Family-based association analysis of alcohol dependence implicates KIAA0040 on Chromosome 1q in multiplex alcohol dependence families

Shirley Y. Hill*, Bobby L. Jones, Nicholas Zezza, Scott Stiffler

Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, USA
Email: *syh50@imap.pitt.edu

Received 27 September 2013; revised 10 October 2013; accepted 11 November 2013

Copyright © 2013 Shirley Y. Hill et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

Background: A previous genome-wide linkage study of alcohol dependence in multiplex families found a suggestive linkage result for a region on Chromosome 1 near microsatellite markers D1S196 and D1S2878. The KIAA0040 gene has been mapped to this region (1q24 - q25). A recent genome-wide association study using SAGE (the Study of Addiction: Genetics and Environment) and COGA (Collaborative Study on the Genetics of Alcoholism) found five SNPs within the KIAA0040 gene significantly associated with alcohol dependence. A meta-analysis using data from these sources also found the KIAA0040 gene significantly associated with alcohol dependence. Methods: Using family data consisting of 1000 individuals with phenotypic data (762 with both phenotype and DNA), finer mapping of a 0.3 cM region that included the KIAA0040 gene and a flanking gene was undertaken using SNPs with minor allele frequency (MAF) ≥ 0.15 and pair-wise linkage disequilibrium (LD) of $r^2 < 0.8$ using the HapMap CEU population. Results: Significant FBAT p-values were observed for six SNPs, four within the KIAA0040 gene and a flanking gene was undertaken using SNPs with minor allele frequency (MAF) ≥ 0.15 and pair-wise linkage disequilibrium (LD) of $r^2 < 0.8$ using the HapMap CEU population. Results: Significant FBAT p-values were observed for six SNPs, four within the KIAA0040 gene (rs2269650, rs2861158, rs1008459, rs2272785) and two adjacent to KIAA0040 (rs10912899 and rs3753555). Five haplotype blocks of varying size were identified using HAPLOVIEW. Analysis using the haplotype-based test function of FBAT revealed one two-SNP block (rs1008459-rs2272785) associated with alcohol dependence. This block showed a pattern of transmission in which one haplotype, CT, with a frequency of 0.577 was found to be over-transmitted to affected offspring (p = 0.017) while another haplotype, AG, with a frequency of 0.238 was found to be under-transmitted to affected offspring (p = 0.006). A three-SNP block (rs1008459-rs2272785-rs3753555) showed an overall significant association (p = 0.011) with alcohol dependence with the haplotype ACT over-transmitted to affected offspring (p = 0.016) and the haplotype GAG under-transmitted (p = 0.002). Conclusions: Family-based association analysis shows the KIAA0040 gene significantly associated with alcohol dependence. The potential importance of the KIAA0040 gene for AD risk is currently unknown. However, the present results support earlier findings from a genome-wide association study.

Keywords: KIAA0040; Alcohol Dependence; Multiplex Families

1. INTRODUCTION

Alcohol dependence is a complex disorder that is characterized by psychological and physical dependence and is often accompanied by chronic consumption of hazardous levels of ethanol. Excessive use of alcohol is the third leading cause of preventable death [1] in the US. The economic and social costs have been estimated to be $184 billion due to alcohol-related accidents, lost productivity, incarceration and other alcohol related morbidity [2]. In spite of the fact that the use of alcohol is quite common, a smaller proportion of the population drink in sufficient quantity and with associated health, family and work-related problems to be considered alcohol dependent AD. The one-year prevalence of AD in the US is 3.8% [3]. The lifetime prevalence of AD has been estimated at 12.5% [4]. Prevalence among male respondents ages 15 - 54 has been reported to be higher with 20.1% of men and 8.2% of women meeting criteria for alcohol dependence [5]. There is now evidence that those individuals with the greatest propensity for AD may carry an increased genetic risk for developing alcohol depend-
ence.

Although there is considerable heritability for alcohol dependence (0.49 - 0.64) in males [6,7] and females (0.56 - 0.59), [8,9] few genes have been identified that reliably confer susceptibility. However, studies employing well-designed sampling strategies over sample families with a high density of cases have revealed important clues for gene finding as seen in the Collaborative Study on the Genetics of Alcoholism (COGA) studies [10,11]. Genome-wide association (GWAS) studies have also revealed potentially important loci but require large samples to detect loci having genome-wide significance. A meta-analysis of two GWAS studies of alcohol dependence totaling 4979 cases and controls has identified three loci with statistical significance of $p < 5 \times 10^{-7}$.

