American Journal of Medical Genetics 124A:19–27 (2004)

# A Genomic Scan for Habitual Smoking in Families of Alcoholics: Common and Specific Genetic Factors in Substance Dependence

**Laura Jean Bierut,[1]* John P. Rice,[1] Alison Goate,[1] Anthony L. Hinrichs,[1] Nancy L. Saccone,[1] Tatiana Foroud,[2] Howard J. Edenberg,[2] C. Robert Cloninger,[1] Henri Begleiter,[3] P. Michael Conneally,[2] Raymond R. Crowe,[4] Victor Hesselbrock,[5] Ting-Kai Li,[2] John I. Nurnberger, Jr.,[2] Bernice Porjesz,[3] Marc A. Schuckit,[6] and Theodore Reich[1]**

[1]*Washington University School of Medicine, St. Louis, Missouri*
[2]*Indiana University Medical Center, Indianapolis, Indiana*
[3]*SUNY HSC at Brooklyn, Brooklyn, New York*
[4]*University of Iowa, Iowa City, Iowa*
[5]*University of Connecticut Health Center, Farmington, Connecticut*
[6]*University of California—San Diego, San Diego, California*

**Smoking is a highly heritable, addictive disorder that commonly co-occurs with alcohol dependence. The purpose of this study is to perform a genomic screen for habitual smoking and comorbid habitual smoking and alcohol dependence in families from the Collaborative Study on the Genetics of Alcoholism (COGA). Subjects were assessed using the Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA) to evaluate alcohol dependence and habitual smoking (smoking one pack per day or more for at least 6 months). Sixty seven multi-generational families with 154 independent sibling pairs affected with habitual smoking were genotyped in a screening sample. Analyses on 79 multi-generational families with 173 independent sibling pairs were repeated in a replication sample. Sibpair analyses were performed using ASPEX. Four chromosomal regions in the screening sample had increased allele sharing among sibling pairs for habitual smoking with a LOD score greater than 1 (chromosomes 5, 9, 11, and 21). The highest LOD score was on chromosome 9 (LOD = 2.02; allele sharing 58.9%). Four chromosomal regions also had modest evidence for linkage to the comorbid phenotype habitual smoking and alcohol dependence (chromosomes 1, 2, 11, 15); and the strongest finding was on chromosome 2 (LOD = 3.30; allele sharing 69.1%). Previously identified areas (chromosomes 1 and 7) implicated in the development of alcohol dependence in this same data set did not provide evidence for linkage to habitual smoking in the screening sample. In the replication data set, there continued to be increased allele sharing near peaks identified in the screening sample on chromosomes 2 and 9, but the results were modest. An area on chromosome 7, approximately 60 cM from a location previously identified in linkage analysis with alcohol dependence, had increased allele sharing for the comorbid habitual smoking and alcohol dependence. These data provide evidence of specific genetic regions involved in the development of habitual smoking and not alcohol dependence. Conversely, genetic regions that influence the development of alcohol dependence do not appear to contribute to the development of habitual smoking. Finally, there is also evidence of an area on chromosome 2 that may reflect a common genetic vulnerability locus to both habitual smoking and alcohol dependence.**

© 2003 Wiley-Liss, Inc.

Grant sponsor: The National Institute on Alcohol Abuse and Alcoholism (NIAAA) (NIH Grant); Grant numbers: U10AA08403, AA00231.

*Correspondence to: Dr. Laura Jean Bierut, M.D., Washington University School of Medicine, Department of Psychiatry, 4940 Children's Place, St. Louis, MO 63110.
E-mail: bierutl@psychiatry.wustl.edu

Received 4 March 2003; Accepted 27 March 2003
DOI 10.1002/ajmg.a.20329

© 2003 Wiley-Liss, Inc.

**KEY WORDS: family study; linkage analysis; alcoholism; smoking; comorbidity; genetic study**

## INTRODUCTION

Smoking is common in the United States, with 24% of the adult population (approximately 47 million people) classified as current smokers [CDC, 2001]. Though the prevalence of smoking has decreased over the last 30 years, this decline has slowed and there has been little change in the prevalence of current smoking since the mid-1990s. More worrisome, there continues to be high rates of smoking activity among adolescents and young adults in the United States, along with a dramatic increase in tobacco use worldwide [CDC, 2002].

Smoking is a leading cause of preventable disability and death, and there are multiple well-documented adverse health effects from smoking, such as heart disease, pulmonary disease, and cancer. The economic burden of smoking is high; $75 billion in direct medical costs and a similar amount in lost productivity [CDC, 2002]. Thus, smoking is a concern with major public health implications and economic costs.

Recently, smoking behavior has been studied from a genetic perspective. Many aspects of smoking behavior, such as initiation (ever smoked), persistence (current smoker), and nicotine dependence (Fagerström nicotine dependence and DSM-IIIR nicotine dependence), cluster in families, and this clustering is related in part to genetic factors [Carmelli et al., 1992; Heath and Martin, 1993; True et al., 1997, 1999; Kendler et al., 1999]. The heritability of nicotine dependence is estimated to be 60%. As a result, several candidate genes that may influence smoking such as nicotine metabolizing genes [Cholerton et al., 1996; Boustead et al., 1997; Pianezza et al., 1998] and neurotransmittor systems [Comings et al., 1996; Shields et al., 1998; Spitz et al., 1998; Lerman et al., 1999; Bierut et al., 2000; Lerman et al., 2001; Vandenbergh et al., 2002] have been studied, but findings have been equivocal and their role in smoking behavior remains unclear.

Individuals who smoke are also at high risk for becoming alcohol dependent, another highly heritable, addictive disorder, with a heritability of 64% [Heath et al., 1997]. Twin studies that have examined the relationship between smoking and alcohol use have demonstrated common and specific genetic factors in the use of both substances [Madden et al., 1995; Swan et al., 1996, 1997]. A twin study of the more severe phenotypes of nicotine and alcohol dependence has also supported common and specific genetic factors in the development of these disorders, and a genetic correlation estimated at 0.68 [True et al., 1999].

