge that it is causing harm (28). FTND stresses somewhat different aspects of smoking; it queries respondents regarding, for example, how soon after waking they smoke their first cigarette, which cigarette in the day would be hardest for the respondent to give up, and smoking quantity (20); these issues are not addressed directly in the DSM-IV criteria. Although both ND and FTND scores are heritable, as they measure somewhat different aspects of nicotine use and dependence, it is not surprising that there may be different contributions to each from different risk loci. Consistent with this idea, the linkage signal we observed in this region (9) was specific to ND, and so is the observed association. Conversely, we reported previously that *DDC* variants are associated much more markedly to FTND than to ND (11).

Our sample was recruited for the presence of opioid and/or cocaine dependence, not for ND, although our subjects had a high prevalence of ND and the sample proved to be very informative for the purpose of ND gene mapping (many of the subjects are polysubstance abusers). Since the sample was not recruited for ND, it would be reasonable to ask how representative this sample is for the trait in general. As we noted previously (9), we have a direct way to address the generalizability of findings from this particular sample, at least with respect to one specific locus, *DDC*. Ma *et al*. (10) reported association of several ND-related traits to markers at this locus in AAs; we (11) confirmed this finding, and extended it to EAs. The clinical sample for the study by Ma *et al*. (10) was recruited for a genetic linkage study of ND; our study (11) was conducted in the same sample as the present study. The *DDC* locus contributes to ND risk in both samples. Some of the linkage peaks we observed (9) were also consistent with those observed in previous studies. Thus, we conclude that our sample is, in a number of respects, comparable with other samples used in ND studies. Even so, we do not claim that this association is specific to ND alone; since many subjects are dependent on multiple substances of abuse, there may be some more specific predisposition to a substance or a set of substances.

Table 3. Sample characteristics

Characteristic	AA	EA	Pooled
Number of families	319	313	632
Average genotyped members/family	2.7 ± 1.0	2.4 ± 0.8	2.6 ± 1.0
Number of subjects	854	761	1615
Sex (% female)	504 (59.0)	366 (48.1)	870 (53.9)
Age (years)	40.9 ± 7.3	37.8 ± 10.2	39.4 ± 8.9
FTND score	3.7 ± 2.6	5.0 ± 2.9	4.3 ± 2.9
Cocaine dependence			
Affected (%)	676 (79.2)	559 (73.5)	1235 (76.5)
Unaffected (%)	115 (13.5)	145 (19.0)	260 (16.1)
Unknown (%)	63 (7.4)	57 (7.5)	120 (7.4)
Opioid dependence			
Affected (%)	214 (25.1)	252 (33.1)	666 (41.2)
Unaffected (%)	579 (67.8)	452 (59.4)	831 (51.5)
Unknown (%)	61 (7.1)	57 (7.5)	118 (7.3)
Tobacco dependence			
Affected (%)	470 (55.0)	515 (67.7)	985 (61.0)
Unaffected (%)	326 (38.2)	192 (25.2)	518 (32.1)
Unknown (%)	58 (6.8)	54 (7.1)	112 (6.9)
Alcohol dependence			
Affected (%)	328 (38.4)	323 (42.5)	651 (40.3)
Unaffected (%)	435 (50.9)	361 (47.4)	796 (49.3)
Unknown (%)	91 (10.7)	77 (10.1)	168 (10.4)

Our results, which were derived from Family-Based Association Test (FBAT) analysis of family samples using unaffected sibs as controls, are unlikely to be the result of stratification within a population group. Our observation of significant results in two distinct populations, also greatly supports the robustness of these results. AA populations may have >20% admixture with EA populations, with estimated EA admixture of ~20% for a New York AA population (36), reasonably comparable geographically to the population we studied, most of which was recruited in New England (17); the pattern of the results we observed (reflected in both the statistical strength of the observed effects, and in the fact that some different SNPs were associated in the EA and AA samples) supports the conclusion that the results in AAs were not driven merely by admixture from EAs.