In a genome-wide scan of multiplex families ascertained through a pair of affected probands [13], we found evidence for linkage in multiple chromosomal regions. The present report is based on efforts to follow up on linkage findings for a region on Chromosome 1q23.3 to 1q25.1 that includes a maximal LOD score of 3.46 ($p = 0.002$) at marker D1S196 and at an adjacent marker D1S2878 with a LOD value of 3.45 ($p = 0.002$). A previous follow-up of this region revealed significant family-based association for the astrotactin neuronal protein (ASTN1) gene [14]. Two nearby genes, KIAA0040 and TNN have been mapped to 1q25.1 and have been recently identified in genome-wide association analyses as being significantly related to alcohol dependence [12,15]. Accordingly, a study of the TNN/KIAA0040 region was undertaken.

2. MATERIALS AND METHODS

2.1. Study Sample

Written consent was obtained from all members of the multiplex families who participated in the study after the nature and purpose of the study was fully explained to them. The consent forms used in the study were approved by the University of Pittsburgh Institutional Review Board.

2.2. Multiplex Families

Multiplex families were selected on the basis of the presence of a pair of alcohol dependent brothers or sisters. The probands were selected from among individuals in treatment for alcohol dependence in the Pittsburgh area. Probands were eligible if they met DSM-III criteria for AD and had a same sex sibling who similarly met criteria for AD. Families were excluded if the probands or any first-degree relative were considered to have a primary diagnosis of drug dependence (preceded alcohol dependence onset by at least 1 year), or the proband or first-degree relative met criteria for schizophrenia, or a recurrent major depressive disorder. Probands and relatives with mental retardation or physical illness precluding participation were excluded. Complete details regarding participant selection may be seen in Hill et al. [13]. The majority of probands (80%) had three or more siblings who contributed DNA, consented to a clinical interview, and provided family history. These large sibships resulted in a total of 648 sib pairs within the proband generation. Across the generations, an average of 5.7 individuals per family was genotyped.

2.3. Generation I and II Diagnoses

All proband pairs and their cooperative relatives (siblings and parents) were personally interviewed using a structured psychiatric interview (Diagnostic Interview Schedule [DIS]). The DIS provides good reliability and validity [16] for alcohol dependence and alcohol abuse by DSM-III and IIIR criteria [17,18] the diagnostic criteria in place when the study began. The DIS also provides an alcoholism diagnosis by Feighner Criteria [19].

2.4. Generation III—Young Adult Assessment for DSM-IV Diagnoses

With the initiation of a third generation follow-up, offspring who had reached their 19th birthday were assessed using the Composite International Diagnostic Interview (CIDI) [20] to determine the presence or absence of a DSM-IV Axis I diagnosis. The CIDI-SAM (Substance Abuse Module) [21] was also administered in order to determine quantity, frequency, and pattern of drug and alcohol use. Interrater reliability for interviewers on the diagnostic instruments used in this study exceeded 90%.

2.5. SNP Selection

Previously, we carried out a genome-wide linkage analysis finding potentially important linkage results for multiple regions including Chromosome 1 [13]. Our study included genotyping in a 26.6 cM region on Chromosome 1 that centered on the microsatellite marker D1S196. A LOD score of 3.46 was obtained using a binary alcohol dependence phenotype and including relevant covariates (age, gender and the personality variable Constraint). Constraint from the Multidimensional Personality Questionnaire measures tendencies to inhibit impulse expression, rejection of unconventional behavior, and risk taking and with genetic variance of 0.58 in twins reared apart [22].