Since smoking and alcohol dependence are correlated diagnoses in individuals, and there is evidence of common and specific genetic factors in the development of these disorders, both traits can be used in linkage analyses. Correlated measurements can improve results by decreasing genetic heterogeneity in affected subjects and by reducing measurement error.

The purpose of the present study is to use a genomic screen to identify chromosomal regions that potentially harbor genes modifying the risk for habitual smoking and the composite phenotype habitual smoking and alcohol dependence in families of alcohol dependent individuals. Data are from the Collaborative Study on the Genetics of Alcoholism (COGA), a large, multi-site genetic investigation of alcohol dependence with well-characterized families. This family study has previously demonstrated a complex interaction between alcohol dependence and habitual smoking, with evidence for common and specific factors in their familial transmission [Bierut et al., 1998]. In addition, a genomic screen for genetic linkage to alcohol dependence has been performed [Reich et al., 1998; Foroud et al., 2000] and several chromosomal areas (chromosomes 1, 2, 3, and 7) have been identified as possibly influencing the development of alcohol dependence. Analyses of sibling pairs affected with alcohol dependence supported linkage to chromosome 1 and 7 in both the initial genomic screen and the replication data set. Analyses also highlighted chromosomal areas (chromosome 2 and 3) in one data set but not the other.

We have extended this investigation to the study of habitual smoking and comorbid habitual smoking and alcohol dependence to address the following questions:

1. Do areas in the genome that are implicated in the development of alcohol dependence also contribute to the development of habitual smoking?
2. Are there areas in the genome that more strongly influence the development of habitual smoking compared to alcohol dependence?
3. Does the composite phenotype of habitual smoking and alcohol dependence modify linkage results?

## MATERIALS AND METHODS

Subjects were recruited as part of COGA, a multicenter family and genetic study of alcohol dependence [Reich et al., 1998]. The initial recruitment sites were: Indiana University, State University of New York at Brooklyn, University of California at San Diego, University of Connecticut, University of Iowa, and Washington University in St. Louis. The local Institutional Review Boards approved the study and written informed consent was obtained from all interviewed subjects.

### Subjects

Probands who met criteria for both DSM-IIIR alcohol dependence [American Psychiatric Association, 1987] and Feighner et al. [1972] definite alcoholism were identified in chemical dependency treatment settings, and their biologic relatives were recruited. Probands were older than 17 years, able to speak English, and had at least two first-degree relatives living in one of the COGA catchment areas. Initial subjects were excluded if they had HIV infection, life threatening illness, severe cognitive impairment, acute psychosis, or habitual IV

drug use (>30 times lifetime or any IV drug use in the last 6 months). All available first-degree family members of probands were invited to participate.

### Assessment

Subjects completed the semi-structured assessment for the genetics of alcoholism (SSAGA) [Bucholz et al., 1994], a highly reliable, semi-structured interview that assessed alcohol dependence, other substance dependence, habitual smoking, and related psychiatric disorders over a lifetime. Using this instrument, individuals were diagnosed with "alcoholism" if they met criteria for both DSM-IIIR alcohol dependence and Feighner definite alcoholism over their lifetime. Nicotine dependence was not evaluated at the initial assessment, so "habitual smoking," defined as ever smoking at least one pack (20 cigarettes) daily for 6 months or more, was used as a proxy for tobacco dependence. A subset of subjects (N = 1,501) was evaluated for both habitual smoking and DSM-IV nicotine dependence, [American Psychiatric Association, 1994], and 71% of those with habitual smoking also met criteria for DSM-IV nicotine dependence. A composite phenotype of habitual smoking and alcohol dependence was defined as having both lifetime diagnoses, and these illnesses need not have occurred concurrently.

To assess the rate of alcohol dependence in more distant relatives, subjects reported on the psychiatric history of all their relatives using the Family History Assessment Module (FHAM) [Rice et al., 1995]. Uninterviewed subjects were considered affected with alcoholism if three or more of their relatives reported symptoms consistent with a diagnosis of alcohol dependence. These family history reports of alcoholism were used to identify severely affected families for more extensive recruitment into the genetic sample.

### Genetic Sample

The genetic sample was collected in a sequential two-stage design. The first group was used as screening sample (Wave I) and the second for replication (Wave II).

Since COGA focused on the search for genes for alcoholism, families with at least three interviewed members diagnosed with alcohol dependence were selected for further study. More distant family members were recruited either as first-degree relatives of an alcohol dependent individual, or by extension over an unaffected individual into a branch of the family with at least two other relatives implicated with alcohol dependence by family history report. All available members of these extensions were recruited, and the process was repeated. Bilineal branches, where both parents were alcohol dependent, were not extended. Habitual smoking, other substance dependence, and psychiatric disorders were not used in the determination of further recruitment.

In the initial screening sample, 105 genetically informative multi-generational families for the study of alcohol dependence were selected for genotyping [Reich et al., 1998]. Parents and non-alcohol dependent siblings were included to maximize genetic information. A subset of 67 multi-generational families was also identified as having multiple individuals affected with habitual smoking. These families were divided into 78 nuclear families with 154 independent sibling pairs affected with habitual smoking and were identified as the initial screening sample in these analyses.

The replicate genetic sample was recruited using the same ascertainment criteria that focused on alcohol dependence. One hundred fifty seven multi-generational families were selected for the study of alcohol dependence, and a subset of 79 multi-generational families was also informative for the study of habitual smoking. These families were divided into 97 nuclear families with 173 habitual smoking independent sibling pairs.

Though habitual smoking and alcohol dependence frequently co-occur in the same individual, the overlap of both diagnoses was not complete. The sample was thus further sub-divided into families with at least one sibling pair affected with both habitual smoking and alcohol dependence, resulting in a screening sample of 88 independent sibling pairs and a replicate sample of 115 sibling pairs.

