Genomewide linkage survey of nicotine dependence phenotypes

Patrick F. Sullivan a,b,*, Po-Hsiu Kuo c,d, B. Todd Webb d, Michael C. Neale d, Jen Vittum d, Helena Furberg a, Dermot Walsh e, Diana G. Patterson f, Brien Riley d, Carol A. Prescott g, Kenneth S. Kendler d

a Departments of Genetics, Psychiatry, & Epidemiology, University of North Carolina, Chapel Hill, NC, USA
b Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden
c Institute of Clinical Medicine, College of Medicine, National Cheng Kung University, Taiwan
d VIPBG, Virginia Commonwealth University, Richmond, VA, USA
e Health Research Board, Dublin, Ireland
f Shaftsbury Square Hospital, Belfast, Northern Ireland
g Department of Psychology, University of Southern California, Los Angeles, CA, USA

Received 28 July 2006; received in revised form 9 September 2007; accepted 11 September 2007
Available online 19 November 2007

Abstract

Background: A comprehensive understanding of the etiology and neurobiology of nicotine dependence is not available. We sought to identify genomic regions that might contain etiologically-relevant loci using genomewide univariate and bivariate linkage analyses.

Methods: We conducted secondary data analyses of 626 all possible sibling pairs ascertained in Ireland and Northern Ireland on the basis of alcohol dependence. A set of 1020 short tandem repeat genetic markers were genotyped in all subjects. The phenotypes analyzed were the Fagerström Test for Nicotine Dependence (FTND), a history of nicotine dependence, the number of symptoms of alcohol dependence (AlcSx), and a history of alcohol dependence. Genomewide linkage analyses were conducted with non-parametric and variance components methods.

Findings: For the bivariate variance component analysis of the continuous FTND and AlcSx scores, multipoint LOD scores were >4 in two genomic regions—an 11 cM region on chr7 (D7S2252–D7S691, empirical p = 0.0006) and an 8 cM region on chr18 flanking D18S63 (empirical p = 0.0007). These findings did not exceed a conservative estimate of study-wide significance. The remaining sets of findings had considerably smaller or less consistent peak signals. Notably, strong linkage signal at D4S1611 for AlcSx from a prior report (PMID 16534506) was not found when jointly analyzed with FTND.

Interpretation: Replication is required. However, chromosomes 7 and 18 may contain genetic loci relevant to the etiology of nicotine-related phenotypes.

© 2007 Elsevier Ireland Ltd. All rights reserved.

Keywords: Nicotine dependence; Genomewide linkage analysis

1. Introduction

Cigarette smoking is a first-rank public health problem (US Department of Health and Human Services, 1989; Centers for Disease Control, 1994b; World Health Organization, 1997). Although the prevalence of cigarette use in the US and so-called “first-world” countries has declined dramatically since 1964 (US Department of Health and Human Services, 1964), there remains a core group of smokers who rarely achieve sustained cessation (Glasgow and Orleans, 1997) despite often evidencing both the desire to quit (Centers for Disease Control, 1994a) and making serious attempts at cessation (US Department of Health and Human Services, 1989; US Department of Health and Human Services, 1990). In many other countries, the prevalence of smoking is rising alarmingly (World Health Organization, 2006).

One line of inquiry into developing new methods for understanding and treating nicotine dependence is to adapt approaches commonly used for elucidating the genetic basis of other complex human traits of biomedical importance (e.g., genomewide linkage and association studies). The rationale for
these approaches are elucidated at length elsewhere—briefly, these studies are predicated on observations that various smoking behaviors are heritable, particularly the core phenotype of nicotine dependence (Sullivan and Kendler, 1999; Li et al., 2003a). For example, a detailed analysis of the literature estimated the heritability in liability of nicotine dependence (or its proxies) to be 0.67 with the remaining variance from shared environmental (0.02) and individual-specific environmental effects (0.31) (Sullivan and Kendler, 1999).