In order to investigate the region further, SNPs were chosen with minor allele frequency (MAF) $\geq 0.15$ and pair-wise linkage disequilibrium (LD) of $r^2 < 0.8$ using the HapMap CEU population at approximately 1 cM intervals in this region. The genotyping and analysis was
completed in three stages. First, we focused on a 19 cM region extending from rs7522166 to rs2816187. This region, bounded by these SNPs was chosen because rs7522166 is 7 cM proximal to D1S196 and rs2816187 is 13 cM distal to D1S196. We genotyped 18 SNPs at approximately 1 cM intervals in this region. Analysis of these 18 SNPs revealed the greatest statistical significance for rs2228008 located in the ASTN1 gene. Finer mapping of this gene at an average distance of 28.9 kb using twelve additional SNPs confirmed the significant result obtained for rs2228008 [14]. Because two nearby genes, KIAA0040 and TNN (1q25.1) have shown highly significant association with alcohol dependence in a GWAS study [15], a study of this region was planned. A total of 18 SNPs were selected for genotyping with 9 SNPs selected to cover a 0.3 cM region extending from rs12094153 to rs3753555 covering the TNN/KIAA0040 region at intervals of no greater than 5 kb, with 8 SNPs selected for their presence within the KIAA0040 gene. The SNPs selected for presence within the gene were chosen based on the reports of Wang et al. [12] and Zuo et al. [15]. Specifically, from among the SNPs evaluated by Zuo et al. [15], the five SNPs having the best statistical significance were chosen (rs6701037, rs6425323, rs1057302, rs1057239, and rs1894709). These SNPs were also reported to be significantly related to alcohol dependence in the Wang et al. [12] meta-analysis.

2.6. DNA Isolation and Genotyping

Genomic DNA was extracted from whole blood with a second aliquot prepared for EBV transformation and cryopreservation. PCR conditions were as described in Hill et al. [13]. Genotyping was completed on a Biotage PSQ 96 MA Pyrosequencer (Biotage AB, Uppsala, Sweden). Each polymorphism was analyzed by PCR amplification incorporating a biotinylated primer. Thermal cycling included 45 cycles at an annealing temperature of 60°C. The Biotage workstation was used to isolate the biotinylated single strand from the double strand PCR products. The isolated product was then sequenced using the complementary sequencing primer.

2.7. Quality Control

SNP genotyping quality control involved ongoing monitoring of SNP signals provided by Qiagen software. Output is provided using three categories for each SNP: pass, fail and check. Data analysis was performed for only those signals meeting the “pass” criterion. Signals that failed or were returned as needing further checking were rerun. If after 3 attempts the SNP did not meet the “pass” criterion, it was eliminated from the analysis and another SNP chosen as a replacement.

2.8. Statistical Methods

The sample included 133 pedigrees consisting of 1000 individuals (49% male and 51% female). Among the 1000 subjects, 542 were affected, 436 were unaffected, and 22 had unknown status.

2.9. Mendelian Inconsistency

The PedCheck program [23] was used to evaluate individual SNPs for Mendelian inconsistencies based on the pedigree structures. As a result of the evaluation, 36 marker genotypes from among the 13,656 were coded as missing to resolve the reported inconsistencies.

2.10. Hardy-Weinberg Equilibrium (HWE)

Estimates of population allele frequencies were calculated using MENDEL version 11 [24]. Files required by the MENDEL program were generated via the program Mega2 [25]. Marker allele frequencies were tested for departures from Hardy-Weinberg equilibrium using the allele frequency option in MENDEL. None of the 18 SNPs analyzed were found to have p-values below the Bonferroni adjusted threshold (<0.003) that would indicate significant HWE departures.

2.11. Genetic Maps

The Genetic Map Interpolator (GMI) software [26] was used to retrieve current physical map positions from Ensembl (Ensembl 68). These physical positions were then used to linearly interpolate genetic map positions based on the Rutgers Combined Linkage-Physical Map [27,28].

2.12. Family-Based Association Test (FBAT)

Transmission rates of marker alleles were examined using the family-based association test program, FBAT [29,30], assuming an additive genetic model with robust variance estimation (-e option) to account for the relatedness. This family-based method is a generalization of the transmission disequilibrium test (TDT) [31], which provides a valid test of association even if admixture is present. FBAT converts each pedigree into nuclear families, which are then treated as independent families for the test statistic calculation. Informative families consisting of parent-child trios are utilized in the FBAT analysis.

Generation I and II individuals were coded as affected if they met criteria for alcohol dependence by DSM-III criteria. Generation III individuals were coded as affected if they met criteria for alcohol abuse or dependence, or drug abuse or dependence. Choice of this broader phenotype for the third generation was based on the greater prevalence of drug use disorders in the third
Also, Generation I and II individuals had been selected for the family study based on the presence of primary alcohol dependence (if drug dependence was present it must have followed the alcohol dependence diagnosis by one year). However, analyses were also conducted using an alcohol abuse/dependence phenotype for the third generation.