Three hundred thirty six genetic markers were genotyped on both the initial screening and replicate samples. These markers were chosen for analysis based on a combination of high informativeness (average heterozygosity 0.72), approximate 10–15 centimorgan (cM) spacing, and ease of use. Genotyping was completed using radioactive or fluorescent detection systems [Reich et al., 1998], and data were checked for Mendelian inheritance of markers. Marker order and distance were estimated from these data [Boehnke, 1991].

### Genetic Analysis

Because habitual smoking is a complex disorder with no clear mode of inheritance, a model independent multi-point sibpair method was used (ASPEX) [Risch et al., 1999]. We recognized that there have been significant changes in the rate of alcohol dependence and habitual smoking across generations [Bierut et al., 1998], so we examined sibling pairs to minimize potentially confounding inter-generational factors. ASPEX was used in analyses because it allows large sibships (greater than size 8) to be included in the analysis. Like other affected sibpair methods, ASPEX tests whether affected sibling pairs share alleles at a specific marker more than the 50% expected rate when there is no linkage. Significantly increased allele sharing (>50%) supports linkage of a disease susceptibility locus to the marker being tested. We focused our results on independent sibling pairs (N-1). ASPEX evaluates multiple markers on the same chromosome simultaneously and uses marker information from all siblings (affected and unaffected) in a family. We used the sibphase program that infers allele sharing when there is ambiguity of allele sharing (identity by state versus identity by descent), and marker frequencies in the sample are then used to estimate sibpair sharing. Since estimating

marker frequency from the data set can lead to biases because of ethnic stratification, we re-ran analyses on all chromosomes with positive findings using families where genotypic data was available from both parents in order to minimize false positive results due to biased allele frequencies. Focusing on families where genetic data is available on two parents reduces a sample's size, but increases confidence in estimates of sharing by affected sibling pairs.

To study the relationship between habitual smoking and alcohol dependence, two phenotypes were examined: habitual smoking and the composite phenotype of habitual smoking and alcohol dependence. These results were also compared with analyses of alcohol dependence alone [Reich et al., 1998; Foroud et al., 2000].

## RESULTS

Table I details the demographic characteristics of the genetic sample and lists the rates of alcohol dependence and habitual smoking. Comorbid alcohol dependence and habitual smoking was more common in individuals than expected by chance alone.

The multipoint linkage analysis for habitual smoking in the initial (Wave I) screening data set gave several modest signals with a LOD scores greater than one on four chromosomes (chromosomes 5, 9, 11, and 21). See Table II. The strongest signal was on chromosome 9 near two consecutive markers (marker D9S1120 (92 cM)—allele sharing 58.9%; LOD = 1.51; marker D9A261 (116 cM)—allele sharing 58.4%; LOD = 1.28), and the LOD score peaked in the intermarker location (103 cM—allele sharing 63.9%; LOD = 2.02). See Figure 1. The combined phenotype of habitual smoking and alcohol dependence showed similar estimates of sibpair allele sharing at this site, but lower LOD scores, a result consistent with the smaller sample size (marker D9S1120 (92 cM)—allele sharing 57.9%; LOD = 0.66; D9A261 (116 cM)—allele sharing 56.5%; LOD = 0.46).

TABLE I. Demographic Characteristics and Lifetime Prevalences of Substance Dependence in the Genetic Sample

| | Wave I (N = 970) | Wave II (N = 1159) |
|---|---|---|
| Men | 454 | 550 |
| Women | 516 | 609 |
| Age | 40.6 years (range 17–91 years) (SD ± 15.05) | 40.7 years (range 18–91) (SD ±14.51) |
| Caucasian | 76.4% | 80.9% |
| African–American | 13.3% | 13.4% |
| Hispanic | 8.5% | 2.8% |
| Other | 1.9% | 2.8% |
| Alcohol dependence | | |
| Men | 67.6% | 69.5% |
| Women | 26.4% | 35.0% |
| Habitual smoking | | |
| Men | 54.4% | 52.6% |
| Women | 29.3% | 36.6% |
| Composite phenotype | | |
| Men | 41.0%* | 42.2%* |
| Women | 11.2%* | 17.4%* |

*Significantly more comorbidity than expected by chance alone $P < 0.001$.

TABLE II. Linkage Results for Habitual Smoking Sibling Pairs in the Wave I Screening Sample

| | Marker | Location (cM) | LOD score | Percent sharing (%) |
|---|---|---|---|---|
| Chromosome 5 | D5S815 | 119 | 1.12 | 57.8 |
| Chromosome 9 | D9S1120 | 92 | 1.51 | 58.9 |
| | D9A261 | 116 | 1.28 | 58.4 |
| | D9S904 | 168 | 1.05 | 58.5 |
| Chromosome 11 | D11S1354 | 87 | 1.64 | 59.9 |
| Chromosome 21 | D21S120 | 0 | 1.51 | 59.6 |

Interestingly, this chromosomal location did not show strong evidence of linkage to alcohol dependence (marker D9S1120 (92 cM)—allele sharing 51.0%; LOD = 0.03; D9A261 (116 cM)—allele sharing 55.2%; LOD = 0.76).

Since habitual smoking and alcohol dependence are frequently comorbid, we repeated the linkage analyses using the composite phenotype. Though the number of sibling pairs is greatly reduced compared to analyses with habitual smoking or alcohol dependence alone, and thus the power to detect linkage is reduced, multiple areas still showed modest signals with LOD scores greater than 1 (chromosomes 1, 2, 11, and 15). See Table III. The strongest signal was on chromosome 2 (marker D2S379 (87 cM)—allele sharing 69.1%; LOD = 3.30), and allele sharing was greater than for either phenotype alone (habitual smoking allele sharing 56.5%; LOD = 0.67; alcohol dependence allele sharing 59.8%; LOD=2.24). See Figure 2.

