Evidence for Genetic Linkage to Alcohol Dependence on Chromosomes 4 and 11 From an Autosome-Wide Scan in an American Indian Population

Jeffrey C. Long,1* William C. Knowler,2 Robert L. Hanson,2 Robert W. Robin,1 Margrit Urbanek,1 Elisa Moore,1 Peter H. Bennett,2 and David Goldman1

1Laboratory of Neurogenetics, National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, Bethesda, Maryland
2Phoenix Epidemiology and Clinical Research Branch, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Phoenix, Arizona

To identify specific genes affecting vulnerability or resistance, we performed a whole-autosomal genome scan for genetic linkage to alcohol dependence in a Southwestern American Indian tribe. Genotypes at 517 autosomal microsatellite loci and clinical evaluations were available for 152 subjects belonging to extended pedigrees and forming 172 sib-pairs. Highly suggestive evidence for linkage emerged for two genomic regions using two- and multipoint sib-pair regression methods; both regions harbored neurogenetic candidate genes. The best evidence is seen with D11S1984 (nominal P = 0.00007, lod ≅ 3.1) on chromosome 11p, in close proximity to the DRD4 dopamine receptor and tyrosine hydroxylase (TH) genes. Good evidence is seen with D4S3242 (nominal P = 0.0002, lod ≅ 2.8) on chromosome 4p, near the β1 GABA receptor gene. Interestingly, three loci in the alcohol dehydrogenase gene cluster on chromosome 4q showed evidence for linkage with two-point analyses, but not multipoint analysis. Am. J. Med. Genet. (Neuropsychiatr. Genet.) 81:216–221, 1998. © 1998 Wiley-Liss, Inc.

INTRODUCTION

Alcohol dependence is the most severe diagnostic category of alcoholism. It indicates that an individual continues drinking despite significant problems related to tolerance, withdrawal, and compulsive drinking behavior [American Psychiatric Association, 1987]. Family [Aston and Hill, 1990; Crum and Harris, 1996], twin [Kendler et al., 1984; Pickens et al., 1991], and adoption [Cloninger et al., 1981; Sigvardsson et al., 1996] studies reveal that alcohol dependence is causally complex, but in part hereditary. According to epidemiologic studies, the prevalence of alcoholism is higher in men than in women and differs among ethnic groups. Regardless of sex or ethnicity, prevalence is negatively associated with education, income, and occupational status, and positively associated with marital instability [Bucholz, 1992; Helzer and Pryzbeck, 1988; Helzer et al., 1991]. Extremely high rates of alcoholism have been observed in some American Indian populations, i.e., approximately 80% in men and 50% in women. Nevertheless, no evidence suggests that alcoholism in American Indians departs from the basic sex/socioeconomic trends observed in other ethnic groups [Brown et al., 1993; Kinzie et al., 1992].

Well-defined populations such as American Indian tribes are advantageous for identifying the causes underlying complex diseases [Lander and Schork, 1994]. Such populations are often more genetically and environmentally homogeneous than the general population; they are geographically restricted, and large families are common. The members of the American Indian tribe studied here have shared a common heritage of biology, culture, language, and location since pre-Columbian times. The tribal name and exact location are not given here because they are unnecessary for the analyses presented.

In a large sample (N = 582) from this American Indian tribe [Long et al., 1996], the odds for alcoholism in a subject are significantly increased by having an
alcoholic first-degree relative. While the male-to-male pattern of risk is not statistically significant, the odds for alcoholism in males are increased about 1.6-fold by having an alcoholic male first-degree relative. In females, there is a statistically significant 2.2-fold increase in odds for alcoholism afforded by having an affected first-degree female relative. Similar results have been observed in a large U.S. metropolitan sample. Interestingly, the odds for alcoholism in females in this sample are the significantly increased by having alcoholic female second- and even third-degree relatives. The risk pattern in females translates to a 40% heritability of alcoholism, which approximates that seen in female twin studies [Kendler et al., 1994; Pickens et al., 1991]. It is also consistent with the hypothesis that for multifactorial disorders, the subgroup with the lower frequency requires a higher genetic loading for expression [Reich et al., 1972, 1979]. More importantly, these findings suggest that genetic factors determining alcoholism in this population are not overwhelmed by a high environmental loading. To find specific genes predisposing to alcohol dependence, we performed a genome-wide scan using typings at 517 auto-
somal microsatellite loci spaced at an average distance of 6.9 cM. Genotypes were available for 152 subjects clinically evaluated for DSM-III-R alcohol dependence [American Psychiatric Association, 1987].