2.13. Gamete-Competition (GC)

We also considered the gamete-competition model [32], a generalization of the transmission disequilibrium test (TDT), to investigate association of marker alleles with alcohol dependence. The gamete-competition model can be used to test for differences in transmission of marker alleles to affected individuals.

2.14. Haplotype Analysis

Linkage disequilibrium (LD) analysis was performed using the HAPLOVIEW program version 4.2 [33]. The LD block structure was defined by calculating D' values pairwise between SNPs. SNP haplotype blocks were created using the HAPLOVIEW default block determination method [34]. Additionally, a sliding window approach was used to identify two and three SNP blocks in order to insure that any blocks within larger haplotype blocks could be analyzed. Haplotype blocks were investigated for family-based association with affected status. A within-family association analysis between alcohol dependence and the revealed haplotypes was performed using haplotype FBAT [35] assuming an additive genetic model and using a robust estimate of variance (-e option).

3. RESULTS

3.1. Association Results

Analysis of 18 SNPs covering a 68.8 Kbp region on Chromosome 1 extending from rs12094153 to rs3753555 revealed six SNPs associated with alcohol dependence with significant FBAT p values (rs2269650, rs2861158, rs1008459, rs2272785, rs10912899, rs3753555). The SNP showing the most significant association with alcohol dependence affected status was rs1008459 (FBAT p = 0.006) located within intron 2 of KIAA0040. Four SNPs are within the KIAA0040 gene. Two of these were also found to be significant using Gamete Competition (GC) analyses. Results for the FBAT and GC analyses are summarized in Table 1. LocusZoom [36] was used to generate a plot of the association test results (Figure 1).

3.2. Haplotype Analysis

Five haplotype blocks were identified by HAPLOVIEW. Pairwise linkage disequilibrium between the SNPs and the LD block structure are shown in Figure 2. Haplotype analyses were performed using two alternative phenotypes for Generation III (see Table 2). The gender distribution by generation may be seen in Table 3. Results of the haplotype analysis can be seen in Table 4. One two SNP block (Block 5) which consisted of one SNP within the KIAA0040 gene (rs2272785) and an adjacent SNP (rs3753555) showed an association with affected status with a p value of 0.041 using a broader SUD phenotype and 0.034 when restricted to alcohol abuse or dependence only. In Block 5, the haplotype CT with a frequency of 0.577, was found to be over-transmitted to affected offspring (p = 0.017) while the haplotype block AG with a frequency of 0.238, was found to be under-transmitted.
### Table 1. Results of FBAT and GC analyses for alcohol dependence in generations I and II with any substance use disorder coded as positive for generation III.

