The role of conduct disorder in the relationship between alcohol, nicotine and cannabis use disorders

J. D. Grant1*, M. T. Lynskey2, P. A. F. Madden3, E. C. Nelson1, L. R. Few1, K. K. Bucholz2, D. J. Statham3, N. G. Martin4, A. C. Heath1 and A. Agrawal1

1 Department of Psychiatry, Washington University School of Medicine, St Louis, MO, USA
2 Institute of Psychiatry, Psychology & Neuroscience, Addictions Department, King’s College London, London, UK
3 University of the Sunshine Coast, Queensland, Australia
4 QIMR Berghofer Medical Research Institute, Queensland, Australia

Background. Genetic influences contribute significantly to co-morbidity between conduct disorder and substance use disorders. Estimating the extent of overlap can assist in the development of phenotypes for genomic analyses.

Method. Multivariate quantitative genetic analyses were conducted using data from 9577 individuals, including 3982 complete twin pairs and 1613 individuals whose co-twin was not interviewed (aged 24–37 years) from two Australian twin samples. Analyses examined the genetic correlation between alcohol dependence, nicotine dependence and cannabis abuse/dependence and the extent to which the correlations were attributable to genetic influences shared with conduct disorder.

Results. Additive genetic ($\alpha^2 = 0.48–0.65$) and non-shared environmental factors explained variance in substance use disorders. Familial effects on conduct disorder were due to additive genetic ($\alpha^2 = 0.39$) and shared environmental ($c^2 = 0.15$) factors. All substance use disorders were influenced by shared genetic factors ($r_g = 0.38–0.56$), with all genetic overlap between substances attributable to genetic influences shared with conduct disorder. Genes influencing individual substance use disorders were also significant, explaining 40–73% of the genetic variance per substance.

Conclusions. Among substance users in this sample, the well-documented clinical co-morbidity between conduct disorder and substance use disorders is primarily attributable to shared genetic liability. Interventions targeted at generally reducing deviant behaviors may address the risk posed by this shared genetic liability. However, there is also evidence for genetic and environmental influences specific to each substance. The identification of these substance-specific risk factors (as well as potential protective factors) is critical to the future development of targeted treatment protocols.

Received 4 October 2014; Revised 16 July 2015; Accepted 17 July 2015

Key words: Alcohol, cannabis, conduct disorder, genetic overlap, nicotine, substance use disorders, twins.
behavioral disinhibition (Krueger et al. 2002; Hicks et al. 2011).

Based on this extant literature, it remains unanswered the extent to which the genetic covariation across the three most common forms of substance use disorder (i.e. alcohol, nicotine and cannabis) is attributable to genes shared with conduct disorder. Two outcomes might be expected. One possibility is that all the genetic influences on the three substances are overlapping with each other and with conduct disorder, with limited evidence for substance-specific or conduct disorder-specific genetic factors. This is unlikely given prior evidence for significant substance-specific genetic influences on substance dependence symptoms and on conduct disorder (Kendler et al. 2003b; Button et al. 2006, 2007). Alternatively, a moderate to high genetic correlation across the three substances might be expected, with evidence for substance-specific genes. Whether the entire proportion of this genetic correlation is due to genes shared with conduct disorder or whether some of this genetic correlation is independent of it (i.e. shared across the substances but not with conduct disorder) has not been examined. This study of the extent to which the genetics of conduct disorder contribute to genetic liability to substance use disorders is critical to gene-finding efforts where increasing sample sizes, often capitalizing on differing definitions of phenotypes with overlapping genetic underpinnings, is a necessity (Hicks et al. 2011). Equally important is the estimation of substance-specific genetic influences. For instance, Kendler et al. (2007) noted that substance-specific genetic variance ranged from 3% for cocaine to 91% for caffeine abuse/dependence symptoms.

To examine this question, we utilized data from 9577 adult male and female twins aged 24–37 years drawn from two samples derived from the Australian Twin Registry. Specifically we fit a quadrivariate model that examined the extent to which genetic and environmental influences were shared across conduct disorder and DSM-IV diagnoses of nicotine dependence (ND) and alcohol dependence (AD), and a modified cannabis abuse/dependence (CAD) diagnosis based on a subset of DSM-IV criteria.

