A Joint Genomewide Linkage Analysis of Symptoms of Alcohol Dependence and Conduct Disorder

Kenneth S. Kendler, Po-Hsiu Kuo, B. Todd Webb, Gursharan Kalsi, Michael C. Neale, Patrick F. Sullivan, Dermot Walsh, Diana G. Patterson, Brien Riley, and Carol A. Prescott

Background: A large linkage peak for alcohol dependence (AD) was detected on chromosome 4q in the Irish Affected Sib Pair Study of Alcohol Dependence (IASPSAD). Are the susceptibility genes underlying this peak specific for AD or do they increase risk for externalizing disorders more generally? Can we, in the IASPSAD, replicate prior evidence for linkage to conduct disorder (CD)?

Methods: The 733 all possible sibling pairs in IASPSAD were typed for 1,020 short-tandem-repeat genetic markers. Univariate and bivariate linkage analyses were conducted by the program sequential oligogenic linkage analysis routines (SOLAR), for both the raw and the transformed number of symptoms of AD (ADsx) and number of symptoms of CD (CDsx). In the bivariate analyses, specificity was assessed by the ratio of the variance accounted for in ADsx and CDsx by the quantitative trait locus.

Results: In the univariate linkage analyses, no evidence for linkage to CDsx was found under the 4q peak for ADsx and the largest peaks for CDsx were seen on chromosomes 1q (LOD = 3.16) and 14p (LOD = 2.36). In the bivariate linkage analysis, the 4q peak had high specificity for AD (AD/CD ratio of 39.9). Several smaller peaks, on chromosomes 1, 7, and 10, had moderate specificity for CD but also impacted on risk for AD, with AD/CD ratios of 0.18 to 0.32.

Conclusions: Genes under the 4q linkage peak for AD in the IASPSAD impact specifically on risk for AD rather than more broadly on risk for externalizing syndromes. Suggestive linkages were found in several locations for CD, 2 of which broadly replicate prior findings. The bivariate analyses identified genomic locations containing susceptibility loci that impacted on risk for both CDsx and ADsx.

Key Words: Alcohol Dependence, Conduct Disorder, Genetics, Linkage.

In a genomewide scan in the Irish Affected Sib Pair Study of Alcohol Dependence (IASPSAD), we recently detected strong evidence for linkage of the symptoms of alcohol dependence (AD) to the short arm of chromosome 4 (4q22–4q32; peak multipoint LOD = 4.59) (Prescott et al., 2006). Weaker evidence for linkage to AD symptoms was seen on chromosomes 2q, 9q, and 18p. A large population-based twin study previously found 2 major sources of genetic risk for AD: (i) genes shared in common with other externalizing disorders and (ii) genes specific for AD (Kendler et al., 2003). These results lead to the first question we seek to address in this article. For the positive linkage results in the IASPSAD—particularly on chromosome 4q—do the gene or genes responsible for these linkage peaks alter liability to a general externalizing spectrum of disorders or are they specific in their risk for AD? To address this question, we used symptoms of conduct disorder (CD) as an index of externalizing disorders as prior twin studies have demonstrated a substantial overlap in the genetic risk factors for AD and CD (Haber et al., 2005; Hicks et al., 2004; Slutske et al., 1998).

We are aware of 2 prior genomewide linkage scans for CD symptoms (Dick et al., 2004; Stallings et al., 2005) with divergent results. The second question we address in this article is whether in the relatively large and culturally and ethnically homogeneous IASPSAD sample, we can detect evidence for linkage to CD symptoms and, if so, whether any of these regions replicate prior mapping results for this phenotype.