<table>
<thead>
<tr>
<th>Marker</th>
<th>cM</th>
<th>Informative families</th>
<th>FBAT p-value</th>
<th>GC p-value</th>
<th>Major/Minor Allele</th>
<th>Nucleotide</th>
<th>Major Allele Frequency</th>
<th>Minor Allele Frequency</th>
<th>Allele Freq. Affected</th>
<th>Allele Freq. Unaffected</th>
<th>Gene</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs12094153</td>
<td>184.752</td>
<td>44</td>
<td>0.591</td>
<td>0.584</td>
<td>1</td>
<td>G</td>
<td>G = 0.522</td>
<td>A = 0.478</td>
<td>0.603</td>
<td>0.397</td>
<td>TNN</td>
<td></td>
</tr>
<tr>
<td>rs4651322</td>
<td>184.753</td>
<td>41</td>
<td>0.438</td>
<td>0.310</td>
<td>1</td>
<td>C</td>
<td>C = 0.673</td>
<td>T = 0.327</td>
<td>0.615</td>
<td>0.385</td>
<td>TNN</td>
<td></td>
</tr>
<tr>
<td>rs12563833</td>
<td>184.760</td>
<td>40</td>
<td>0.065</td>
<td>0.051</td>
<td>2</td>
<td>A</td>
<td>G = 0.743</td>
<td>A = 0.257</td>
<td>0.289</td>
<td>0.711</td>
<td>TNN</td>
<td></td>
</tr>
<tr>
<td>rs1018829</td>
<td>184.765</td>
<td>35</td>
<td>0.086</td>
<td>0.389</td>
<td>1</td>
<td>G</td>
<td>C = 0.832</td>
<td>T = 0.168</td>
<td>0.770</td>
<td>0.230</td>
<td>TNN</td>
<td></td>
</tr>
<tr>
<td>rs2018318</td>
<td>184.763</td>
<td>33</td>
<td>0.177</td>
<td>0.488</td>
<td>1</td>
<td>C</td>
<td>G = 0.826</td>
<td>T = 0.174</td>
<td>0.791</td>
<td>0.209</td>
<td>TNN</td>
<td></td>
</tr>
<tr>
<td>rs6701037</td>
<td>184.769</td>
<td>43</td>
<td>0.526</td>
<td>0.512</td>
<td>1</td>
<td>A</td>
<td>C = 0.518</td>
<td>A = 0.482</td>
<td>0.566</td>
<td>0.434</td>
<td>TNN</td>
<td></td>
</tr>
<tr>
<td>rs10912899</td>
<td>184.766</td>
<td>44</td>
<td><strong>0.021</strong></td>
<td>0.255</td>
<td>2</td>
<td>T</td>
<td>G = 0.650</td>
<td>A = 0.350</td>
<td>0.411</td>
<td>0.589</td>
<td>TNN</td>
<td></td>
</tr>
<tr>
<td>rs6425323</td>
<td>184.768</td>
<td>43</td>
<td>0.469</td>
<td>0.506</td>
<td>1</td>
<td>C</td>
<td>T = 0.518</td>
<td>C = 0.482</td>
<td>0.570</td>
<td>0.430</td>
<td>TNN</td>
<td></td>
</tr>
<tr>
<td>rs1057302</td>
<td>184.769</td>
<td>43</td>
<td>0.469</td>
<td>0.540</td>
<td>1</td>
<td>A</td>
<td>C = 0.513</td>
<td>T = 0.487</td>
<td>0.575</td>
<td>0.425</td>
<td>KIAA0040 Exon 5</td>
<td></td>
</tr>
<tr>
<td>rs2269650</td>
<td>184.771</td>
<td>45</td>
<td><strong>0.037</strong></td>
<td>0.076</td>
<td>2</td>
<td>A</td>
<td>C = 0.677</td>
<td>A = 0.323</td>
<td>0.336</td>
<td>0.664</td>
<td>Exon 5</td>
<td></td>
</tr>
<tr>
<td>rs1057239</td>
<td>184.779</td>
<td>43</td>
<td>0.604</td>
<td>0.160</td>
<td>1</td>
<td>G</td>
<td>T = 0.509</td>
<td>C = 0.491</td>
<td>0.559</td>
<td>0.441</td>
<td>KIAA0040 Exon 5</td>
<td></td>
</tr>
<tr>
<td>rs1894709</td>
<td>184.773</td>
<td>43</td>
<td>0.497</td>
<td>0.098</td>
<td>1</td>
<td>G</td>
<td>A = 0.509</td>
<td>C = 0.491</td>
<td>0.557</td>
<td>0.443</td>
<td>KIAA0040 Intron 4</td>
<td></td>
</tr>
<tr>
<td>rs2269655</td>
<td>184.779</td>
<td>35</td>
<td>0.272</td>
<td>0.617</td>
<td>1</td>
<td>G</td>
<td>G = 0.881</td>
<td>T = 0.119</td>
<td>0.824</td>
<td>0.181</td>
<td>KIAA0040</td>
<td></td>
</tr>
<tr>
<td>rs2861158</td>
<td>184.776</td>
<td>37</td>
<td><strong>0.033</strong></td>
<td>0.052</td>
<td>2</td>
<td>A</td>
<td>G = 0.743</td>
<td>A = 0.257</td>
<td>0.245</td>
<td>0.755</td>
<td>KIAA0040 Exon 4</td>
<td></td>
</tr>
<tr>
<td>rs16447872</td>
<td>184.779</td>
<td>35</td>
<td>0.266</td>
<td>0.648</td>
<td>1</td>
<td>T</td>
<td>A = 0.881</td>
<td>G = 0.119</td>
<td>0.836</td>
<td>0.164</td>
<td>KIAA0040 Intron 3</td>
<td></td>
</tr>
<tr>
<td>rs1008459</td>
<td>184.787</td>
<td>37</td>
<td><strong>0.006</strong></td>
<td><strong>0.013</strong></td>
<td>2</td>
<td>G</td>
<td>A = 0.735</td>
<td>G = 0.265</td>
<td>0.268</td>
<td>0.732</td>
<td>KIAA0040 Intron 2</td>
<td></td>
</tr>
<tr>
<td>rs2272785</td>
<td>184.795</td>
<td>38</td>
<td><strong>0.015</strong></td>
<td><strong>0.031</strong></td>
<td>2</td>
<td>A</td>
<td>C = 0.712</td>
<td>A = 0.288</td>
<td>0.268</td>
<td>0.733</td>
<td>KIAA0040 Intron 2</td>
<td></td>
</tr>
<tr>
<td>rs3753555</td>
<td>184.796</td>
<td>45</td>
<td><strong>0.021</strong></td>
<td>0.082</td>
<td>2</td>
<td>G</td>
<td>A = 0.588</td>
<td>C = 0.422</td>
<td>0.448</td>
<td>0.552</td>
<td>KIAA0040</td>
<td></td>
</tr>
</tbody>
</table>