MATERIALS AND METHODS

Sampling Strategy

Elder tribal members who are regarded as matri-
archs or patriarchs and who possess a good knowledge of family structures provided information on large mul-
tigenerational genealogies. Three of these were se-
lected for analysis on the basis of availability of adult
members, geographic accessibility, and willingness of
at least one family member to advocate for the study.
Participants were genealogy members, age ≥21 years,
and eligible for tribal enrollment (≥¼ tribal heritage).
Due to the high prevalence of alcoholism there was no
need to ascertain subjects through affected probands,
and recruitment was blind to the clinical histories of
subjects or their relatives.

Focus groups comprised of tribal staff and commu-
nity members reviewed testing instruments and ques-
tionnaires for potential cultural biases and general
suitability to the population. They also made recom-
mendations on our final selection of testing instru-
ments and on the conduct of interviews. Informed con-
sent was obtained under a human subjects research
protocol approved by the Tribal Council and the Insti-
tutional Review Board of NIAAA. Our final sample
consisted of 582 subjects, and the participation rate was
93% of the eligible persons contacted.

Testing Instruments, Interviews, and
Psychiatric Diagnoses

Research diagnoses for alcohol dependence were
based on: 1) a semistructured psychiatric interview us-
ing the Schedule for Affective Disorders and Schizo-
phrenia-Lifetime Version (SADS-L), 2) medical, educa-
tional, court, and other records, and 3) corroborative
information from family members. The SADS-L was
administered to all subjects by a psychologist experi-
enced in psychiatric assessment with this tribe and
other American Indian populations.

Diagnostic and Statistical Manual of Mental Disor-
ders, Third Edition-Revised (DSM-III-R) [American
Psychiatric Association, 1987] diagnoses were made
from the SADS-L by following operationally defined
criteria and using the instructions of Spitzer et al.
[1989]. Diagnoses were made from the SADS-L inter-
view data by two blind raters: a clinical social worker
and a clinical psychologist. Diagnostic differences were
resolved in a consensus conference which included a
senior psychiatrist experienced in diagnosis of Ameri-
can Indian people. All analyses reported hereafter in
this presentation are based on a subsample of 152 sub-
jects who also participated in a long-term epidemiologic
study of chronic disease, and for whom genotypes at
517 microsatellite loci were available. The 152 clini-
cally evaluated subjects belong to 32 interrelated
nuclear families, and they form 172 sib-pairs. Among
the sib-pairs, 18 were concordantly unaffected, 57 were
discordant, and 97 were concordantly affected.

Genotypings

The typings were determined for 1,311 individuals as
a part of a study on other diseases for which our re-
search subjects substantially overlapped. Each locus
was amplified by PCR, using either fluorescent- or ra-
dioactive-labeled primers, and most typings were per-
fomed following standard protocols by the Marshfield
Medical Research Foundation (Marshfield, WI) [Dubovsky et al., 1995; Schwengel et al., 1994]. The loci
analyzed cover all of the autosomes with an average
separation of 6.9 cM and largest gap of 25.6 cM. The
CRI-MAP program [Lander and Green, 1987] was used
to create maps with distances estimated from these
data. In most cases, the order of the marker loci was
taken as that on maps provided by the Marshfield
Medical Research Foundation. In a few instances (chro-
mosomes 1, 4, 7, and 19), a slightly different marker
order was used, as there was strong evidence to sup-
port such an order. Allele frequencies were estimated
by gene counting on all 1311 individuals.

Linkage Analyses

Two-point linkage analysis was conducted using the
nonparametric sib-pair regression method developed
by Haseman and Elston [1972], as programmed in the
SIBPAL module of the S.A.G.E. package [S.A.G.E.,
1994]. While this sib-pair linkage method was origi-
nally explained for a continuous trait, it is also appli-
cable to binary traits [Elston, 1997]. Among the chief
advantages of the Haseman-Elston method are that it
is conceptually based on allelic identity by descent; it
is powerful with common disorders because it combines
information from all affected/unaffected classes of sib-
pairs; and it provides a significance test corrected for
multiplex sibships. The statistical validity of P values
produced by the Haseman-Elston test has been demon-
strated for small samples [Wilson and Elston, 1993].
In addition, we performed limited computer simula-
tions in order to confirm the accuracy of $P$ values produced by the SIBPAL program for this study. Based on a total of 50,000 trials using allelic configurations at five of the tested loci (10,000/locus) and the observed pedigrees and trait distributions, the overall agreement was good. Nevertheless, a slight anti-conservative trend was noted (SIBPAL—empirical): (0.0500→0.05290), (0.0100→0.01402), (0.0010→0.00244), and (0.0001→0.00034). Since a computer simulation to definitively estimate $P$ values, for all loci and allelic configurations, and for two-point and multipoint analysis, is beyond the scope of this study, we chose to report all $P$ values as “nominal $P$ values.”