MATERIALS AND METHODS

Sample

Data collection, which occurred from 1998 to 2002, was conducted as a collaborative effort between Virginia Commonwealth University, the Health Research Board in Dublin, and Shaftsbury Square Hospital in Belfast (Prescott et al., 2005). All interviewed participants provided informed consent before assessment and sample collection. The study protocol and consent procedures were approved by the VCU Institutional Review Board, Western IRB, the Health Research Board of the Irish Republic, and the human subjects committees of the treatment facilities from which participants were recruited (where such committees existed).
Proband ascertainment was by convenience sampling centered on community alcoholism treatment facilities and public and private hospitals in the Republic of Ireland and Northern Ireland. Probands were eligible if they (i) met DSM-IV criteria for AD (American Psychiatric Association, 1994); (ii) all 4 grandparents had been born in Ireland, Northern Ireland, Scotland, Wales, or England; and (iii) had at least 1 full sibling who met criteria for AD. Individuals with other substance dependence and psychiatric disorders were not excluded but we did assess the chronological relationship between the onsets of these disorders and AD. We attempted to enroll all living biological parents for whom the probands provided permission to contact.

Probands, siblings, and parents were interviewed by clinically trained research interviewers (often with extensive clinical experience with alcoholism) usually in participants’ homes or a treatment facility. A small proportion of siblings who lived outside Ireland were interviewed by telephone. Symptoms of AD (ADsx, range 3–7) were assessed with the SSAGA interview (version 11) (Bucholz et al., 1994). We assessed 14 of the 15 CD criteria in DSM-IV (American Psychiatric Association, 1994) (all except “forced someone into sexual activity”) by personal interview using an adaptation of questions from the Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID) (Spitzer et al., 1987). We examined in this article the sum of the endorsed CD criteria that we abbreviate CDsx.

For each symptom, we inquired about whether the criterion was only met when under the influence of alcohol or drugs. If this was the case, the criterion was not scored as positive for these analyses.

All interviewed probands (N = 591) and parents (N = 213, except for 1 mother) provided DNA samples. Among 610 affected siblings, there were only 6 individuals who did not provide DNA samples. Individuals not willing to donate blood and those interviewed by telephone were asked to contribute buccal epithelial cells collected with 4 cytology brushes. A total of 15% of the sample gave brush samples only. Blood samples from 66 volunteer control subjects recruited in Ireland were used to obtain allele frequency estimates. Controls were screened and excluded if they reported a history of heavy drinking or problem alcohol use.

Laboratory Methods

Details of the laboratory methods and error checking are provided elsewhere (Prescott et al., 2006). DNA was extracted using standard techniques. A 1,020-marker autosomal genomewide screen was conducted by deCODE Genetics (http://www.decode.com/) using their standard panel of polymorphic microsatellite markers. The average marker spacing was 4 cM, and the average heterozygosity in our sample was 72.5% (minimum 6.3%, maximum 91.8%). Genotyping reliability was assessed in 17 individuals for whom we submitted 2 blinded samples for genotyping. The between-sample agreement reliability was assessed in 17 individuals for whom we submitted 2 blinded samples for genotyping. The between-sample agreement ranged from 98.44% to 100% and averaged 99.68%.

The genome scan was conducted with samples from 1,423 individuals, including 1,357 family members and 66 control subjects. We used GRR (Abecasis et al., 2001) to check for familial relationships/nonmaternity and marker error. There were 39 families whose samples were excluded from analyses because 1 or more individuals had genotypes inconsistent with the reported family structure. After data cleaning, 1,289 samples were included in the linkage analyses reported here—474 families (2–12 members) with 81.1, 14.2, and 1.9% containing 2, 3, or 4 persons. There were 733 “all possible” sibling pairs.

The genotyping success rate among these individuals was 90.2%, reflecting 1,185,544 of a possible 1,314,780 genotypes (1,289 subjects×1,020 markers). In total, we had 1,081 valid samples for both univariate and bivariate linkage analyses.

Statistical Methods

The univariate CDsx and bivariate CDsx–ADsx multipoint linkage analyses were conducted in sequential oligogenic linkage analysis routines (SOLAR) 2.1.4 (Almasy and Blangero, 1998). A variance components method was utilized in the linkage models so that the variance and covariation of traits could be partitioned into that which results from the quantitative trait locus (QTL), residual or background additive genetics, and residual effects (nonshared environment and measurement error). The LOD score and parameter estimates of components were obtained from SOLAR. The bivariate LOD score was obtained based on the asymptotically mixture distributions of \( \chi^2_2, \beta_1, \) and \( \beta_2 \) under the null hypothesis of no linkage of either trait (Self and Liang, 1987).