This is one of the most robust reports of genetic association to ND to date. We studied a genomic region now supported by our own linkage findings and by some previous association analyses; however, our strongest evidence for association did not coincide with what might have been considered the best-supported candidate loci *a priori*. We cannot rule out additional risk loci that map to *DRD2* or *NCAM*, but it seems likely that, at least for the ND phenotype, the major susceptibility variant in this region is located in *ANKK1* or *TTC12*. It seems very unlikely that these genes that either clearly (*DRD2*) or presumptively (*ANKK1, NCAM1*) act in neurotransmitter pathways related to addiction are chromosomal neighbors based on chance alone; we hypothesize that the proximity of these genes in this cluster is in some way related to their physiology. Functional studies of the loci implicated

Table 4. SNP characteristics

Marker number	dbSNP rs number	GENE	SNP type (aka)	Chromosome position (bp)	Distance from previous marker (bp)	Alleles	MAF[a] (AA)	MAF[a] (EA)	MAF[a] (all)
1	rs4547132	NCAM1	Intron	112338023	.....	C/T	0.37	0.33	0.35
2	rs2011505	NCAM1	Intron	112493446	155423	A/G	0.47	0.49	0.48
3	rs1245119	NCAM1	Intron	112506871	13425	G/C	0.49(G)	0.35	0.42
4	rs2043602	NCAM1	Intron	112548936	42065	T/C	0.46	0.37	0.41
5	rs2117912	NCAM1	Intron	112565679	16743	T/C	0.07	0.21	0.15
6	rs1821693	NCAM1	Intron	112582751	17072	G/A	0.36	0.44(G)	0.47
7	rs686934	NCAM1	Intron	112590658	7907	C/T	0.36	0.42(C)	0.48
8	rs584427	NCAM1	Intron	112609206	18548	G/T	0.13	0.46	0.32
9	rs646558	NCAM1	Intron	112611117	1911	C/A	0.48(C)	0.28	0.39
10	rs586903	NCAM1	Intron	112616156	5039	A/C	0.26	0.20	0.23
11	rs2303377	NCAM1	Intron	112616711	555	T/C	0.24	0.37	0.31
12	rs605843	NCAM1	Intron	112630444	13733	T/C	0.47	0.28	0.36
13	rs2288158	NCAM1	Intron	112638886	8442	T/G	0.11	0.15	0.13
14	rs2298703	NCAM1	Intron	112639815	929	G/A	0.04	0.23	0.15
15	rs598026	NCAM1	Intron	112644460	4645	A/G	0.28	0.15	0.21
16	rs2156485	NCAM1	Intron	112648767	4307	C/T	0.05	0.26	0.17
17	rs618114	NCAM1	Intron	112650283	1516	G/A	0.24	0.39	0.33
18	rs593217	NCAM1	intergenic	112658699	8416	T/C	0.29	0.43	0.37
19	rs688011	NCAM1	intergenic	112659380	681	G/A	0.23	0.27	0.25
20	rs7103866	NCAM1	intergenic	112669978	10598	G/A	0.11	0.33	0.23
21	rs4517559	TTC12	intergenic	112685412	15434	A/G	0.50(A)	0.40	0.44
22	rs2236709	TTC12	Intron	112691986	6574	A/G	0.27	0.29	0.28
23	rs1893699	TTC12	Intron	112697734	5748	T/G	0.39	0.39	0.39
24	rs723077	TTC12	M73L	112699378	1644	A/C	0.18	0.46	0.34
25	rs2303380	TTC12	Intron	112705919	6541	A/G	0.35	0.40	0.37
26	rs948176	TTC12	Intron	112709691	3772	A/G	0.38	0.19	0.28
27	rs2288159	TTC12	Intron	112716539	6848	G/T	0.42	0.19	0.30
28	rs719804	TTC12	Intron	112739985	23446	A/G	0.23	0.24	0.23
29	rs2282511	TTC12	intergenic	112749387	9402	C/A	0.26	0.35	0.31
30	rs4938012	ANKK1	Intron	112764864	15477	G/A	0.25	0.36	0.31
31	rs4938013	ANKK1	exon (S)	112769680	4816	C/A	0.24	0.36	0.31
32	rs4938015	ANKK1	Intron	112769854	174	C/T	0.39(C)	0.39	0.49
33	rs7118900	ANKK1	A239T	112772031	2177	G/A	0.35	0.23	0.29
34	rs11604671	ANKK1	G318R	112773269	1238	C/T	0.11	0.45	0.31
35	rs4938016	ANKK1	G442R	112775225	1956	C/G	0.47(C)	0.35	0.43
36	rs1800497	ANKK1	E713K (DRD2-TaqI A)	112776038	813	G/A	0.39	0.23	0.30
37	rs6277	DRD2	exon (S) C957T	112788669	12631	A/G	0.12	0.49	0.33
38	rs1800498	DRD2	Intron (TaqI D)	112796798	8129	G/A	0.21	0.45(G)	0.40
39	rs1079597	DRD2	intron (TaqI B)	112801496	4698	G/A	0.18	0.19	0.18
40	rs1799732	DRD2	Promoter (ins/del)	112851462	49966	A/G	0.50(A)	0.15	0.31
41	rs6589377	DRD2	Intergenic	112860946	9484	A/G	0.20	0.37	0.30
42	rs4482060	DRD2	Intergenic	112863421	2475	A/T	0.32	0.42	0.38
43	rs4938023	DRD2	Intergenic	112880057	16636	G/T	0.20	0.36	0.29