Genotypes from 29 relatives who were not psychiatrically interviewed were used for family structure testing and to improve estimates of allelic identity by descent. Our sample contained 172 clinically evaluated sib-pairs, but the families included several large multiplex sibships; therefore, significance tests were conducted with about 70 effective degrees of freedom as determined by the standard weighting formula employed in SIBPAL. While multiplex sibships reduce the number of degrees of freedom available for hypothesis tests, they are generally beneficial for analysis because they provide information on the genotypes of missing parents, and thereby improve estimates of the proportion of shared genes that are identical by descent. SIBPAL uses a likelihood method for nuclear families that fully uses this information. With respect to these analyses, it is important to note that the 517 locus typings allowed for extensive parentage testing, so that our family relationships were essentially confirmed. As a second check, we note that the proportions of identical-by-descent alleles shared by siblings at these loci do not differ systematically from the expected, i.e., 50%. SIBPAL also provides an estimate of the proportion of alleles shared by descent in all classes of affected/unaffected sib-pairs. These proportions are useful for demonstrating the effects underlying the regression results.

Multipoint linkage analysis was performed by the sib-pair interval mapping procedure of Fulker et al. [1995], using a SAS program (written by R.L.H.). This is an extension of the Haseman-Elston procedure that simultaneously takes into account all marker information available on a chromosome. It is particularly useful for dense maps with marker loci that vary with respect to information content.

RESULTS

As demonstrated in Table I, alcohol dependence is highly prevalent, it is more common in men than women, and the mean age of participating subjects is late relative to the onset of heavy drinking.

The 517 Haseman-Elston sib-pair regression linkage analyses revealed the 30 loci listed in Table II with nominal $P$ values below 0.05. The results are presented by regression $P$ value because the Haseman-Elston method is the most powerful nonparametric linkage technique for nuclear families [Davis and Weeks, 1997]. Table II also provides the probabilities of allelic identity by descent (ibd), as computed by SIBPAL, for the three sib-pair classes (unaffected, discordant, and affected). As expected for genetic linkage, estimated ibd in discordant sib-pairs is <0.50 for all loci with a significant regression. Moreover, estimated ibd in concordant affected pairs is >0.50 for all but one locus with a significant regression.

Linkage to alcohol dependence is strongly suggested for D11S1984, which is located at chromosome 11p15.5. For D11S1984, the nominal regression $P$ value is 0.00007. This locus also shows a strong reduction in ibd, sharing in discordant sib-pairs (estimated ibd = 0.3950, nominal $P < 0.01$) and a strong excess of ibd, sharing in concordant affected sibpairs (estimated ibd = 0.5889, nominal $P < 0.01$). From Lander and Kruglyak [1995], a nominal $P = 0.00007$ translates to a nominal lod score of 3.1 (although it does not imply that a usual pedigree linkage analysis was performed). Observations this extreme would be observed by chance in only about one in six full-genome scans on a perfect genetic map that included the X chromosome. However, this is a conservative reference point because only autosomal genes are analyzed here, and because such extreme nominal $P$ values would be less frequent with a more sparse genetic map. For example, we would expect an observation at this significance level in only 1 in 33 full-genome scans using a 10-cM map.

Ten loci with nominal $P < 0.05$ are located on chromosome 4. Seven of these loci are on the short arm; the lowest nominal regression $P$ value (0.0002), equivalent to a lod score of about 2.8, is observed at D4S3242. This locus also shows a strong reduction in ibd sharing in discordant sib-pairs (estimated ibd = 0.3997, nominal $P < 0.01$) and a significant excess of ibd sharing in both discordant unaffected sib-pairs (estimated ibd = 0.6011, nominal $P < 0.05$) and concordant affected sibpairs (estimated ibd = 0.5564, nominal $P < 0.05$). We would expect one observation this extreme by chance in about 2–3 full-genome scans with a perfectly informative map, and less frequently with a more sparse map. In addition, we observed three closely linked loci with nominal $P$ values in the 0.01 $< P < 0.05$ range on chromosome 4q, D4S1647, D4S2356, and D4S3240. These loci map into the type I alcohol dehydrogenase (ADH) gene cluster. ADH catalyzes the first step in the principal ethanol metabolic pathway.