Because the distributions of both CDsx and ADsx were relatively nonnormal, we transformed them using the Proc Rank procedure in SAS (SAS Institute, 2000) and examined in linkage analysis both the raw and rank-transformed data. In SOLAR, a function of covariate screening was featured to find relevant covariates in linkage analysis. Age and gender were significant covariates (proportion of variance due to them was 0.13) for CDsx, and gender was a significant covariate for ADsx as we included them in the linkage analyses. After transformation and covariate adjustment, the residual kurtosis was -0.33 for CDsx and -0.12 for ADsx—both within the normal range.

In the bivariate linkage analysis, SOLAR provides an estimate of the proportion of variance for each of the traits accounted for by the QTL. As we were interested in the relative specificity of the observed peaks for CD versus AD, we report a ratio of these 2 variances (AD/CD). Values substantially greater than and less than 1 reflect a QTL with specificity, respectively, for AD and for CD. Values close to 1, by contrast, indicate nonspecificity, with the QTL contributing approximately equally to liability to CD and AD.

We report raw \( p \) values for all linkage peaks. To evaluate whether the nonnormality of trait distribution would influence the accuracy of these values, we used the “lodadj” function featured in SOLAR to adjust the distribution of LOD scores under the assumption of trait normality by running a simulation of no linkage using 10,000 replicates. For the univariate linkage peaks obtained using raw data, we report pointwise empirical \( p \) values based on this simulated distribution.

RESULTS

In the IASPSAD sample, the means (SD) for the ADsx and CDsx were, respectively, 6.3 (1.1) and 2.8 (2.9). A total of 35.7% of the sample (42.7% of males and 22.8% of females) met DSM-IV criteria for CD. The correlation between ADsx and CDsx was modest: 0.19.

The results from our genomewide multipoint univariate linkage analyses of CDsx and ADsx using both the raw and the rank-transformed scores are shown in Figs. 1 and 2, respectively. The locations of all peak LOD scores \( \geq 1.00 \) along with the nearest marker and \( p \) values are listed in Table 1.

Using raw scores, 8 regions with LOD scores in excess of 1.00 are found for CDsx. The strongest evidence for linkage is on chromosome 1 with a maximal LOD score of 3.16 near marker D1S196 (\( p = 0.0001 \)). The next largest peak, with a peak LOD of 2.36, is on 14p (\( p = 0.001 \)). All the remaining peaks, on chromosomes 2, 7, 8, and 10, have LOD scores below 1.50 and \( p \) values ranging from 0.01 to 0.03. When rank-transformed CD symptom scores were used, the evidence for linkage increased in some locations (on chromosomes 2, 7, and 10) and decreased in other (chromosomes 1, 8, and 14) (Table 1).
The empirical p values agreed quite well with those obtained from SOLAR, being slightly more conservative for the lowest values (0.0005 and 0.002 for the regions on chromosomes 1 and 14, respectively) and nearly identical for the other values (e.g., p = 0.010, 0.022, and 0.012 for the 2 peaks on chromosome 2 and the peak on chromosome 7, respectively).

For ADsx, linkage analysis with the raw scores reveals only 3 regions with LOD scores over 1.00, the largest of which is 1.60 on 4q (p = 0.007). Empirical p values for these peaks were all slightly smaller than the parallel values obtained from SOLAR (e.g., p = 0.004 for the 4q peak). Using rank-transformed AD symptom scores results in a substantial increase in evidence for linkage to 4q (peak LOD of 2.80, p = 0.0003) as well as evidence for other linked regions on chromosomes 2 and 18. A comparison of Figs. 1 and 2 reveals little evidence for overlap of linkage peaks for ADsx and CDsx using either the raw or the rank-transformed scores.