We are aware of 11 published genomewide linkage studies of smoking-related phenotypes from seven independent samples. One sample was ascertained on the basis of smoking behavior (Straub et al., 1999; Sullivan et al., 2004). Four samples informative for smoking behavior were created from population-based parent studies in Framingham, MA (Goode et al., 2003; Li et al., 2003b; Wang et al., 2005), Mission Indians near San Diego, CA (Ehlers and Wilhelmsen, 2006), a longitudinal study in California (Swan et al., 2006), and the Australian Twin Registry (Morley et al., 2005). The remaining studies were all from secondary data analyses of samples ascertained on the basis of alcohol dependence (as with this report) (Bergen et al., 1999; Duggirala et al., 1999; Bierut et al., 2004) or panic disorder (Gelernter et al., 2004). The autosomal findings from these studies are depicted in Fig. 1. A few genomic regions overlap but the extant data are not strikingly consistent.

Our goal in this report was to attempt to add meaningfully to this literature by conducting a secondary data analysis of genomewide linkage data for nicotine dependence-related phenotypes in a large sample genotyped with a relatively large number of microsatellite markers. Ascertainment and primary...
2. Methods

2.1. Clinical data collection

Clinical data collection occurred between 1998 and 2002 as a joint collaboration among Virginia Commonwealth University, the Health Research Board in Dublin, and Shafsbury Square Hospital in Belfast (Prescott et al., 2005, 2006). All interviewed participants provided informed consent prior to 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).

Probands were selected 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 for inclusion if they met DSM-IV criteria for alcohol dependence (American Psychiatric Association, 1994) and if all four grandparents had been born in Ireland, Northern Ireland, Scotland, Wales, or England. In addition, the proband was required to have at least one full sibling who met the same inclusion criteria. Individuals with other substance dependence and psychiatric disorders were not excluded but we did assess the chronological relationship between the onset of these disorders and alcohol dependence. 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. The symptoms of nicotine dependence (during the period of lifetime maximum tobacco use) were assessed with an adapted version (Kendler et al., 1999) of the Fagerström Test for Nicotine Dependence (FTND) (Heatherton et al., 1991). A history of nicotine dependence (ND) was considered present if the subject had an FTND score ≥7 (Kendler et al., 1999). The phenotypic status of individuals who had never smoked was considered as unknown. DSM-IV alcohol dependence (AD) and the number of symptoms of alcohol dependence (AlcSx, based on the DSM-IV alcohol dependence criteria, range 3–7) during a subject’s lifetime period of heaviest use were assessed with the SSAGA interview (version 11) (Bucholz et al., 1994).

2.2. Laboratory

Laboratory Methods are described in detail elsewhere (Prescott et al., 2006). Genotyping was blinded to all phenotypic data. Briefly, an autosomal genomewide screen was conducted by deCODE Genetics (http://www.decode.com/) using a panel of 1020 microsatellite markers (average spacing of 4 cM) with an average heterozygosity of 0.725 (range 0.063–0.918) as calculated from our linkage sample using PEDSTATS (Wigginton and Abecasis, 2005). Based on 17 duplicated samples, between-sample agreement averaged 99.68% (range 98.44–100%).

Full detail of the data cleaning process is provided in (Prescott et al., 2006). Briefly, data cleaning was accomplished using GRR (Abecasis et al., 2001) and Merlin (Abecasis et al., 2002). After data cleaning, 1289 samples were included in the linkage analyses reported here—470 families (2–12 members) with 81.1%, 14.2%, and 1.9% containing 2, 3, or 4 persons. There were 626 all possible sibling pairs.

2.3. Statistical analyses

Univariate singlepoint and multipoint linkage analyses were performed using Merlin (Abecasis et al., 2002). Singlepoint analysis uses linkage information from each marker independently whereas multipoint analysis incorporates linkage information from nearby markers. Marker allele frequencies were estimated from 66 blood bank volunteers from Ireland and Northern Ireland. For ND, we used non-parametric linkage and we report the NPL LOD and asymptotic p-values generated by Merlin. For FTND symptom counts, we used Merlin–Regress. This method can be applied to selected samples but requires specification of the trait distribution parameters in the general population. We used the trait distribution (mean = 4.32 and variance = 7.24) from Virginia adult twin samples and weighted the gender ratio according to that of the present sample.