The two-point linkage analyses described above were followed with the multipoint extension [Fulker et al., 1995] of the Haseman-Elston linkage test. As shown in Figure 1 for chromosomes 4 and 11, the multipoint results closely parallel the two-point linkage findings. On chromosome 4, a position for an alcoholism-linked gene is sharply localized to the short arm, at p11–p13, and nearby D4S3242. The multipoint evidence for the other

<table>
<thead>
<tr>
<th>TABLE I. Sample Structure and Alcohol Dependence Phenotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Male</td>
</tr>
</tbody>
</table>

aSubject age (years) at clinical evaluation ± standard error. bAge (years) at which heavy drinking began ± standard error.
linkage on chromosome 4, near \textit{ADH}, is enigmatic because the evidence for linkage shown at the individual loci diminishes when they are analyzed in concert. Multipoint analysis does little to enhance the significant finding on chromosome 11. This is unsurprising because no flanking loci are distal to \textit{D11S1984}.

**DISCUSSION**

Table II demonstrates that all three sib-pair categories (unaffected, discordant, and concordant) contributed to the significant linkage findings. In fact, discordant sib-pairs are likely the most informative sib-pair category, which is expected given the high base rate of alcohol dependence in this population. This underscores the value of using the regression method of Haseman and Elston [1972], which combines information from all three categories. Methods which rely solely on affected sib-pairs are less powerful [Davis and Weeks, 1997] and potentially compromise linkage analysis of common heritable disorders.

Several interesting candidate genes related to neurotransmission are located in the vicinity of \textit{D11S1984} at the distal short arm of chromosome 11. Most prominent is the \textit{DRD4} dopamine receptor gene, which has recently been associated with novelty-seeking in Israelis and Euroamericans [Epstein et al., 1996; Benjamin et al., 1996] (but see Malhotra et al. [1996]). There are functional repeat polymorphisms at \textit{DRD4} that can be directly assessed. However, allele frequencies are different between American Indian and European populations [Chang et al., 1996], and this can influence the power to detect linkage, linkage disequilibrium, and phenotypic associations [Thompson et al., 1988]. The tyrosine hydroxylase gene is also on chromosome 11 [O'Malley and Rotwein, 1988], near \textit{D11S1984}. Tyrosine hydroxylase is the rate-limiting enzyme in dopamine biosynthesis. Therefore, variability in its coding or regulatory sequences may be related to individual differences in alcoholism vulnerability. Finally, it should be mentioned that the tryptophan hydroxylase gene, which encodes the rate-limiting enzyme in serotonin biosynthesis, is also located on chromosome 11p [Nielsen et al., 1992]. However, this gene is approximately 30 cM from \textit{D11S1984} towards the centromere.

The strongest chromosome 4 candidate locus is the \textit{b1 GABA} receptor gene (\textit{GABARB1}), which is close to \textit{D4S2424} [Dean et al., 1991]. \textit{GABA} (\gamma-aminobutyric acid) is an inhibitory neurotransmitter, and \textit{GABA} receptors are allosterically modulated by benzodiazepines, barbiturates, and ethanol. While no functional alleles are known at \textit{GABARB1}, it will be worthwhile to screen for them. The positive \textit{ADH} findings cannot be dismissed. There are well-known alleles at genes in the \textit{ADH} gene family that enhance catalytic activity.
Chromosome 11

Chromosome 4

Fig. 1. Log$_{10}$ P values for two-point and multipoint linkage analyses for alcohol dependence vs. microsatellite loci. Squares represent two-point results for alcohol dependence and specific microsatellite marker loci. Lines represent the multipoint extension of Fulker et al. [1995] of the method of Haseman-Elston [1972]. Results are shown only for chromosomes 4 and 11.

Acknowledgments

We thank the subjects who participated in this research study, and community members who provided invaluable advice and assistance. Thanks also go to L. Akhtar and L. Phillips for laboratory assistance, to S. Kobes for computer and data management, and to B. Albaugh, D. Breault, B. Chester, and J. Jaranson for assistance with psychiatric diagnosis. A. Bergen read and commented on earlier drafts of this paper. The genotypings performed by the Marshfield Clinic were funded by Glaxo-Wellcome Plc. A collaborative research and development agreement between Glaxo-Wellcome Plc and the National Institute of Diabetes and Digestive and Kidney Diseases (P.H.B., principal investigator) also provided partial support. Some of the 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.

References