The results of the bivariate linkage analysis, conducted on both raw and rank-transformed data, are seen in Fig. 3 and summarized in Table 2. Using raw data, 6 bivariate linkage peaks exceeded LOD scores of 1.0, with the strongest being 3.25 on chromosome 1 (p = 0.0001). These peaks could be divided into 3 groups in their specificity for AD and CD. First, only 1 peak—on 4q—possessed high specificity for AD (i.e., AD/CD variance ratio of ~40). Second, 2 peaks, on 2p and 14p, had high specificity for CD with AD/CD variance ratios of ≤0.02. Third, 3 peaks, on chromosomes 1, 7, and 10, possessed modest specificity for CD, with AD/CD variance ratios estimated to be between 0.18 and 0.26. This means that these regions were estimated to contribute ~4 to 5 times as much variance for CDsx as for ADsx.

Using the rank-transformed data, only 1 new region (on 2q) had a LOD score exceeding 1. The pattern of the specificity of these regions was similar to that seen with the raw scores, with one exception. The region on 2q had an AD/CD variance ratio of 2.14, indicating only moderate specificity for AD. The linkage peak on chromosome 4 remained highly specific for AD.

**DISCUSSION**

The first goal of these analyses was to determine whether the linkage signals we have found previously for AD in the IASPSAD reflected genes that predisposed to a general externalizing syndrome or were more specific in their...
impact on the liability to AD. By far the strongest linkage signal detected for ADsx was on chromosome 4q (Prescott et al., 2006). We replicated this linkage in these analyses, albeit at lower LOD scores, using a different statistical method. Our univariate linkage analysis of CDsx revealed no significant linkage in this 4q region. Our bivariate analysis also indicated that this linkage peak had high specificity for AD. The gene or genes responsible for the 4q linkage peak do not appear to alter liability to a general externalizing syndrome but rather appear to more specific in impacting on risk for AD. These analyses do not address the pharmacologic specificity of the genes in this region that impact on risk for AD. However, in another report, we examine linkage in our sample to symptoms of nicotine dependence (Sullivan et al., 2006) and find no significant linkage for this phenotype in the 4q region. This result provides further confirmation of our findings that the genes underlying this QTL are relatively specific in their impact on liability to AD.

Our original report also detected weaker evidence for linkage (LODs between 1.0 and 1.5) for ADsx in 4 regions: 2q37, 9q21, 9q34, and 18p11. Here, we found evidence for LOD scores of greater than 1 at 2 of these sites (chromosomes 2q and 18p) in our univariate analyses and also found LOD scores of 0.95 and 0.85 on chromosome 9q21 using, respectively, our rank and raw data. For the 2q peak, our bivariate analysis with rank-transformed data suggested that this had moderate specificity for AD but also impacted on CD risk. Neither the 9q21 nor the 18p peak exceeded LOD scores of 1 in our bivariate analysis. A small peak was seen in 9q with a maximum LOD of 0.55 and 0.66 in the raw and rank-transformed data, respectively. The evidence for linkage diminished modestly in most

### Table 1. Univariate Linkage Analysis Using Raw and Rank-Transformed Data for Symptoms of Conduct Disorder (CD) and Alcohol Dependence (AD)