Since the strongest evidence of linkage for these substance dependence phenotypes was on chromosome 9 for habitual smoking and chromosome 2 for the combined habitual smoking and alcohol dependence phenotype, we repeated analyses on these chromosomes in the replicate sample. See Figures 3 and 4. There continued to be increased allele sharing at markers near

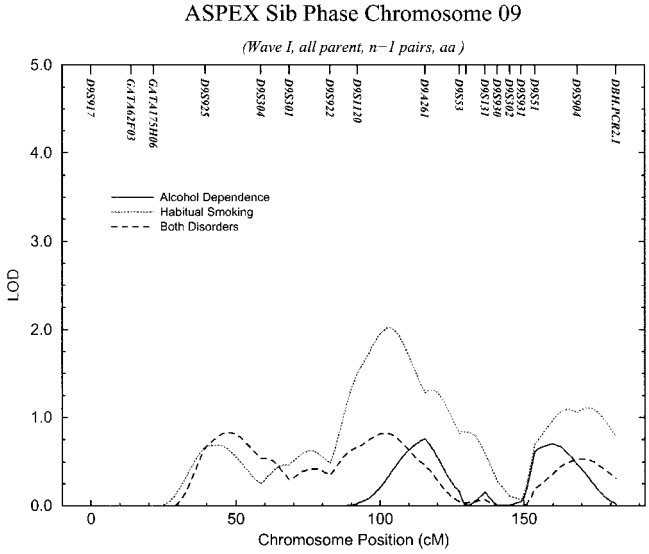


Fig. 1. Chromosome 9 linkage results on the Wave I screening sample.

TABLE III. Linkage Results for Comorbid Habitual Smoking and Alcohol Dependent Sibling Pairs in the Wave I Screening Sample

| | Marker | Location (cM) | LOD score | Percent sharing (%) |
|---|---|---|---|---|
| Chromosome 1 | D1S224 | 106 | 1.07 | 59.8 |
| | D1S1588 | 142 | 1.14 | 60.7 |
| | D1S1675 | 165 | 1.36 | 61.2 |
| Chromosome 2 | D2S379 | 87 | 3.30 | 69.1 |
| Chromosome 11 | D11S1354 | 87 | 1.05 | 60.7 |
| Chromosome 15 | D15S217 | 14 | 1.25 | 59.7 |

the multipoint LOD score peak, though not as strong as in the initial screening sample. Chromosome 9 had slightly increased allele sharing for habitual smoking at the peak identified in the screening sample (marker D9S1120 (92 cM)—allele sharing = 53.1%; LOD = 0.19) and a larger peak 35 cM distal (127 cM—allele sharing 54.5%; LOD = 0.53). Chromosome 2 had increased allele sharing for habitual smoking (marker D2S379 (87 cM)—allele sharing 57.6%; LOD = 0.92) at the same location as in the screening sample and sharing reached its maximum 24 cM away (marker D2S1790 (110 cM)—allele sharing = 59.1%; LOD = 1.79). In the comorbid habitual smoking and alcohol dependent sample, there continued to be increased allele sharing, but it was similar to the allele sharing for habitual smoking alone (marker D2S279 (87 cM)—allele sharing = 56.0%; LOD = 0.39; marker D2S1790 (110 cM)—allele sharing = 57.9%; LOD = 0.88). See Table IV. Only one other chromosomal region gave modest (greater than 1) LOD scores in the replication sample (marker D7S490 (147cm)—allele sharing = 58.0%; LOD = 1.13 for comorbid habitual smoking and alcohol dependence). This region is 60 cM from the previously identified area on chromosome 7 for alcohol dependence. None of the phenotypes (alcohol dependence, habitual smoking, or the comorbid habitual smoking and alcohol dependence)

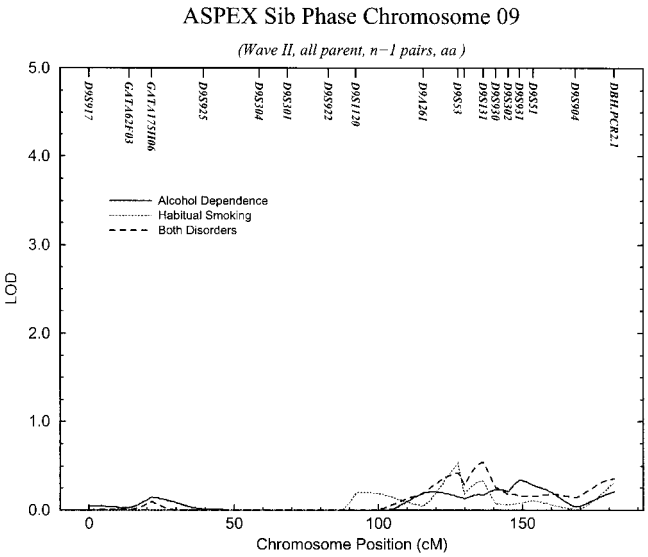


Fig. 3. Chromosome 9 linkage results on the Wave II replication sample.

in the initial sample provided evidence for linkage at this region (allele sharing 48.5–50.6%).

We also examined other genomic areas previously identified as having increased sharing for alcohol dependence [Reich et al., 1998; Foroud et al., 2000]. In the initial screening sample, chromosomes 1 and 7 had elevated allele sharing for alcohol dependence [Reich et al., 1998]. Habitual smoking showed had no evidence of increased allele sharing on either chromosome 1 (marker D1S1588 (142 cM)—allele sharing 50.2%; LOD = 0.0) or chromosome 7 (marker D7S1793 (81 cM)—allele sharing 50.4%; LOD = 0.0) at the regions previously identified with the alcohol dependence phenotype. The combined phenotype had increased allele sharing similar to that of alcohol dependence alone on chromosome 1 (marker D1S1588 (142 cM)—allele