Method and materials

Sample

Twin pairs from two cohorts of the volunteer Australian Twin Registry were included in the present analyses (Maciejevski et al. 2014). As described in detail elsewhere, telephone diagnostic interviews were completed for one cohort in 1996–2000 (see Knopik et al. 2004 for details) and for the other in 2005–2009 (Lynskey et al. 2012). The samples had comparable ages at the time of interview (mean = 29.94 and mean = 31.85, respectively). After removing individuals of unknown zygosity and those with no data for any of the four outcomes, the sample for the present analyses included 9577 individuals (6255 of 6257 interviewed individuals in the first cohort; 3322 of 3348 interviewed individuals in the second cohort). The 3982 complete twin pairs included: 1096 monozygotic (MZ, or identical) female twin pairs, 665 MZ male pairs, 817 dizygotic (DZ, or fraternal) female pairs, 513 DZ male pairs, and 891 male–female twin pairs; data from 1613 individuals without co-twin data, which contribute to estimates of prevalence (but not variance decomposition), were also included. Zygosity was determined using standard items about physical similarity which have been shown in a subset of the present respondents to have 95% agreement with DNA analysis (Medland et al. 2009). Verbal informed consent was obtained from all participants, and the procedures were approved by the Human Research Protection Office at Washington University and the Human Research Ethics Committee at the Queensland Institute of Medical Research.

Measures

Both cohorts completed a diagnostic telephone interview based on the Semi-structured Assessment of the Genetics of Alcoholism [Australian version, SSAGA-OZ (Bucholz et al. 1994; Heath et al. 1997)], which included DSM-IV (APA, 1994) assessments of conduct disorder, and ND, AD, and CAD.

Conduct disorder

Respondents were asked about 15 behaviors associated with DSM-IV conduct disorder (e.g. bullying, threatening, stealing, rule-breaking). Individuals who indicated that ≥3 behaviors (without requiring impairment) occurred within a 12-month period prior to age 18 years were coded as having conduct disorder. None of the items used in the conduct disorder section was related to substance use or misuse.

Nicotine

Respondents who had smoked at least 100 cigarettes lifetime or who had smoked 21–99 cigarettes and had smoked for at least ≥2 days a week for ≥3 weeks, completed a detailed DSM-IV nicotine dependence (ND) assessment. Individuals who endorsed ≥3 dependence symptoms (out of seven) within a 12-month period were coded as having ND. Because individuals who do not initiate using a substance (i.e. never user) should be considered
unmeasured for their liability to developing substance use disorders (Neale et al. 2006), individuals who had never tried cigarettes were coded as missing for ND ($n = 1001$). Individuals with <3 dependence symptoms, those whose symptoms did not cluster within a 12-month period, and individuals who had tried cigarettes but had only experimented with nicotine (i.e. had smoked <100 cigarettes lifetime and had not smoked for at least $\geq 2$ days per week, $n = 3900$) were coded as unaffected for ND.

### Alcohol

Respondents were asked about their lifetime alcohol consumption. Individuals who had consumed at least five drinks in a 24-h period at least once completed a detailed DSM-IV alcohol dependence (AD) assessment and were included in analyses of AD. Those individuals who endorsed $\geq 3$ AD symptoms (out of seven) within a 12-month period were coded as having a history of AD. Non-drinkers ($n = 105$) and those who had not consumed at least five drinks in a 24-h period ($n = 908$) were coded as missing for AD analyses.

### Cannabis

Individuals who had used cannabis $\geq 11$ times or who had used cannabis at least monthly during the period of heaviest use were asked questions tapping DSM-IV cannabis abuse and dependence (CAD). The first cohort was only asked about two abuse symptoms (hazardous use and role interference) and four dependence symptoms (tolerance, using more than intended, continued use despite emotional/psychological problems caused/exacerbated by use, and repeated desire or inability to cut down on use). Although the second cohort was asked additional abuse and dependence questions, the present analyses used only the symptoms assessed in both cohorts. Respondents were identified as having CAD if they endorsed any abuse symptom or they endorsed $\geq 2$ dependence symptoms (symptoms were not required to have clustered within a 12-month period). Because individuals who do not initiate using a substance (i.e. never user) should be considered unmeasured for their liability to developing substance use disorders (Neale et al. 2006), individuals who had never tried cannabis were coded as missing for CAD ($n = 3499$). Individuals who completed the CAD section but did not meet the criteria for CAD listed above as well as those who had minimal cannabis exposure (i.e. had tried cannabis but used it <11 times and had not used it at least monthly during their period of heaviest use; $n = 3114$) were coded as unaffected for CAD.