[a]Minor allele frequency is the frequency of the second allele unless noted in the parenthesis.

here may, we hope, provide new insights into pathophysiological mechanisms of nicotine dependence.

MATERIALS AND METHODS

Subject recruitment and assessment

ASPs with cocaine or opioid dependence defined according to DSM-IV criteria (37) were ascertained at four clinical sites in the United States: University of Connecticut Health Center (Farmington, CT), Yale University School of Medicine (APT Foundation; New Haven, CT), McLean Hospital (Harvard Medical School; Belmont, MA) and Medical University of South Carolina (Charleston, SC). Subjects diagnosed with a major psychotic illness (e.g. schizophrenia or schizoaffective disorder) were excluded. Additional siblings and parents of the ASPs were recruited whenever possible regardless of affection status. Characteristics of the sample included are described in Table 3.

Subjects were interviewed using the SSADDA for psychiatric diagnosis as described previously (16,18). The reliability of the SSADDA for the diagnosis of ND was very good to excellent, with test–retest reliability $\kappa = 0.97$ (based on 120 subjects) and interrater reliability $\kappa = 0.77$ (based on 173 subjects) (18). The diagnosis of ND was based on application to the SSADDA data of a computer algorithm using DSM-IV diagnostic criteria. The FTND (19,20) (range of 0–10 points) is embedded in the SSADDA. Diagnostic information on tobacco or alcohol dependence was not involved in pedigree selection or extension. Subjects were classified as AA or EA