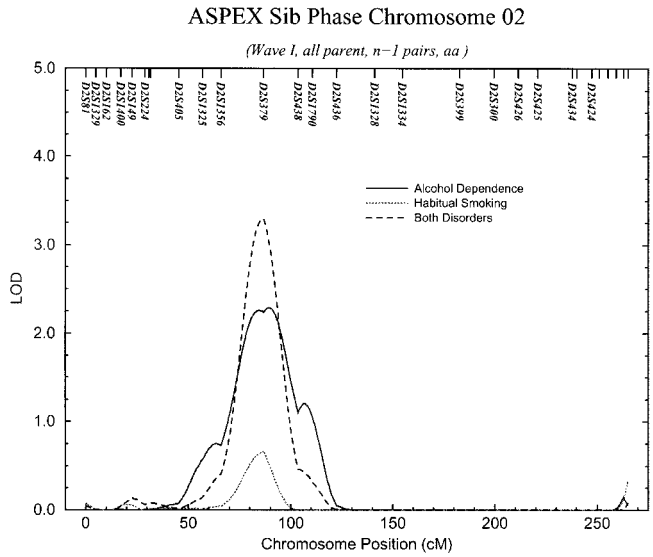


Fig. 2. Chromosome 2 linkage results on the Wave I screening sample.

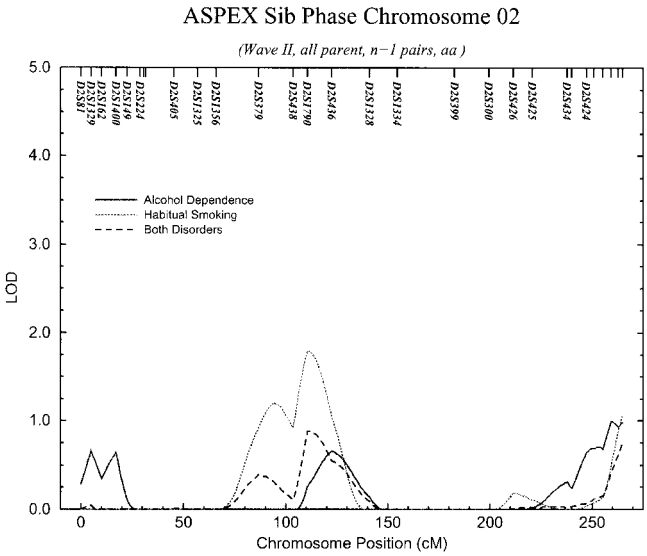


Fig. 4. Chromosome 2 linkage results on the Wave II replication sample.

TABLE IV. Allele Sharing for Sibling Pairs at Peaks on Chromosomes 2 and 9

| | D2S379 (%) | D9S1120 (%) |
|---|---|---|
| Habitual smoking sibling pairs | | |
| Wave I | 56.5 | 58.9 |
| Wave II | 57.5 | 53.1 |
| Alcohol dependence sibling pairs | | |
| Wave I | 59.8 | 51.0 |
| Wave II | 47.6 | 48.3 |
| Composite phenotype sibling pairs | | |
| Wave I | 69.1 | 57.9 |
| Wave II | 56.0 | 49.3 |

sharing 60.7%; LOD = 1.14) and there was no evidence of linkage on chromosome 7 (marker D7S1793 (82 cM)—allele sharing 50.4%; LOD = 0.02). In the replicate sample, positive results were found on chromosome 3 for alcohol dependence [Foroud et al., 2000]. Both habitual smoking and the combined phenotype had increased allele sharing similar to that of alcohol dependence (marker D3S1766 (76 cM)—alcohol dependence—allele sharing 57.0%; LOD = 1.34; habitual smoking—allele sharing 56.4%; LOD = 0.70; combined phenotype—allele sharing 57.0%; LOD = 0.57), however because of the smaller sample sizes the LOD scores are below 1.

Since we were concerned that ethnic stratification may be confounding these results, we repeated analyses using families in which both parents were genotyped in a combined analysis of the screening and replication sample. The combined sample was used since requiring two parents genotyped reduced the sample size by half. This analysis strengthened the finding on chromosome 2 for the composite phenotype (combined data set: comorbid habitual smoking and alcohol dependence D2S379 (87 cM) allele sharing 68.6%; LOD = 4.3; N = 111 independent sibling pairs), and allele sharing on chromosome 9 remained the same (combined data set: habitual smoking D9S1120 (92 cM) allele sharing 57.2%; LOD; 1.18; N = 174 independent sibling pairs) though the LOD score decreased because of the smaller sample size used.

## DISCUSSION

Since genetic factors are implicated in the development of smoking and there is evidence of shared genetic factors in the development of dependence, we performed a genomic screen for habitual smoking in families of alcoholics. Our results support previous studies that there are both common and specific factors in the development of smoking and alcohol dependence [Bierut et al., 1998; True et al., 1999].

We found chromosomal regions that have increased allele sharing among habitual smoking sibling pairs and are thus implicated in the development of habitual smoking in families of alcoholics in the initial screening sample. These same chromosomal regions showed little or no increase in allele sharing over the expected 50% rate in alcohol dependent sibling pairs. For instance, on chromosome 9 there was increased allele sharing for sibling pairs affected with habitual smoking, whereas there was no elevation in allele sharing for sibling pairs affected with alcohol dependence. The composite phenotype of habitual smoking and alcohol dependence had similar allele sharing as the analyses for habitual smoking alone. Thus, this region may have a susceptibility locus that is specific for the development of habitual smoking.

Similarly, there were genomic areas implicated in the development on alcohol dependence, and not habitual smoking. Chromosomes 1 and 7 were reported [Reich et al., 1998; Foroud et al., 2000] as demonstrating evidence of linkage for alcohol dependence. In contrast, in this same data set, there was no evidence for increased allele sharing among habitual smoking sibling pairs at these regions. The combined phenotype of habitual smoking and alcohol dependence had increased allele sharing among sibling pairs similar to results in the analysis of alcohol dependence alone or no increased allele sharing. These results provide evidence for specific genetic factors in the development of alcohol dependence that do not impact the development of habitual smoking.

In the replication sample, chromosome 3 provided modest evidence for linkage for alcohol dependence [Foroud et al., 2000], This region also demonstrated increased allele share for habitual smoking and the combined phenotype. Allele sharing was essentially identical for all three phenotypes examined, though the LOD scores differed because of the differences in sample size. This may be evidence of a common genetic factor that contributes to either smoking or alcohol dependence.