<table>
<thead>
<tr>
<th>Ch</th>
<th>Position</th>
<th>Nearest marker</th>
<th>LOD</th>
<th>p Value</th>
<th>Ch</th>
<th>Position</th>
<th>Nearest marker</th>
<th>LOD</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>171</td>
<td>D1S196</td>
<td>3.16</td>
<td>0.0001</td>
<td>1</td>
<td>172</td>
<td>D1S196</td>
<td>2.86</td>
<td>0.0003</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>D2S193</td>
<td>1.44</td>
<td>0.0010</td>
<td>2</td>
<td>3</td>
<td>D2S193</td>
<td>1.71</td>
<td>0.0005</td>
</tr>
<tr>
<td>3</td>
<td>103</td>
<td>D2S2116</td>
<td>1.12</td>
<td>0.0023</td>
<td>3</td>
<td>103</td>
<td>D2S2116</td>
<td>1.14</td>
<td>0.0013</td>
</tr>
<tr>
<td>4</td>
<td>110</td>
<td>D4S1572</td>
<td>1.60</td>
<td>0.0007</td>
<td>4</td>
<td>110</td>
<td>D4S1572</td>
<td>1.60</td>
<td>0.0007</td>
</tr>
<tr>
<td>4</td>
<td>111</td>
<td>D4S1572</td>
<td>2.80</td>
<td>0.0003</td>
<td>4</td>
<td>111</td>
<td>D4S1572</td>
<td>1.60</td>
<td>0.0007</td>
</tr>
<tr>
<td>5</td>
<td>184</td>
<td>D4S1607</td>
<td>1.04</td>
<td>0.0029</td>
<td>5</td>
<td>184</td>
<td>D4S1607</td>
<td>1.04</td>
<td>0.0029</td>
</tr>
<tr>
<td>5</td>
<td>185</td>
<td>D4S1607</td>
<td>1.12</td>
<td>0.0024</td>
<td>5</td>
<td>185</td>
<td>D4S1607</td>
<td>1.12</td>
<td>0.0024</td>
</tr>
<tr>
<td>7</td>
<td>183</td>
<td>D7S2447</td>
<td>1.75</td>
<td>0.0005</td>
<td>7</td>
<td>183</td>
<td>D7S2447</td>
<td>1.17</td>
<td>0.0023</td>
</tr>
<tr>
<td>7</td>
<td>183</td>
<td>D7S2447</td>
<td>1.17</td>
<td>0.0023</td>
<td>7</td>
<td>183</td>
<td>D7S2447</td>
<td>1.75</td>
<td>0.0005</td>
</tr>
<tr>
<td>11</td>
<td>68</td>
<td>D11S4191</td>
<td>1.27</td>
<td>0.0018</td>
<td>11</td>
<td>68</td>
<td>D11S4191</td>
<td>1.27</td>
<td>0.0018</td>
</tr>
<tr>
<td>11</td>
<td>67</td>
<td>D11S4191</td>
<td>1.16</td>
<td>0.0016</td>
<td>11</td>
<td>67</td>
<td>D11S4191</td>
<td>1.16</td>
<td>0.0016</td>
</tr>
<tr>
<td>14</td>
<td>2</td>
<td>D14S277</td>
<td>1.25</td>
<td>0.0026</td>
<td>14</td>
<td>2</td>
<td>D14S277</td>
<td>2.36</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

Ch, chromosome; cM, centimorgan.

The second goal of this article was to determine whether regions of significant linkage for CDsx could be found in this sample and if so whether any of our findings replicated prior linkage results from this phenotype. Using raw CDsx scores, we found 8 regions with LOD scores in excess of 1.00. The 2 largest peaks had, respectively, LOD scores of 3.16 (on 1q) and 2.36 on 14p with p values, respectively, in the range of 0.0001 to 0.0005 and 0.001 to 0.002. The evidence for linkage diminished modestly in most
regions when examined using rank-transformed CD symptom.

Much controversy surrounds the interpretation of linkage results for complex traits. While our findings for CDsx, especially on 1q and 14p, do not reach the levels of significance required to be declared “significant,” they are almost certainly “suggestive” of linkage and therefore worthy of further inquiry (Lander and Kruglyak, 1995).

In the first of the 2 prior linkage studies of CD symptoms, Dick et al. (2004) examined retrospectively reported CDsx and CD diagnosis in 2,282 individuals from 114 families from the initial sample from the Collaborative Study on the Genetics of Alcoholism (COGA). In the second study of which we are aware, Stallings et al. (2005) examined CDsx in 249 adolescent and young adult sibling pairs from 191 families ascertained through adolescent probands in treatment for substance abuse and/or delinquency. Dick et al. (2004) reported positive multipoint LOD scores (which they defined as ≥ 1.50) on chromosomes 2, 3, 12, and 19 for the CD diagnosis and 1 and 19 for CD symptoms. By contrast, Stallings et al. (2005), using multipoint analyses, found LOD scores for CDsx in excess of 1 at 9q34 and 17q12.