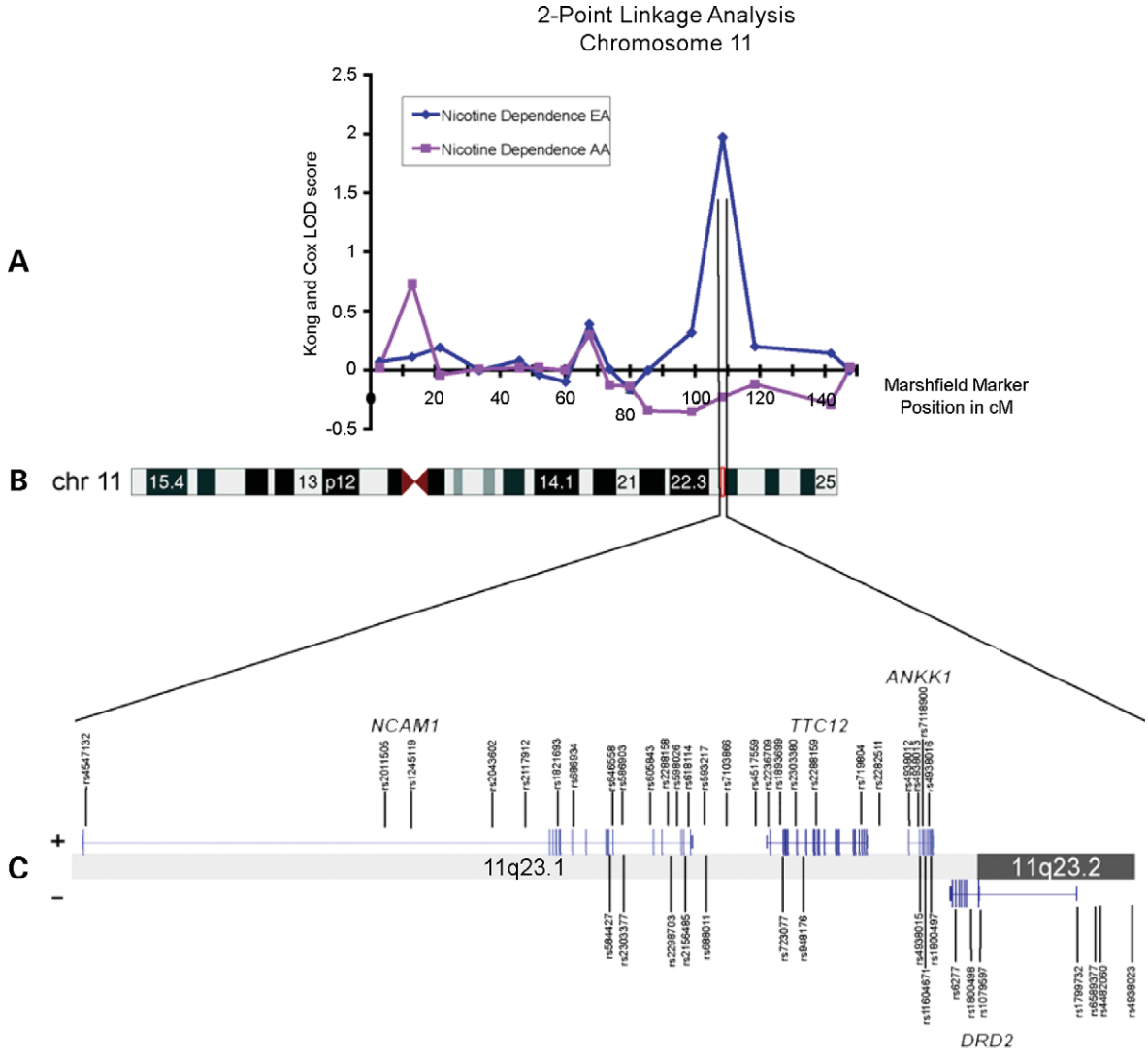

Figure 2. Nicotine dependence linkage analysis (9), and map of SNPs genotyped.

based on a Bayesian model-based clustering method using genetic marker information as previously described (9). The parts of the sample recruited at the four sites were similar in terms of percent diagnosed with ND and mean FTND scores.

DNA was obtained from immortalized cell lines in most cases, but for a small number of subjects, DNA was obtained directly from blood or saliva. Subjects gave informed consent as approved by the institutional review board at each clinical site, and a certificate of confidentiality for the work was obtained from NIH (NIDA).

SNP selection and genotyping

Forty-three SNPs covering the region containing *NCAM1*, TTC12, *ANKK1* and *DRD2* were selected, most from the ABI SNPbrowser (38) and the NCBI SNP database (http://www.ncbi.nlm.nih.gov/projects/SNP/) (Table 4). Figure 2 locates this gene cluster with respect to our previously observed linkage signal on chromosome 11. Overall, we genotyped 20 markers in *NCAM1*, nine markers in *TTC12*, seven markers in *ANKK1* and seven markers in *DRD2*. Prior published reports (for markers mapped to DRD2), information content, minor allele frequency, function potential, LD structure and validation evidence were considered in SNP selection. Detailed information about these SNPs is summarized in Table 4. SNPs were genotyped with a fluorogenic 5′ nuclease assay method, that is the TaqMan technique (39), using the ABI PRISM 7900 Sequence Detection System (ABI, Foster City, CA, USA); or using PCR-RFLP methods.