Since a complicating factor in the study of addiction is the high rate of comorbidity, we repeated analyses focusing on individuals affected with both illnesses. The composite phenotype of habitual smoking and alcohol dependence gave three results: allele sharing similar to habitual smoking; allele sharing similar to alcohol dependence; and allele sharing increased over either independent phenotype. For instance, allele sharing on chromosome 2, which was elevated for both habitual smoking and alcohol dependence, increased dramatically when the composite phenotype was examined. There may be a genetic factor on chromosome 2 that confers susceptibility to a combined habitual smoking and alcohol dependent phenotype, or the composite phenotype may represent a more genetically homogeneous characteristic.

Assessing the significance of multiple testing is difficult, and so a replication sample was built into the study design. Chromosome 2 showed the strongest evidence of linkage in the initial screening sample. In the replication sample, there continued to be increased allele sharing and an increased LOD score. However, the peak for the LOD score was shifted 35 cM. Also, in the initial screening sample allele sharing was greatest for the combined phenotype of habitual smoking and alcohol dependence, whereas in the replication sample, the allele sharing was similarly elevated for habitual

smoking (57.6%) and the comorbid phenotype (56.0%). Since the sample size for the habitual smoking sibling pairs was one and one-half times larger than the comorbid sibling pairs sample, the LOD score was greater for habitual smoking in the replication sample. Chromosome 9 showed increased allele sharing and LOD score in the screening sample and a continued modest elevation in the replication sample. Though these findings of increased allele sharing in both the screening and replicate sample are intriguing, LOD scores were modest and below the guidelines set by Lander and Kruglyak [1995] for confirming linkage, and we cannot say confidently that there is replication of these genetic findings.

We were concerned that ethnic stratification may be confounding these results since allele sharing is estimated from the study population when there is ambiguity regarding identity by descent. We repeated analyses using families in which both parents were genotyped in a combined analysis of the screening and replication sample. The combined sample was used since requiring two parents genotyped reduced the sample size by half. This analysis strengthened the finding on chromosome 2 for the composite phenotype, and continued to support the finding on chromosome 9. Thus, it seems unlikely that these findings are false positives due to misspecification of allele frequencies.

Because of the differences in genetic findings between the initial screening and replication sample, we searched these two data sets for clinical variables that may explain these discrepant results. With all phenotypic characteristics analyzed, the two samples showed no significant differences. There were no differences in age of onset of habitual smoking or alcohol dependence in the two groups. Comorbid disorders (such as other substance dependence, major depressive disorder, conduct disorder, or antisocial personality disorder) did not differentiate these two samples. Though the initial screening sample and the replication sample appear phenotypically similar, there remain discrepant genetic results.

Smoking data from the initial screening sample was previously analyzed in part during the Genetic Analysis Workshop 11. Bergen et al. [1999] examined a dichotomous variable "ever smoking" defined as a lifetime history smoking daily for a month or more and smoking more than 100 cigarettes in one's lifetime and analyzed the genetic markers using single point sibling pair analysis. Three chromosomal regions (chromosomes 6, 9, and 14) had increased allele sharing at two adjacent markers, and an additional 12 independent markers had significantly increased allele sharing at a $P < 0.01$ level. Several findings are consistent with our multipoint analysis of habitual smoking. The chromosome 9 and 21 findings overlap for the "ever smoking" and "habitual smoking" definitions. In addition, the region on chromosome 2 for the composite phenotype habitual smoking and alcohol dependence also gave a significant result for the "ever smoking" phenotype. These similarities in findings are reassuring, but differences are also expected since varying thresholds for smoking and discrete analytic techniques (single point versus multipoint) were used. Duggirala et al. [1999] used another approach to the genetic analysis of smoking in these data. Smoking behavior was categorized as a quantitative variable of number of cigarette packs smoked per day per year, and a multipoint variance components method was used for genetic analyses. A very high LOD score was found on chromosome 5 (D5S1354), however, as the authors discuss, there is a large gap between markers at this location and at the end of the chromosome, which reduced their enthusiasm for this putative genetic locus. These differing results highlight how an identical data set can produce varying results based on phenotypic definition and analytic method.

An independent genomic survey of nicotine dependence has been performed [Straub et al., 1999]. An area of interest on chromosome 2 was identified for nicotine dependence, however, it was about 50 cM from our peak on chromosome 2. Our findings were strongest on chromosome 2 for the composite phenotype of habitual smoking and alcohol dependence, whereas Straub et al. examined only nicotine dependence. Straub did not emphasize any findings on chromosome 5, 9, 11, or 21 (our findings of a multi-point LOD score greater than 1 for habitual smoking) since their results fell below a threshold for follow up. Similarly, areas highlighted by Straub (chromosomes 4, 10, 16, 17, and 18) did not have significantly increased allele sharing in our data. A major difference in our results may be due to our different definitions of affected. We used a phenotype of "habitual smoking," that is smoking one pack of cigarettes or more a day for 6 months or more in families with many alcohol dependent members, whereas Straub et al. used a definition of nicotine dependence defined by the Fagerström Tolerance Questionnaire [Fagerstrom and Schneider, 1989] in a general population sample. In a subset of our data set, we have information on the Fagerström Tolerance and Nicotine Dependence (FTND) questionnaire, a correlate to the full Fagerström Tolerance Questionnaire [Heatherton et al., 1991], and we found that approximately 75% of our individuals identified with habitual smoking are also nicotine dependence by the FTND. Though all these definitions of nicotine dependence are correlated, there is not complete overlap of these disorders. These differences in phenotype definition are likely a major cause of our conflicting results.

It is difficult to estimate the potential genetic contribution of these identified areas to the development of habitual smoking, alcohol dependence, or the combined phenotype. These studied phenotypes are the result of complex interactions between numerous genes and environmental factors. An estimation of the potential contribution of these genetic factors requires models with an approximation of the number of genes involved, interaction between genes, and environmental factors. With many unknown variables, it is difficult to approximate the potential genetic contribution of these areas.