Bivariate linkage analyses (adjusted for sex and age) were conducted for both binary (ND–AD) and ordinal traits (FTND score–AlcSx) using Mx (Neale et al., 2004). Due to computational constraints, we only conducted multipoint linkage analyses. Variance components linkage analyses were applied in the models that partitioned the co-variance of the two traits into quantitative trait locus (QTL) effects, familial effects (residual additive genetic and shared environment), and residual effects (non-shared environment and measurement error). We assumed that the underlying liability of the two traits was bivariate normal and a maximum likelihood method was used for parameter estimation. Differences between log likelihood ratios of models with and without QTL effects were used to estimate the 2 d.f. bivariate LOD scores. We transformed these scores to make them comparable to the 1 d.f. univariate LOD score, and obtained the corresponding p-value based on the asymptotic mixture distributions of χ²1, χ²2, and χ²3 under the null hypothesis of no linkage for either trait (Self and Liang, 1987). Empirical p-values for linkage peaks were obtained by permutation for the four principal linkage peaks (due to computational limitations) (North et al., 2002). Bivariate analyses were also corrected for the ascertainment bias inherent in selecting on the lifetime presence of AD. Not performing such a correction would tend to underestimate the co-variance between the two traits. Correction for ascertainment bias requires population data on the probability of AD; as these epidemiological data were not available for Ireland, we used US lifetime prevalence estimates of 20.2% for males and 8.2% for females (Kessler et al., 1994). However, the use of US data is qualified by the known differences in yearly capita alcohol consumption (US 8.511, UK 10.41, and Ireland 14.51) (World Health Organization, 2004).

The study-wide significance level is set to be 1 × 10⁻⁵, approximately the results of dividing 0.05 by the number or markers (1020) times the number of statistical tests (4). This is quite conservative as no account is made for correlations among markers due to linkage disequilibrium and co-variance among the phenotypes studied. Empirical “gene-dropping” estimations were not practical due to computing constraints.

3. Results

For the subjects included in these analyses (N = 1289), the mean age was 41.9 years (S.D. 9.7), 64.3% were male, nearly all met criteria for lifetime AD (there were seven subjects who barely missed meeting criteria for AD), and 61.9% met criteria for ND. The median FTND score was 7 with an interquartile range of 5–9. The median AlcSx count was 7 with an interquartile range of 6–7. There was a significant albeit modest phenotypic correlation between FTND and AlcSx (Spearman ρ = 0.15, p < 0.0001). In a phenotypic multivariate model, the FTND score during the lifetime period of maximal smoking was significantly predicted by younger age at interview (p = 0.02), greater AlcSx (p = 0.003), and the sex × AlcSx interaction (p = 0.03, at comparable levels of AlcSx, males have higher...
FTND scores than women). Additional information about this sample is presented elsewhere (Prescott et al., 2005).

Fig. 1 shows the four sets of linkage results for this investigation. For all four sets of analyses, lines plot the multipoint results. Multipoint linkage analyses should generally provide the most powerful analyses of the linkage data as they maximize the use of information from each marker plus nearby markers (Elston, 1992; Kruglyak and Lander, 1995; Weeks et al., 1995; Kruglyak et al., 1996). Due to computational constraints, the marker-by-marker singlepoint results are presented only for ND and the FTND measure. Note that singlepoint and multipoint results are usually correlated—singlepoint–multipoint Spearman $\rho$ for ND was 0.73 and 0.74 for FTND). A more complete discussion of singlepoint and multipoint linkage analyses is provided elsewhere (Sullivan et al., 2003).