### Statistical analyses

Quantitative genetic analyses were used to estimate the relative contributions of genetic and environmental factors to outcomes, and the overlap between diagnoses. As detailed elsewhere (Neale & Cardon, 1992), twin models allow for the estimation of additive genetic factors (A), non-additive genetic factors (D), shared environmental factors (C, environmental influences that make twins similar to each other), and non-shared environmental factors (E, environmental influences not shared by twins, as well as error variance). Genetic influences are indicated when the identical twin (MZ) correlation is greater than the fraternal twin (DZ) correlation. If all twin pair similarity were attributable to A, the MZ correlation would be about twice the DZ correlation, because MZ twins share all of their genes and DZ twins share half of their segregating genes (on average). D is indicated when the DZ correlation is less than half the MZ correlation (because MZ twins again share all non-additive influences but DZ twins only share one quarter of such influences, on average), and C is indicated when the DZ correlation is more than half the MZ correlation. Models containing only twins reared together cannot estimate both C and D simultaneously, and a decision as to which should be estimated is made based on correlations between identical and fraternal twins.

A quadrivariate Cholesky (lower triangular) decomposition was used to assess the degree of genetic and environmental influence on each measure, as well as the overlap across conduct disorder, ND, AD, and CAD. A full Cholesky model allows for influences on the first variable to also load on all subsequent variables, and for novel influences on each subsequent variable to load on all remaining variables. The E parameters in Fig. S1 [see also work by Neale and colleagues (Neale & Cardon, 1992; Neale, 2004)]. Conduct disorder was entered into the Cholesky model first. Given that individuals become at risk for developing dependence when they initiate substance use and that not all users develop dependence, ordering of the substance diagnosis variables in the Cholesky model (ND, AD, CAD) was determined by the most typical pattern of substance initiation. Of those who had tried all three substances, 60.0% tried cigarettes before both alcohol and cannabis, and an additional 18.3% initiated multiple substances in the same year, with cigarettes being one of those substances; 20.0% initiated alcohol first; 0.8% initiated cannabis first. Of individuals who had tried two of the three substances, 94.6% had used cigarettes and alcohol; 60.7% of those
who had used two substances initiated cigarettes before alcohol, 21.1% initiated alcohol before cigarettes, and 12.9% initiated cigarettes and alcohol within the same year. This ordering of variables in the Cholesky model (conduct disorder, ND, AD, CAD) was also supported by the mean ages of initiation (see Table 1).

The fit of the full Cholesky model was compared to a saturated model, which allowed us to test whether the assumptions of twin modeling were met. Sub-models dropping specific paths were used to test the significance of familial influences, and the overlap in influences across measures. To minimize the large number of tests that would need to be conducted to examine combinations of paths that could be dropped from the model, omnibus tests that constrained a family of paths were fit to the data in a manner that fit with the correlation patterns and with conceptually meaningful tests (single-parameter tests, presented in Supplementary Table S2, confirmed that no individual path in an omnibus test was statistically significant). The significance of paths was tested by comparing the fit of each sub-model to the preceding model and calculating the difference in −2 times the log-likelihood of the prior model and the sub-model, which is interpreted as a χ² test for the given degrees of freedom (df). Models were fitted using the statistical package Mx (Neale, 2004) using full information maximum-likelihood estimation with raw data. Analyzing raw data in Mx (Neale, 2004) allowed us to utilize data from conduct disorder for everyone, and substance dependence data from each individual for the substances he/she had used more times than the cut-off established. All quadrivariate models included gender and cohort as covariates.