In conclusion, we expect multiple genes of modest effect to be involved in the development of this complex phenotype and for replication to be difficult. We have taken a stepwise approach to the study of smoking in our analysis of genetic data. These data support a chromosomal region that may be specific in the development of

habitual smoking and not alcohol dependence (chromosome 9), and conversely areas that may be specific to the development of alcohol dependence and not habitual smoking (chromosome 1 and 7). There is also evidence for shared genetic risk to both habitual smoking and alcohol dependence (chromosome 3). The composite phenotype of habitual smoking and alcohol dependence may have a unique genetic risk (chromosome 2). These different genetic findings with correlated characteristics underscore the importance of careful phenotypic assessments for the genetic analysis of complex traits. Though this is a large sample, the number of sibling pairs for the analysis of a complex disorder is modest (320 independent sibling pairs for habitual smoking and 203 independent sibling pairs for the composite phenotype), and these results must be interpreted with caution. Multiple analyses were performed and there must be concerns about false positive results. Only the identification of genes involved in the development of these disorders will provide definitive confirmation of any linkage results.

## ACKNOWLEDGMENTS

We would like to acknowledge the assistance of Sarah Bertelsen and Joseph Mullaney in data analysis. Some results of this paper were obtained by using the program package S.A.G.E., which is supported by a U.S. Public Health Service Resource Grant (1 P41 RR03655) from the National Center for Research Resources. The Collaborative Study on the Genetics of Alcoholism (COGA) (H. Begleiter, SUNY HSCB Principal Investigator, T. Reich, Washington University, Co-Principal Investigator) includes nine different centers where data collection, analysis, and storage take place. The nine sites and Principal Investigators and Co-Investigators are: Howard University (R. Taylor); Indiana University (H. Edenberg, J. Nurnberger Jr, P.M. Conneally, T. Foroud); Rutgers University (J. Tischfield); Southwest Foundation (L. Almasy); State University of New York Health Sciences Center at Brooklyn (B. Porjesz, H. Begleiter); University of California at San Diego (M. Schuckit); University of Connecticut (V. Hesselbrock); University of Iowa (R. Crowe, S. Kuperman); Washington University in St. Louis (T. Reich, C.R. Cloninger, J. Rice, A. Goate). Lisa Neuhold serves as the NIAAA Staff Collaborator. This national collaborative study is supported by the NIH Grant U10AA08403 from the National Institute on Alcohol Abuse and Alcoholism (NIAAA). U.S. Public Health Services Grant AA00231 also supported this work. Presented in part at the World Congress of Psychiatric Genetics, Bonn, Germany, 1998.

## REFERENCES

American Psychiatric Association. 1987. Diagnostic and statistical manual of mental disorders. Third edition (revised). Washington, DC: American Psychiatric Press.

American Psychiatric Association. 1994. Diagnostic and statistical manual of mental disorders. Fourth edition. Washington, DC: American Psychiatric Press.

Bergen AW, Korczak JF, Weissbecker KA, Goldstein AM. 1999. A genome-wide search for loci contributing to smoking and alcoholism. Genet Epidemiol 17S1:S55–S60.

Bierut LJ, Dinwiddie SH, Begleiter H, Crowe RR, Hesselbrock V, Nurnberger JI Jr, Porjesz B, Schuckit MA, Reich T. 1998. Familial transmission of substance dependence: Alcohol, marijuana, cocaine, and habitual smoking: A report from the Collaborative Study on the Genetics of Alcoholism. Arch Gen Psychiatry 55:982–988.

Bierut LJ, Rice JP, Edenberg HJ, Goate A, Foroud T, Cloninger CR, Begleiter H, Conneally PM, Crowe RR, Hesselbrock V, Li T-K, Nurnberger JI Jr, Porjesz B, Schuckit MA, Reich T. 2000. Family-based study of the association of the dopamine D2 receptor gene (DRD2) with habitual smoking. Am J Med Genet 90:299–302.

Boehnke M. 1991. Allele frequency estimation from data on relatives. Am J Hum Genet 48:22–25.

Boustead C, Taber H, Idle JR, Cholerton S. 1997. CYP2D6 genotype and smoking behaviour in cigarette smokers. Pharmacogenetics 7:411–414.

Bucholz KK, Cadoret R, Cloninger CR, Dinwiddie SH, Hesselbrock VM, Nurnberger JI Jr, Reich T, Schmidt I, Schuckit MA. 1994. A new, semi-structured psychiatric interview for use in genetic linkage studies: A report on the reliability of the SSAGA. J Studies Alcohol 55:149–158.

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

CDC. 2001. Cigarette smoking among adults—United States, 1999. Morbidity Mortality Weekly Report 50:869–873.

CDC. 2002. Trends in cigarette smoking among high school students—United States, 1991–2001. Morbidity Mortality Weekly Report 51: 409–412.

Cholerton S, Boustead C, Taber H, Arpanahi A, Idle JR. 1996. CYP2D6 genotypes in cigarette smokers and non-tobacco users. Pharmacogenetics 6:261–263.

Comings DE, Ferry L, Bradshaw-Robinson S, Burchette R, Chiu C, Muhleman D. 1996. The dopamine D2 receptor (DRD2) gene: A genetic risk factor in smoking. Pharmacogenetics 6:73–79.

Duggirala R, Almasy L, Blangero J. 1999. Smoking behavior is under the influence of a major quantitative trait locus on human chromosome 5q. Genet Epidemiol 17S1:S139–S144.

Fagerstrom KO, Schneider NG. 1989. Measuring nicotine dependence: A review of the Fagerstrom Tolerance Questionnaire. J Behav Med 12:159–182.

Feighner JP, Robins E, Guze SB, Woodruff RAJ, Winokur G, Munoz R. 1972. Diagnostic criteria for use in psychiatric research. Arch Gen Psychiatry 26:57–63.