Statistical analysis

The PedCheck program was used to detect Mendelian inconsistencies in the genotype data (40). Consistency with Hardy–Weinberg equilibrium expectations (HWEE) was tested for each SNP in a set of unrelated non-smokers for each population group using a χ^2 test. One SNP (rs4938012) showed a significant deviation from HWEE in the EA

group, and also showed numerous Mendelian errors. The Taqman genotypes were therefore discarded and genotypes were obtained instead via a PCR-RFLP assay; the new data were consistent with HWEE and generated no Mendelian errors. Subsequently, a total of 16 genotyping inconsistencies were identified out of 68 370 assays for 43 SNPs across all DNA samples (i.e. $<0.02\%$ of all assays), and these results were excluded from all subsequent statistical analyses (however, our sample had limited power to detect errors in this way). Because AA and EA samples are distinct populations with different histories, the allele frequencies in some markers differ prominently between them, and there is the potential for genetic heterogeneity, analyses were initially conducted separately for each population sample. To reduce potential confounding, the FTND score was adjusted for age and sex in each population sample by computing standardized residuals using SAS (version 9.0). Pairwise LD between all markers was assessed using the Haploview program (41) and haplotypes blocks were discerned according to the criteria of Gabriel *et al*. (42).

Association of individual SNPs with ND and FTND score was evaluated using the FBAT program (43) assuming an additive model under the null hypothesis of linkage and no association. The issue of testing multiple individual markers was addressed by applying a false discovery rate correction to the empirically derived *P*-values.The experiment-wide significance threshold required to keep type I error rate at 5% was 0.002, according to SNPSpD software (44,45). HBAT (46), the haplotype extension routine in the FBAT program, was used to evaluate haplotypic associations. Each SNP included in the haplotype analysis 'tags' an LD block determined empirically for both population groups separately. Overall association for a combination of SNPs was tested using a multi-allelic global test. For significantly associated combinations of SNPs (global $P < 0.05$), haplotype-specific associations were tested using the diallelic mode in HBAT. *P*-values were estimated by the Monte Carlo sampling method under the null hypothesis of linkage and no association. The significance level of haplotype-specific tests was adjusted using Bonferroni correction, that is the α-level was divided by the number of major haplotypes with a frequency greater than 5.0%.

ACKNOWLEDGEMENTS

We thank the families who volunteered to participate in this research study. This work was supported by NIH (NIDA) grants R01 DA12690, R01 DA12849, K24 DA15105 and K02 DA00326; M01 RR06192; and the U.S. Department of Veterans Affairs (the VA Connecticut–Massachusetts Mental Illness Research, Education and Clinical Center [MIRECC]). Ann Marie Lacobelle, Lisa Naito, Lingjun Zuo and Yakov Lozovatsky provided excellent technical assistance. John Farrell provided excellent database support. We also thank the Rutgers University Cell and DNA Repository, the contractor for the NIDA Center for Genetic Studies, co-directed by Dr Jay Tischfield and Dr John Rice.

Conflict of Interest statement. None declared.