Ethical standards

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

Results

Demographic and substance use characteristics are shown in Table 1. Alcohol initiation was endorsed by
Table 1. Demographic and substance use characteristics of 9577 members of Australian twin pairs aged 24–37 years

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>n</th>
<th>(%)</th>
<th>Years, mean (s.d.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age at interview</td>
<td>30.60 (2.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean age at conduct disorder onset</td>
<td>13.84 (2.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean age at cigarette initiation</td>
<td>14.06 (3.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean age at nicotine dependence onset</td>
<td>21.96 (4.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean age at alcohol initiation</td>
<td>15.83 (2.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean age at alcohol dependence onset</td>
<td>21.87 (3.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean age at cannabis initiation</td>
<td>18.50 (3.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean age at cannabis dependence onset*</td>
<td>21.36 (3.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>3972 (41.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monozygotic</td>
<td>4102 (42.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSM-IV conduct disorder</td>
<td>1100 (11.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nicotine use and dependence</td>
<td>8470 (88.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ever smoked a cigarette</td>
<td>2770 (28.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Of ever smokers</td>
<td>2770 (32.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol use and dependence</td>
<td>9456 (98.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ever had a full drink of alcohol</td>
<td>8544 (89.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSM-IV alcohol dependence</td>
<td>2219 (23.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population prevalence</td>
<td>2219 (26.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Of those having consumed at least 5 drinks in a 24-h period</td>
<td>1683 (17.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cannabis use and abuse/dependence</td>
<td>6017 (63.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ever used cannabis</td>
<td>1683 (28.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cannabis abuse/dependence (modified DSM-IV)</td>
<td>1683 (17.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population prevalence</td>
<td>1683 (17.7)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data from n = 9577 individuals; n = 9501 individuals for conduct disorder, n = 8470 for nicotine dependence, 8544 for alcohol dependence, and n = 6017 for cannabis abuse/dependence.

*Age at clustering for n = 311 individuals in the second cohort who met the modified DSM-IV cannabis abuse/dependence diagnosis (CAD). The n = 1683 with CAD included n = 1122 from the first cohort (age at diagnosis not assessed); of the n = 561 cohort 2 individuals with the modified CAD diagnosis, age was not assessed for the n = 177 with abuse only or the n = 73 with the modified ≥2 symptom dependence diagnosis but not the full DSM-IV dependence diagnosis with ≥3 symptoms and clustering.

almost 99% of the respondents, with 89% of the sample reporting having consumed at least five drinks in a 24-h period and 26% of those having consumed ≥5 drinks in a 24-h period reporting a lifetime history of AD. Over 88% of respondents had tried cigarettes, with 33% of initiators having a history of ND. Cannabis initiation was endorsed by 63% of respondents, with 28% of initiators having a history of CAD. Mean age of onset of conduct disorder was 13.84 years, which predated the average age of initiation for cigarettes (mean = 14.06), alcohol (mean = 15.83), and cannabis (mean = 18.50).

Because all univariate genetic analyses indicated that the parameter estimates could be equated for men and women [Δχ²(3) range: 2.30–6.38, all p > 0.05], univariate analyses were collapsed across gender and gender was included as a covariate (see Supplementary Table S1 for the correlations for all five zygosity groups). MZ and DZ twin-pair correlations, tested using SAS software version 9.2 (SAS Institute Inc., 2002–2008), indicated significant familiality for all measures (see Table 2). Univariate genetic models confirmed that additive genetic influences and non-shared environmental influences were significant for all measures, with shared environmental influences not reaching significance for any measure, but reaching p = 0.06 for conduct disorder (see Table 2).