Foroud T, Edenberg HJ, Goate A, Rice J, Flury L, Koller DL, Bierut LJ, Conneally PM, Nurnberger JI, Bucholz KK, Li TK, Hesselbrock V, Crowe R, Schuckit M, Porjesz B, Begleiter H, Reich T. 2000. Alcoholism susceptibility loci: Confirmation studies in a replicate sample and further mapping. Alcohol Clin Exp Res 24:933–945.

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

Heath AC, Bucholz KK, Madden PA, Dinwiddie SH, Slutske WS, Bierut LJ, Statham DJ, Dunne MP, Whitfield JB, Martin NG. 1997. Genetic and environmental contributions to alcohol dependence risk in a national twin sample: Consistency of findings in women and men. Psychol Med 27:1381–1396.

Heatherton TF, Kozlowski LT, Frecker RC, Fagerstrom KO. 1991. The Fagerstrom Test for Nicotine Dependence: A revision of the Fagerstrom Tolerance Questionnaire. Br J Addict 86:1119–1127.

Kendler KS, Neale MC, Sullivan P, Corey LA, Gardner CO, Prescott CA. 1999. A population-based twin study in women of smoking initiation and nicotine dependence. Psychol Med 29:299–308.

Lander E, Kruglyak L. 1995. Genetic dissection of complex traits: Guidelines for interpreting and reporting linkage results. Nat Genet 11:241–247.

Lerman C, Caporaso NE, Audrain J, Main D, Bowman ED, Lockshin B, Boyd NR, Shields PG. 1999. Evidence suggesting the role of specific genetic factors in cigarette smoking. Health Psychol 18:14–20.

Lerman C, Caporaso NE, Bush A, Zheng Y-L, Audrain J, Main D, Shields PG. 2001. Tryptophan hydroxylase gene variant and smoking behavior. Am J Med Genet 105:518–520.

Madden PA, Heath AC, Starmer GA, Whitfield JB, Martin NG. 1995. Alcohol sensitivity and smoking history in men and women. Alcohol Clin Exp Res 19:1111–1120.

Pianezza ML, Sellers EM, Tyndale RF. 1998. Nicotine metabolism defect reduces smoking. Nature 393:7501.

Reich T, Edenberg HJ, Goate A, Williams JT, Rice JP, Van Eerdewegh P, Foroud T, Hesselbrock V, Schuckit MA, Bucholz K, Porjesz B, Li T-K, Conneally PM, Nurnberger JI Jr, Tischfield JA, Crowe RR, Cloninger CR, Wu W, Shears S, Carr K, Crose C, Willig C, Begleiter H. 1998. Genome-wide search for genes affecting the risk for alcohol dependence. Am J Med Genet (Neuropsychiatric Genetics) 81:207–215.

Rice JP, Reich T, Bucholz KK, Neuman RJ, Fishman R, Rochberg N, Hesselbrock VM, Nurnberger JI Jr, Schuckit MA, Begleiter H. 1995. Comparison of direct interview and family history diagnoses of alcohol dependence. Alcohol Clin Exp Res 19:1018–1023.

Risch N, Spiker D, Lotspeich L, Nouri N, Hinds D, Hallmayer J, Kalaydjieva L, McCague P, Dimiceli S, Pitts T, Nguyen L, Yang J, Harper C, Thorpe D, Vermeer S, Young H, Hebert J, Lin A, Ferguson J, Chiotti C, Wiese-Slater S, Rogers T, Salmon B, Nicholas P, Myers RM, et al. 1999. A genomic screen of autism: Evidence for a multilocus etiology. Am J Hum Genet 65:493–507.

Shields PG, Lerman C, Audrain J, Bowman ED, Main D, Boyd NR, Caporaso NE. 1998. Dopamine D4 receptors and the risk of cigarette smoking in African-Americans and Caucasians. Cancer Epidemiol Biomarkers Prev 7:453–458.

Spitz MR, Shi H, Yang F, Hudmon KS, Jiang H, Chamberlain RM, Amos CI, Wan Y, Cinciripini P, Hong WK, Wu X. 1998. Case-control study of the D2 dopamine receptor gene and smoking status in lung cancer patients. J Natl Cancer Inst 90:358–363.

Straub RE, Sullivan PF, Ma Y, Myakishev MV, Harris-Kerr C, Wormley B, Kadambi B, Sadek H, Silverman MA, Webb BT, Neale MC, Bulik CM, Joyce PR, Kendler KS. 1999. Susceptibility genes for nicotine dependence: A genome scan and follow-up in an independent sample suggest that regions on chromosomes 2, 4, 10, 16, 17, and 18 merit further study. Mol Psychiatry 4:129–144.

Swan GE, Carmelli D, Cardon LR. 1996. The consumption of tobacco, alcohol, and coffee in Caucasian male twins: A multivariate genetic analysis. J Substance Abuse 8:19–31.

Swan GE, Carmelli D, Cardon LR. 1997. Heavy consumption of cigarettes, alcohol, and coffee in male twins. J Studies Alcohol 58:182–190.

True WR, Heath AC, Scherrer JF, Waterman B, Goldberg J, Lin N, Eisen SA, Lyons MJ, Tsuang MT. 1997. Genetic and environmental contributions to smoking. Addiction 92:1277–1287.

True WR, Xian H, Scherrer JF, Madden PA, Bucholz KK, Heath AC, Eisen SA, Lyons MJ, Goldberg J, Tsuang M. 1999. Common genetic vulnerability for nicotine and alcohol dependence in men. Arch Gen Psychiatry 56:655–661.

Vandenbergh DJ, Bennett CJ, Grant MD, Strasser AA, O'Connor R, Stauffer RL, Vogler GP, Kozlowski LT. 2002. Smoking status and the human dopamine transporter variable number of tandem repeats (VNTR) polymorphism: Failure to replicate and finding that never-smokers may be different. Nicotine Tob Res 4(3):333–340.