Using the UCSC browser (http://genome.ucsc.edu), we determined that the marker nearest to the 2p11.2 peak reported by Dick et al. (D2S1331) maps to 113 cM from the p terminus in our sample, within 10 cM of a peak found with our raw CD symptom count at 103 cM (Table 1). No substantially positive LOD scores (e.g., > 0.50) are seen in our data for CDsx in positive regions reported by Dick et al. on chromosomes 3, 12, or 19 (Fig. 1). However, the marker nearest the positive LOD score reported by Dick et al. (D1S1606) only for CD diagnosis on chromosome 1 (D1S1606) maps to ~205 cM from the p terminus in our data—approximately 35 cM distal to the highest LOD score we find for both the raw and transformed CDsx (Table 1). No substantially positive LOD scores are seen in our data for CDsx in positive regions reported by Stallings et al. on chromosomes 9q and 17q (Fig. 1).

Given the large confidence intervals for linkage peaks in complex disorders with these sample sizes (Roberts et al., 1999), it could be argued that we have modest replications for 2 of the 5 regions for CD found by Dick et al. (2004) and 0 of 2 for the regions reported by Stallings et al. (2005). Methodologically, our study resembles the Dick et al. report in that both studies utilized retrospective reports of CD in families originally ascertained for a high density of AD much more than the study by Stallings et al. (2005). In aggregate, both our results and the findings of Dick et al. suggest that further efforts to localize susceptibility genes for CD might profitably focus on the regions of 1q and 2p that produced positive results in both studies.

Table 2. Bivariate Linkage Results Using Raw and Rank-Transformed Data for Symptoms of Conduct Disorder (CD) and Alcohol Dependence (AD)

<table>
<thead>
<tr>
<th>Raw data</th>
<th>Rank-transformed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ch</td>
<td>Position cM</td>
</tr>
<tr>
<td>1</td>
<td>170</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>110</td>
</tr>
<tr>
<td>7</td>
<td>180</td>
</tr>
<tr>
<td>10</td>
<td>150</td>
</tr>
<tr>
<td>14</td>
<td>5</td>
</tr>
</tbody>
</table>

Ch, chromosome; cM, centimorgan.
Very recently, Dick et al. (2006) reported in the COGA sample association with variants in the GABRA2 gene on chromosome 4p and risk for CD in adolescence and AD in adulthood. These intriguing findings illustrate the potential role of molecular genetic methods to clarify pathways of comorbidity. Our linkage results are not congruent with these findings—which would have predicted a linkage peak for both CDsx and ADsx in this region of 4p. However, the significance of these negative results is uncertain because association methods can detect considerably smaller effect sizes than is possible with linkage.

These results should be viewed in the context of 3 potential methodologic limitations. First, as is typical for complex trait linkage, our linkage peaks were relatively large and contain many genes. Our evidence for effects for ADsx and CDsx from the same QTL could result from the effect of different genes in proximity to one another. Second, the magnitude of most of the LOD scores obtained here is modest. Some of them may represent false-positive findings. Third, our affected sib pair sample was ascertained for AD. The results found for CD symptoms found in this particular group may not extrapolate to other samples of conduct-disordered individuals. Furthermore, as would be expected given its mode of ascertainment, this sample contained considerably greater variation for CDsx than for ADsx. Our bivariate analyses may have more closely resembled the univariate CDsx than the univariate ADsx analyses because of the greater amount of information contained in this sample for symptoms of CD than for symptoms of AD.

ACKNOWLEDGMENTS
Margaret Devitt, Lisa Halberstadt, and Victor Robinson made critical contributions to the project. We thank Ruth Barrington and the Irish Health Research Board for their support of the project.

REFERENCES