Fig. 1 depicts the results of univariate genomewide linkage analyses of FTND and ND and bivariate analyses of FTND–AlcSx and ND–AD in the context of all prior genomewide linkage studies. Table 1 provides the numeric results for the top 1% of scores in any phenotype and analytic method. All results are shown for any position that had at least one score in the top 1%.

For univariate analyses of continuous FTND scores, the maximum singlepoint NPL LOD score was 2.11 at D22S1177 (chr22, 45.348 cM); however, as all other markers from chr22, 25–65 cM had LODs <0.5, the multipoint NPL LOD score at this position was 0.08. The maximum multipoint NPL LOD score for the FTND score was 1.67 (singlepoint 0.29) at D3S2385 (chr3, 33.554 cM). In univariate analyses of lifetime ND, the maximum singlepoint NPL LOD score was 2.05 at D11S903 (chr11, 60.942 cM). The multipoint ND analyses also highlighted the region that was most significant in the ND singlepoint analyses along with an additional region (chr22, 17.6–24.5 cM, peak multipoint NLP LOD of 3.02). Univariate analyses of AlcSx and AD

<table>
<thead>
<tr>
<th>Chr</th>
<th>Marker</th>
<th>FTND singlepoint</th>
<th>FTND multipoint</th>
<th>ND singlepoint</th>
<th>ND multipoint</th>
<th>FTND-AlcSx multipoint</th>
<th>ND-AD multipoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>D3S1263</td>
<td>0.62</td>
<td>1.39</td>
<td>-0.02</td>
<td>0.89</td>
<td>0.90</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>D3S2385</td>
<td>0.29</td>
<td>1.67</td>
<td>0.04</td>
<td>1.23</td>
<td>0.75</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>D3S2386</td>
<td>0.50</td>
<td>1.28</td>
<td>0.27</td>
<td>1.12</td>
<td>0.51</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>D3S1293</td>
<td>1.40</td>
<td>1.55</td>
<td>0.35</td>
<td>1.46</td>
<td>0.56</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>D3S1558</td>
<td>0.75</td>
<td>1.13</td>
<td>0.36</td>
<td>1.26</td>
<td>2.37</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>D3S3646</td>
<td>1.23</td>
<td>1.17</td>
<td>0.42</td>
<td>1.05</td>
<td>2.73</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td>D3S1558</td>
<td>0.33</td>
<td>1.19</td>
<td>0.15</td>
<td>1.03</td>
<td>2.88</td>
<td>1.08</td>
</tr>
<tr>
<td>5</td>
<td>D5S2501</td>
<td>1.52</td>
<td>0.70</td>
<td>1.93</td>
<td>1.55</td>
<td>1.93</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>D5S404</td>
<td>1.56</td>
<td>1.07</td>
<td>0.89</td>
<td>1.40</td>
<td>1.99</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>D5S410</td>
<td>1.54</td>
<td>0.35</td>
<td>1.88</td>
<td>1.46</td>
<td>1.60</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>D5S2050</td>
<td>1.14</td>
<td>0.32</td>
<td>1.71</td>
<td>1.16</td>
<td>1.41</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>D5S408</td>
<td>0.11</td>
<td>0.24</td>
<td>0.20</td>
<td>0.86</td>
<td>4.06</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>D5S2006</td>
<td>0.39</td>
<td>0.24</td>
<td>0.70</td>
<td>0.86</td>
<td>4.21</td>
<td>0.31</td>
</tr>
<tr>
<td>6</td>
<td>D6S458</td>
<td>1.59</td>
<td>0.42</td>
<td>0.45</td>
<td>0.36</td>