REFERENCES

1. SAMHSA, Office of Applied Studies, National Survey on Drug Use and Health, (Detailed Tables) 2003 and 2004 (http://oas.samhsa.gov/nsduh/2k4nsduh/2k4tabs/Sect2peTabs1to57.htm#tab2.1b).
2. Li, M.D., Ma, J.Z. and Beuten, J. (2004) Progress in searching for susceptibility loci and genes for smoking-related behaviour. *Clin. Genet.*, **6**, 382–392.
3. Gelernter, J., Liu, X., Hesselbrock, V., Page, G.P., Goddard, A. and Zhang, H. (2004) Results of a genomewide linkage scan: support for chromosomes 9 and 11 loci increasing risk for cigarette smoking. *Am. J. Med. Genet. B Neuropsychiatr. Genet.*, **128B**, 94–101.
4. Bierut, L.J., Rice, J.P., Goate, A., Hinrichs, A.L., Saccone, N.L., Foroud, T., Edenberg, H.J., Cloninger, C.R., Begleiter, H., Conneally, M.P. *et al*. (2004) A genomic scan for habitual smoking in families of alcoholics: common and specific genetic factors in substance dependence. *Am. J. Med. Genet. A.*, **124**, 19–27.
5. Vink, J.M., Beem, A.L., Posthuma, D., Neale, M.C., Willemsen, G., Kendler, K.S., Slagboom, P.E. and Boomsma, D.I. (2004) Linkage analysis of smoking initiation and quantity in Dutch sibling pairs. *The Pharmcogenomics J.*, **4**, 274–282.
6. Wang, D., Ma, I.Z. and Li, M.D. (2005) Mapping and verification of susceptibility loci for smoking quantity using permutation linkage analysis. *The Pharmacogenomics J.*, **5**, 166–172.
7. Ehlers, C.L. and Wilhelmsen, K.C. (2006) Genomic screen for loci associated with tobacco usage in Mission Indians. *BMC Med. Genet.*, **7**, 1–8.
8. Swan, G.E., Hops, H., Wilhelmsen, K.C., Lessov-Schlaggar, C.N., Cheng, L.S., Hudmon, K.S. Amos, C.I., Feiler, H.S., Ring, H.Z., Andrews, J.A. *et al*. (2006) A genome-wide screen for nicotine dependence susceptibility loci. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.*, **141B**, 354–360.
9. Gelernter, J., Panhuysen, C., Weiss, R., Brady, K., Poling, J., Krauthammer, M., Farrer, L. and Kranzler, H.R. Genomewide linkage scan for nicotine dependence: identification of a novel chromosome 5 risk locus. *Biol Psychiatry*. doi:10.1016/j.biopsych.2006.08.023. Published online ahead of print November 1, 2006.
10. Ma, J.Z., Beuten, J., Payne, T.J., DuPont, R.T., Elston, R.C. and Li, M.D. (2005) Haplotype analysis indicates an association between the DOPA decarboxylase (DDC) gene and nicotine dependence. *Hum. Mol Genet.*, **14**, 1691–1698.
11. Yu, Y., Panhuysen, C., Kranzler, H.R., Hesselbrock, V., Rounsaville, B., Weiss, R., Brady, K., Farrer, L.A. and Gelernter, J. (2006) Intronic variants in the DOPA decarboxylase (DDC) gene are associated with smoking behavior in European-Americans and African-Americans. *Hum. Mol. Gen.* doi:10.1093/hmg/ddl144. Advance access published on June 1, 2006.
12. Hyman, S.E., Malenka, R.C. and Nestler, E.J. (2006) Neural mechanisms of addiction: the role of reward-related learning and memory. *Annu. Rev. Neurosci*, **29**, 565–598.
13. Pidoplichko, V.I., DeBiasi, M., Williams, J.T. and Dani, J.A. (1997) Nicotine activates and desensitizes midbrain dopamine neurons. *Nature*, **390**, 401–404.
14. Hurd, Y.L., Suzuki, M. and Sedvall, G.C. (2001) D1 and D2 dopamine receptor mRNA expression in whole hemisphere sections of the human brain. *J. Chem. Neuroanat.*, **22**, 127–137.
15. Munafo, M.R., Clark, T.G., Johnstone, E.C., Murphy, M.F.G. and Walton, R.T. (2004) The genetic basis for smoking behavior: a systematic review and meta-analysis. *Nicotine Tob. Res.*, **6**, 583–597.
16. Gelernter, J., Panhuysen, C., Weiss, R., Brady, K., Hesselbrock, V., Rounsaville, B. Poling, J., Wilcox, M., Farrer, L. and Kranzler, H.R. (2005) Genomewide linkage scan for cocaine dependence and related traits: significant linkages for a cocaine-related trait and cocaine-induced paranoia. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.*, **136**, 45–52.
17. Gelernter, J., Panhuysen, C., Wilcox, M., Hesselbrock, V., Rounsaville, B., Poling, J., Weiss, R., Sonne, S., Zhao, H., Farrer, L. *et al*. (2006) Genomewide linkage scan for opioid dependence and related traits. *Am. J. Hum. Genet.*, **78**, 759–769.
18. Pierucci-Lagha, A., Gelernter, J., Feinn, R., Cubells, J.F., Pearson, D., Pollastri, A., Farrer, L. and Kranzler, H.R. (2005) Diagnostic reliability of the Semi-Structured Assessment for Drug Dependence and Alcoholism (SSADDA). *Drug Alcohol. Depend.*, **80**, 303–312.