1. major and 2. minor allele (designation based on NCBI build 37.3 allele frequencies for European Caucasians). Bolded nucleotides are the reverse strand. Using a more restrictive phenotype for Generation III that included only alcohol abuse and alcohol dependence p-values are: rs10912899, p = 0.033; rs2269650, p = 0.033; rs2861158, p = 0.037; rs1008459, p = 0.014; rs2272785, p = 0.062; rs3753555, p = 0.020.
Table 2. Diagnostic status by generation.

<table>
<thead>
<tr>
<th>Generation</th>
<th>Age Mean ± SD</th>
<th>Affected</th>
<th>Unaffected</th>
<th>Unknown</th>
<th>Total</th>
<th>Genotyped</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generation I</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol Dependence</td>
<td>59.4 (9.1)</td>
<td>115</td>
<td>127</td>
<td>6</td>
<td>248</td>
<td>121</td>
</tr>
<tr>
<td>Generation II</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol Dependence</td>
<td>36.4 (7.8)</td>
<td>335</td>
<td>162</td>
<td>6</td>
<td>503</td>
<td>392</td>
</tr>
<tr>
<td>Generation III</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUD Phenotype</td>
<td>23.9 (5.8)</td>
<td>92</td>
<td>147</td>
<td>10</td>
<td>249</td>
<td>248</td>
</tr>
<tr>
<td>Alcohol Abuse/Dependence Phenotype</td>
<td>23.9 (5.8)</td>
<td>47</td>
<td>195</td>
<td>7</td>
<td>249</td>
<td>248</td>
</tr>
<tr>
<td>Total</td>
<td>542</td>
<td>436</td>
<td>22</td>
<td>1000</td>
<td>761</td>
<td></td>
</tr>
</tbody>
</table>

Three individuals from an ancestral generation (great grandparents of third generation subject) were also genotyped, but are not included in the tables.

Table 3. Gender distribution by generation for individuals with phenotype.

<table>
<thead>
<tr>
<th>Generation</th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generation I</td>
<td>124</td>
<td>124</td>
<td>248</td>
</tr>
<tr>
<td>Generation II</td>
<td>250</td>
<td>253</td>
<td>503</td>
</tr>
<tr>
<td>Generation III</td>
<td>115</td>
<td>134</td>
<td>249</td>
</tr>
<tr>
<td>Total</td>
<td>490</td>
<td>511</td>
<td>1000</td>
</tr>
</tbody>
</table>

Table 4. Haplotype analysis.

<table>
<thead>
<tr>
<th>Markers</th>
<th>Haplotype</th>
<th>Block</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2272785-rs3753555</td>
<td>C-T</td>
<td>0.577</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>A-G</td>
<td>0.238</td>
<td>0.006</td>
</tr>
<tr>
<td>rs1008459-rs2272785-rs3753555</td>
<td>A-C-T</td>
<td>0.585</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>G-A-G</td>
<td>0.205</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Note: Affected status for Generation III includes any SUD; Affected status for Generation III is alcohol dependence only.