<td>1.18</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>D7S2252</td>
<td>-0.06</td>
<td>0.07</td>
<td>-0.05</td>
<td>0.04</td>
<td>4.13</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>D7S2250</td>
<td>0.07</td>
<td>0.12</td>
<td>-0.04</td>
<td>-0.08</td>
<td>4.61</td>
<td>0.36</td>
</tr>
<tr>
<td>8</td>
<td>D8S512</td>
<td>0.96</td>
<td>0.01</td>
<td>1.60</td>
<td>0.27</td>
<td>0.34</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>D9S1828</td>
<td>1.25</td>
<td>0.69</td>
<td>0.55</td>
<td>1.31</td>
<td>0.22</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>D9S1776</td>
<td>0.55</td>
<td>1.06</td>
<td>0.82</td>
<td>1.46</td>
<td>0.51</td>
<td>1.08</td>
</tr>
<tr>
<td></td>
<td>D9S934</td>
<td>0.60</td>
<td>0.75</td>
<td>0.25</td>
<td>1.33</td>
<td>0.71</td>
<td>1.23</td>
</tr>
<tr>
<td>11</td>
<td>D11S2368</td>
<td>0.12</td>
<td>0.00</td>
<td>1.56</td>
<td>0.40</td>
<td>0.14</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>D11S1785</td>
<td>0.57</td>
<td>0.24</td>
<td>1.62</td>
<td>1.76</td>
<td>0.62</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>D11S903</td>
<td>0.32</td>
<td>0.23</td>
<td>2.05</td>
<td>1.84</td>
<td>0.55</td>
<td>0.05</td>
</tr>
<tr>
<td>12</td>
<td>D12S83</td>
<td>1.62</td>
<td>1.53</td>
<td>0.65</td>
<td>1.74</td>
<td>3.58</td>
<td>0.70</td>
</tr>
<tr>
<td>13</td>
<td>D13S219</td>
<td>1.47</td>
<td>0.75</td>
<td>0.36</td>
<td>0.98</td>
<td>0.98</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>D13S219</td>
<td>0.25</td>
<td>0.89</td>
<td>0.07</td>
<td>0.81</td>
<td>1.17</td>
<td>1.15</td>
</tr>
<tr>
<td></td>
<td>D13S325</td>
<td>0.98</td>
<td>0.80</td>
<td>0.05</td>
<td>0.36</td>
<td>1.01</td>
<td>1.80</td>
</tr>
<tr>
<td></td>
<td>D13S326</td>
<td>0.17</td>
<td>0.47</td>
<td>0.16</td>
<td>0.18</td>
<td>0.92</td>
<td>1.61</td>
</tr>
<tr>
<td></td>
<td>D13S272</td>
<td>0.14</td>
<td>0.32</td>
<td>0.19</td>
<td>0.24</td>
<td>0.84</td>
<td>1.14</td>
</tr>
<tr>
<td>18</td>
<td>D18S63</td>
<td>0.58</td>
<td>0.32</td>
<td>0.01</td>
<td>0.22</td>
<td>4.48</td>
<td>0.11</td>
</tr>
<tr>
<td>21</td>
<td>D21S1905</td>
<td>1.16</td>
<td>0.62</td>
<td>1.84</td>
<td>2.75</td>
<td>1.64</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>D21S1902</td>
<td>0.26</td>
<td>1.12</td>
<td>0.46</td>
<td>3.02</td>
<td>1.65</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>D21S272</td>
<td>1.48</td>
<td>1.23</td>
<td>1.02</td>
<td>2.92</td>
<td>1.33</td>
<td>0.86</td>
</tr>
<tr>
<td>22</td>
<td>D22S1177</td>
<td>2.11</td>
<td>0.98</td>
<td>0.66</td>
<td>0.47</td>
<td>0.71</td>
<td>0.08</td>
</tr>
</tbody>
</table>

* Shown above are the genomic locations with one or more results in the top 1% of scores for any of the phenotypes shown (multipoint for all four phenotypes plus singlepoint for FTND and ND). The shaded boxes indicate the top results. Also shown are the results for the other analyses at the same location in order to indicate convergence across different analytic methods and phenotypes.
are presented elsewhere (Prescott et al., 2006) but the striking chr4 signal for AlcSx (multipoint LOD = 4.59, \( p = 0.0000021 \)) is depicted in Fig. 1.