19. Fagerstrom, K.O. (1978) Measuring degree of physical dependence to tobacco smoking with reference to individualization of treatment. *Addict. Behav.*, **3**, 35–241.
20. Heatherton, T.F., Kozlowski, L.T., Frecker, R.C. and Fagerstrom, K.O. (1991) The Fagerstrom Test for Nicotine Dependence: a revision of the Fagerstrom Tolerance Questionnaire. *Br. J. Addict.*, **86**, 1119–1127.
21. McConville, C.M., Formstone, C.J., Hernandez, D., Thick, J. and Taylor, A.M. (1990) Fine mapping of the chromosome 11q22–23 region using PFGE, linkage and haplotype analysis; localization of the gene for ataxia telangiectasia to a 5cM region flanked by NCAM/DRD2 and STMY/CJ52.75, phi 2.22. *Nucleic Acids Res.*, **18**, 4335–4343.
22. Rubinek, T., Yu, R., Hadani, M., Barkai, G., Nass, D., Melmed, S. and Shimon, I. (2003) The cell adhesion molecules N-cadherin and neural cell adhesion molecule regulate human growth hormone: a novel mechanism for regulating pituitary hormone secretion. *J. Clin. Endocrinol. Metab.*, **88**, 3724–3730.
23. Vawter, M.P., Usen, N., Thatcher, L., Ladenheim, B., Zhang, P., VanderPutten, D.M., Conant, K., Herman, M.M., van Kammen, D.P., Sedvall, G. *et al*. (2001) Characterization of human cleaved N-CAM and association with schizophrenia. *Exp. Neurol.*, **172**, 29–46.
24. Deak, K.L., Boyles, A.L., Etchevers, H.C., Melvin, E.C., Siegel, D.G., Graham, F.L., Slifer, S.H., Enterline, D.S., George, T.M., Vekemans, M. *et al*. (2005) SNPs in the neural cell adhesion molecule 1 gene (NCAM1) may be associated with human neural tube defects. *Hum. Genet.*, **117**, 133–142.
25. Arai, M., Itokawa, M., Yamada, K., Toyota, T., Arai, M., Haga, S., Ujike, H., Sora, I., Ikeda, K. and Yoshikawa, T. (2004) Association of neural cell adhesion molecule 1 gene polymorphisms with bipolar affective disorder in Japanese individuals. *Biol. Psychiatry*, **55**, 804–810.
26. Abrous, D.N., Adriani, W., Montaron, M.F., Aurousseau, C., Rougon, G., Le Moal, M. and Piazza, P.V. (2002) Nicotine self-administration impairs hippocampal plasticity. *J. Neurosci.*, **22**, 3656–3662.
27. Neville, M.J., Johnstone, E.C. and Walton, R.T. (2004) Identification and characterization of ANKK1: a novel kinase gene closely linked to DRD2 on chromosome band 11q23.1. *Hum. Mutat.*, **23**, 540–545.
28. Duan, J., Wainwright, M.S., Comeron, J.M., Saitou, N., Sanders, A.R., Gelernter, J. and Gejman, P.V. (2003) Synonymous mutations in the human dopamine receptor D2 (DRD2) affect mRNA stability and synthesis of the receptor. *Hum. Mol. Gen.*, **12**, 205–216.
29. Jacobsen, L.K., Pugh, K.R., Menci, W.E. and Gelernter, J. (2006) C957T Polymorphism of the Dopamine D2 Receptor Gene Modulates the Effect of Nicotine on Working Memory Performance and Cortical Processing Efficiency. *Psychopharmacology*. doi 10.1007/s00213-006-0469-1.
30. International HapMap Consortium (2003) The International HapMap Project. *Nature*, **426**, 789–796.
31. Gelernter, J., O'Malley, S., Risch, N., Kranzler, H., Krystal, J., Merikangas, K., Kennedy, J. and Kidd, K.K. (1991) No association between an allele at the D_2 dopamine receptor gene (DRD2) and alcoholism. *JAMA*, **266**, 1801–1807.