4. DISCUSSION

Within-family association (FBAT and GC) analyses were performed for 18 SNPs in a region of Chromosome 1. Based on the FBAT within-family association analyses, our results suggest that variation in the KIAA0040 gene is associated with risk for alcohol dependence in families with multiple cases of alcohol dependence. These results support findings from a genome-wide association study [15] and a meta-analysis that includes data from studies that utilizes both case/control and within family association analyses for alcohol dependence for SNPs within the KIAA0040 gene [12].

Zuo et al. [15] reported significant results for five SNPs within the KIAA0040 gene (rs10572239, rs1894709, rs6701037, rs6425323, rs1057302). It is noteworthy that this GWAS study also finds one SNP having only alcohol dependence was present showed an overall probability of (p = 0.011) and a haplotype specific p-value for GAG of 0.003.
Cis-Acting regulatory effects and the rs2269650 SNP with a p-value of $8.6 \times 10^{-5}$. This SNP showed significant results in the present family-based association analysis as well. However, the five top ranked SNPs reported in Zuo et al. [15] were not significant in our within-family association analysis. Two of these SNPs lie in the region proximal to the KIAA0040 gene (rs6701037 and rs6425323) and distal to the TNN gene while three other SNPs lie in Exon 5 (rs1057302 and rs1057239) and intron 4 (rs1894709). Consideration of the meta-analysis of family data provided by Wang et al. [12] shows a replication in the present family data for one SNP (rs1008459) with a reported FBAT value of 0.0367. However, it should be noted that three SNPs (rs6701037, rs2269655, and rs6425323) that reached genome-wide significance in the analysis of Wang et al. [12] were not significant in the present study. However, our results for rs10912899 which lies between two of these SNPs, rs6701037 and rs6425323, did show significance ($p = 0.02$) based on our FBAT analysis.

Haplotype analysis revealed one two-SNP block with a p-value of 0.006 and one three-SNP block with a p-value of 0.002. The Zuo et al. [15] SNPs showing genomewide significance appear to cluster in Exon 5, whereas the current results also suggest the importance of Exon 4 (rs2861158) and intron 2 where we found a two-SNP and a three-SNP haplotype respectively with p-values suggesting their importance.

The biological significance of the current findings is unknown because the KIAA0040 gene encodes a protein whose function is unknown. There is evidence that the KIAA0040 protein product may represent one of the tumor antigens expressed on colorectal cancer cells and recognized by tumor reactive T-cells (CT28 line) [37]. The KIAA0040 gene is flanked by two genes TNN and TNR with a plausible role in alcohol dependence. A previous study has reported significance of SNPs within the TNN gene and AD [15]. Because the TNN gene lies only 8.9 kb from KIAA0040, the TNN gene is of interest. Also, the TNN gene encodes a protein, tenascin-N, which is involved in neurite outgrowth and cell migration in the hippocampus [38]. Weaker evidence for a role of the TNN gene in AD was seen in the present analysis with two SNPs (rs12563833 and rs1018829) showing marginal FBAT significance.

The present results should be considered in the context of some limitations. Although this study represents a follow-up on a linkage peak originally reported for this region of Chromosome 1, the peak observed was relatively large [13]. Because the peak was broad, it may be expected that a number of genes are within this peak. KIAA0040 was not included in our original planned analysis as its function has not been defined. However, KIAA0040 has shown genome-wide significance for alcohol dependence in a large case/control data set [15].

Another issue concerns whether or not the present family-based findings did, indeed, replicate the top GWAS SNPs reported by Zuo et al. [15] and Wang et al. [12]. Of the five SNPs reported as reaching genome-wide significance, two SNPs, rs10912899 and rs2269650, in the present study were within 300 - 400 base pairs of two SNPs reported by Zuo et al. [15] and Wang et al. [12] to have genome-wide significance (rs6701037 and rs1057302). It is noteworthy that Wang et al. [12] reporting on their meta-analytic family study data were not able to completely replicate individual SNPs from the GWAS findings. Of the four SNPs reported by Wang et al. [12], one was not significant, and two had nominal p-values, though one was highly significant.