For the bivariate analysis of the continuous FTND and AlcSx scores, multipoint LOD scores substantially exceeded four in two genomic regions—an 11 cM region on chr7 (D7S2252–D7S691, maximum 5.06, empirical \( p = 0.0006 \)) and an 8 cM region on chr18 flanking D18S63 (maximum 4.48, empirical \( p = 0.0007 \)). Both regions did not meet the study-wide significance level of 0.00001. Inspection of the standardized coefficients from these analyses suggested that the linkage evidence at each of these loci came mostly from the FTND score (\( \sim 2/3 \)) but with a substantial contribution from AlcSx (\( \sim 1/3 \)). For the bivariate analyses of discrete ND and AD phenotypes, all multipoint LOD scores were <2.

Variance component analyses for quantitative phenotypes can yield biased results. First, although there was no evidence of inflation in estimates of linkage in univariate FTND analyses (median of all genome LOD scores was 0.0), there was evidence of upward bias in the bivariate FTND–AlcSx analyses (median of all genome LOD scores was 0.51). Therefore, the empirical \( p \)-values rather than the LOD scores should be used to describe our findings. Second, variance components analyses are susceptible to influential single observations. Occasionally, a few sibling pair can contribute the majority of linkage signal, which raises the concern of artifact (Mitchell et al., 1999). To investigate the degree to which this influenced our findings, we conducted a sensitivity analysis by analyzing the contribution of each family to the overall LOD scores for the bivariate FTND–AlcSx analyses. Briefly, no single family accounted for more than 15% of the total \( \chi^2 \) at any position suggesting that the bivariate FTND–AlcSx variance components findings are not due to an overly influential sibship.

Notably, the chr4 region that was positive in our prior report for AlcSx (Prescott et al., 2006) showed no substantial signal for nicotine-related phenotypes (LODs for ND and FTND were both negative) or when analyzed jointly with nicotine-related phenotypes (ND–AD LOD = 0.05, FTND–AlcSx LOD = 0.252).

To facilitate the incorporation of these findings in meta-analyses, genomewide linkage results are available in the online version of this report.

4. Discussion

In this report, we conducted secondary data analyses for lifetime nicotine-related phenotypes using genomewide linkage scan data (1020 microsatellite markers) from 626 all possible sibling pairs ascertained on the basis of lifetime alcohol dependence. There were four sets of genomewide linkage analyses—univariate analyses of continuous FTND scores, univariate analyses of a lifetime history of ND, bivariate analyses of FTND–AlcSx, and bivariate analyses of ND–AD. We wish to highlight four genomic regions.

First, the most striking finding was from the bivariate FTND–AlcSx linkage analyses with an empirical \( p \)-value of 0.0006 on chr7 from D7S2252 to D7S691. Although the most significant findings did not meet a conservative estimate of study-wide significance (and thus could be due to chance), sensitivity analyses did not suggest an artifactual result. We note that this finding is congruent with those from Li et al. for average cigarettes per day in the Framingham Heart Study population (Li et al., 2003b) and near those of Bergen et al. for smoking initiation in the COGA sample (Bergen et al., 1999). Bioinformatic investigation of the known and RefSeq genes in the 10 mb region from D7S2252 to D7S691 revealed no genes with clear first-order connections to either nicotine or alcohol dependence. Identification or exclusion of any etiologically-relevant locus in this region will require more exhaustive “fine-mapping” genotyping in a case–control sample.

Second, there was some evidence for linkage from bivariate FTND–AlcSx analyses on chr18 (empirical \( p = 0.0007 \)). The chr18 peak is in the vicinity of findings from our prior linkage study of nicotine dependence (Straub et al., 1999; Sullivan et al., 2004). Third, there may be an additional peak for the bivariate FTND–AlcSx bivariate phenotype on distal chr5 although this region is not near any prior finding. Although other regions had elevated LOD scores in these bivariate analyses, we suggest cautious interpretation of isolated results. Note that we provide our results as a Supplemental table to facilitate inclusion in meta-analyses.