32. Gelernter, J., Kranzler, H. and Satel, S. (1999) No association between D_2 dopamine receptor (DRD2) alleles or haplotypes and cocaine dependence or severity of cocaine dependence in European and African Americans. *Biol. Psychiatry*, **45**, 340–345.
33. Gelernter, J. and Kranzler, H. (1999) D_2 dopamine receptor gene (DRD2) allele and haplotype frequencies in alcohol dependent and control subjects: no association with phenotype or severity of phenotype. *Neuropsychopharmacology*, **20**, 640–649.
34. Gelernter, J., Goldman, D. and Risch, N. (1993) The A1 allele at the D_2 dopamine receptor gene and alcoholism: a reappraisal. *JAMA*, **269**, 1673–1677.
35. Xu, K., Lichtermann, D., Lipsky, R.H., Franke, P., Liu, X., Hu, Y., Cao, L., Schwab, S.G., Wildenauer, D.B., Bau, C.H. *et al*. (2004) Association of specific haplotypes of D2 dopamine receptor gene with vulnerability to heroin dependence in 2 distinct populations. *Arch. Gen. Psychiatry*, **61**, 597–606.
36. Parra, E.J., Marcini, A., Akey, J., Martinson, J., Batzer, M.A., Cooper, R., Forrester, T., Allison, D.B., Deka, R., Ferrell, R.E. *et al*. (1998) Estimating African American admixture proportions by use of population-specific alleles. *Am. J. Hum. Genet.*, **63**, 1839–1851.
37. American Psychiatric Association (1994) *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn. DC American Psychiatric Press, Washington.
38. De La Vega, F.M., Isaac, H. and Scafe, C.R. (2006) A tool for selecting SNPs for association studies based on observed linkage disequilibrium patterns. *Pac. Symp. Biocomput.*, **11**, 487–498.
39. Shi, M.M., Myrand, S.P., Bleavins, M.R. and de la Iglesia, F.A. (1999) High throughput genotyping for the detection of a single nucleotide polymorphism in NAD(P)H quinone oxidoreductase (DT diaphorase) using TaqMan probes. *Mol. Pathol.*, **52**, 295–299.
40. O'Connell, J.R. and Weeks, D.E. (1998) PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am. J. Hum. Genet.*, **63**, 259–266.
41. Barrett, J.C., Fry, B., Maller, J. and Daly, M.J. (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, **21**, 263–265.
42. Gabriel, S.B., Schaffner, S.F., Nguyen, H., Moore, J.M., Roy, J., Blumenstiel, B., Higgins, J., DeFelice, M., Lochner, A., Faggart, M. *et al*. (2002) The structure of haplotype blocks in the human genome. *Science*, **296**, 2225–2229.
43. Horvath, S., Xu, X. and Laird, N.M. (2001) The family based association test method: strategies for studying general genotype–phenotype associations. *Eur. J. Hum. Genet.*, **9**, 301–306.
44. Nyholt, D.R. (2004) A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am. J. Hum. Genet.*, **74**, 765–769.
45. Li, J. and Ji, L. (2005) Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity*, **95**, 221–227.
46. Horvath, S., Xu, X., Lake, S.L., Silverman, E.K., Weiss, S.T. and Laird, N.M. (2004) Family-based tests for associating haplotypes with general phenotype data: application to asthma genetics. *Genet. Epidemiol.*, **26**, 61–69.