The issue of representativeness of our findings may be considered a limitation. The multiplex families on which the present report is based were ascertained through affected sib pairs. Multiplex families appear to differ from alcohol dependent families in the general population by having greater transmission of alcohol dependence across generations. Follow-up of third generation offspring from multiplex families shows an exceptionally high rate of AD and associated substance use by young adulthood [39,40]. This suggests that multiplex family samples may provide an efficient means of identifying genes because of the greater likelihood that genes may be segregating within these families that confer greater susceptibility to early onset alcohol dependence and related substance use disorders [41].

Another possible limitation of our results is that some of the diagnoses were based on DSM-III criteria and others on DSM-IV. The DSM-III system was the current system in place when Generation I and II was recruited. The definition of alcohol dependence provided by DSM-III requires the presence of tolerance and physical dependence. With the initiation of a longitudinal follow-up for Generation III, subjects were assessed using DSM-IV criteria. These criteria require three or more symptoms within a 12-month period that may include tolerance or physical dependence but do not require the presence of these symptoms if other symptoms are present. Accordingly, the third generation may have met criteria for alcohol dependence based on social or occupational impairment, use in spite of physical impairment, persistent desire to use, or inability to cut down or control use.

An additional limitation of our analysis was the need to include a broader phenotypic definition for the Generation III individuals. While it would be possible to code any Generation III individual as unaffected if they did not meet criteria for alcohol dependence, it appeared that this would incorrectly reflect individual’s addiction susceptibility where significant abuse or dependence on drugs was present. An additional analysis in which the
third generation offspring was coded for alcohol dependence only (yes/no) provided essentially the same results.

Because the power to detect association increases with the number of observations available for related individuals [42], the family-based association analysis was strengthened by having a larger number of family members that includes three generations. However, the younger age (mean of 24 years) of the third generation family members means that not all have moved through the period of risk. As a result, some individuals coded as not dependent may eventually convert to affected status. However, in spite of this limitation we observed a significant relationship between SNPs within the KIAA0040 gene and alcohol dependence suggesting that this gene may have clinical importance in the etiology of alcohol dependence.

A comment is needed regarding the family-based approach which was taken in this report. Risch and Merikangas [43] were among the first to suggest that association studies are sometimes more powerful than linkage analyses. Since that time, there has been a shift toward large genome-wide association studies (GWAS) instead of family-based methods where sample sizes are more modest. Some have questioned whether GWAS methods that are designed to detect common rather than rare variants will explain a substantial portion of heritability in psychiatric disorders [44]. This view has been amplified by others who argue that GWAS may detect common variants with statistically significant results but only modest population attributable risk in comparison to focused investigations of families where genes can be found with high predictive value [45]. Perhaps, the use of multiple statistical genetic methods is preferable when focused investigations of families where genes can be found with high predictive value [45].

In summary, the present results using family-based association found evidence that the KIAA0040 gene is related to risk for alcohol dependence and supports the GWAS results offered by Zuo et al. [15]. Future work is needed to uncover the function of this gene and its potential role in the risk/protection from alcohol dependence.

5. ACKNOWLEDGEMENTS

The authors wish to acknowledge the continued support of Daniel E. Weeks, Ph.D. in data analysis and interpretation. Also, we wish to acknowledge the contribution of the family members who have contributed DNA and continue to be involved in clinical follow up. Also, we want to acknowledge the contributions of many clinical staff in assessing research participants to determine phenotypic data. A special thanks to Brian Holmes for assistance in manuscript preparation. The financial support for the study reported in this manuscript was provided by NIAAA Grants AA018289, AA05909, AA08082, and AA015168 to SYH. The authors have no conflicts of interest to report.

REFERENCES

http://dx.doi.org/10.1001/jama.291.3.293
http://dx.doi.org/10.1016/j.drugalldip.2004.02.004
http://dx.doi.org/10.1001/archpsyc.64.7.830
http://dx.doi.org/10.1001/archpsyc.1997.0183016003100
http://dx.doi.org/10.1007/BF01066682
http://dx.doi.org/10.1017/S0033291797005643

Copyright © 2013 SciRes.
http://dx.doi.org/10.1001/jama.1992.03490140085040

http://dx.doi.org/10.1002/(SICI)1096-8628(19990508)81:3<207::AID-AJMG1>3.0.CO;2-T

http://dx.doi.org/10.1086/383283


man Genetics, 66, 1168-1172. http://dx.doi.org/10.1086/302826