Fourth, the ND multipoint analyses suggested the importance of a region on chr21 with a peak NPL LOD of 3.02. This region has not been implicated in any prior study.

The other sets of univariate and bivariate analyses showed more modest findings that were at the level that might be expected in a dense genome scan from the play of chance (Lander and Kruglyak, 1995). Moreover, there was no striking overlap with those from other genomewide surveys for nicotine-related phenotypes. Therefore, these additional linkage peaks are likely to contain false positive findings along with an unknown proportion of true positive findings. Formal meta-analyses of all extant data would be useful in attempting to delineate these possibilities (Levinson et al., 2003).

There was no evidence that the linkage signal from univariate analysis of AlcSx on chr4 (peak multipoint LOD 4.59 at D4S1611, \( p = 0.0000021 \) (Prescott et al., 2006) was shared with nicotine-related phenotypes in univariate or bivariate analyses suggesting that putative genetic variation in this region is principally alcohol-related rather than influencing a broader addiction phenotype.

These findings must be considered in light of a number of potential limitations. First, there was no built-in replication sample and a clear lesson from the history of complex trait genetics is that replication is essential. Second, this sample was ascertained on the basis of alcohol dependence. Strictly speaking, positive linkage findings are for increased allele sharing identical by descent for sibling similarity for nicotine-related phenotypes in the context of considerable ethanol consumption and alcohol dependence. The relevance of these findings to nicotine dependence more generally is unknown and unclear. An additional consequence is that there was an attenuation of the correlation between alcohol and nicotine use. Third, as explicated in Section 2, variance components linkage analyses required assumptions of trait normality and ascertainment correction required assump-
tions about population parameters. These are known imprecisely for our sample, and our assumptions may have influenced our results.

Conflict of interest

Dr Sullivan has received research support from Eli Lilly for work on schizophrenia. Dr Webb is currently partially supported by an industry research grant from Philip Morris USA, a major tobacco company; however, he was not supported by PMUSA while conducting the work contained in the manuscript. Dr Webb declares that his relationship with PMUSA has no influence on the choice of analyses or interpretation of the results and that no collaborators at PMUSA were a part of this research. No other authors have received funding from the tobacco industry. All other authors declare that they have no conflicts of interest.

Acknowledgements

We are grateful to the study participants and their families for their time and effort. We thank Dr Hermine Maes for statistical consultation. Margaret Devitt, Lisa Halberstadt, John Myers, Victor Robinson, and Dawn Thiselton made critical contributions to the project. We thank Ruth Barrington and the Irish Health Research Board for their support of the project.

Contributors: Dr Sullivan supervised fieldwork, designed the analytic approach, and wrote the manuscript. Dr Ku performed primary linkage analysis and assisted in interpreting the results. Dr Webb performed quality control procedures for genotypes, performed primary linkage analysis and assisted in interpreting the results. Dr Neale assisted in preparing the data analysis plan. Dr Furfur conducted literature reviews used in this paper. Ms Vittum assisted in laboratory procedures. Dr Walsh oversaw the Irish fieldwork. Dr Patterson ascertained subjects and supervised fieldwork. Dr Riley directed the laboratory component of this project. Dr Prescott was the PI for genotyping grant and was involved in study design, data collection & analysis. Dr Kendler obtained funding, directed fieldwork, and consulted on statistical analyses. All authors contributed to and have approved the final manuscript.

Role of funding source: Funding for this study was provided by NIAAA grant AA011408, NIDA grant DA011287, NIMH grant MH065322, and NCI grant CA085739. The NAAA, NIDA, NIMH, and NCI had no further role: in study design; in obtaining funding, directed fieldwork, and consulted on statistical analyses. All authors contributed to and have approved the final manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.drugalcdep.2007.09.015.

References


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. 17 (Suppl. 1), S139–S144.