A saturated model was used to estimate the tetrachoric correlations in Mx (Neale, 2004), and served as a baseline from which to test the assumptions of the genetic models (~2 times the log-likelihood = 31 827.135 with 72 estimated parameters). The initial genetic model, which included additive genetic, shared environmental, and non-shared environmental paths in a Cholesky parameterization, indicated that the Cholesky parameterization was reasonable for the current data (Δχ² = 19.023, Δdf = 30, p = 0.94 compared to the saturated model). Given that the correlation pattern suggested genetic influence on all measures, with smaller and less consistent evidence for shared environmental influences, we first tested the significance of the shared environmental paths on and from ND and AD. Eliminating all shared environmental influences on and from ND (c21, c22, c32, and c42 in Supplementary Fig. S1) and AD (c31, c33, and c43 in Supplementary Fig. S1) resulted in a more parsimonious model without a significant decrement in fit (total Δχ² = 5.948, Δdf = 7, p = 0.55; also see Supplementary Table S2 for the single-parameter tests of change in model-fit for this and all additional tests). A test of the overlap in shared environmental influences on conduct disorder and CAD indicated that the overlap was not significant (path c41 in Supplementary Fig. S1; Δχ² = 2.656, Δdf = 1, p = 0.99). A sub-model specifying that all genetic overlap between the substances could be explained by genetic influences shared with conduct disorder fit the data well (i.e. deleting paths a43, a42, and a32 from Supplementary Fig. S1; Δχ² = 0.664, Δdf = 3, p = 0.88; also see Supplementary Table S2). From this model, confidence intervals and further model testing confirmed that it was possible to delete the remaining CAD-specific shared environmental path (path c44 in Supplementary Fig. S1; Δχ² = 2.481, Δdf = 1, p = 0.12),...
but that all remaining genetic influences were statistically significant, as indicated by a significant decrement in fit when any path was deleted ($\Delta \chi^2$ range: 20.093–262.394 with 1 df, all $p < 0.001$). Although modest in magnitude, non-shared environmental influences overlapped across measures, with 5 of 6 paths being statistically significant ($\Delta \chi^2$ range: 5.846–28.275 with 1 df, all $p < 0.005$); the non-shared environmental path from conduct disorder to ND, which could be dropped without a significant decrement in fit ($\Delta \chi^2 = 0.169$ with 1 df, all $p = 0.68$), was retained in the final model but is indicated by the dashed line in Fig. 1. The significance of the three substance-to-substance non-shared environmental parameters (e32, e42, and e43 from Supplementary Fig. S1; also see Supplementary Table S2) indicates that individual-specific effects were shared across substances independent of the effects each substance shared with conduct disorder.

The standardized parameter estimates from the best-fitting model are shown in Fig. 1, with the proportions of variance presented in Table 3 and the genetic and environmental correlations in Table 4. As shown in Table 3, familial influences accounted for substantial variance in all measures, with genetic factors explaining 39% of the variance in conduct disorder, and 62%, 48%, and 65% of the variance in ND, AD, and CAD, respectively. Shared environmental influences accounted for an additional 15% of the variance in conduct disorder. Genetic correlations were significant and moderate to substantial (ranging from 0.38–0.77; above the diagonal in Table 4), especially between conduct disorder and ND and CAD, where over 50% of the genetic influences were overlapping. Importantly, the genetic correlation across substances was entirely attributable to genes shared with conduct disorder. Non-shared environmental influences were predominantly measure specific, as shown by the modest correlations between individual-specific environmental factors (range: 0.03–0.28; below the diagonal in Table 4).

### Discussion

In a large cohort of adult Australian twins, we show that the previously well-documented shared genetic etiology across ND, AD, and CAD was entirely attributable to genes contributed by a shared liability to conduct disorder. Our study extends previous research in that we: (a) examined conduct disorder, ND, AD, and CAD simultaneously in an adult sample [researchers have examined these substances together (Kendler et al. 2007, 2008), and have examined conduct disorder

### Table 2. Twin-pair correlations and final univariate genetic models (and 95% confidence intervals) for conduct disorder, nicotine dependence, alcohol dependence, and cannabis abuse/dependence

<table>
<thead>
<tr>
<th></th>
<th>$r_{MZ}$</th>
<th>$r_{DZ}$</th>
<th>A</th>
<th>C</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conduct disorder</td>
<td>0.55*</td>
<td>0.39*</td>
<td>0.32*</td>
<td>0.20*</td>
<td>0.48*</td>
</tr>
<tr>
<td>Nicotine dependence</td>
<td>0.63*</td>
<td>0.31*</td>
<td>0.62*</td>
<td>–</td>
<td>0.38*</td>
</tr>
<tr>
<td>Alcohol dependence</td>
<td>0.51*</td>
<td>0.22*</td>
<td>0.47*</td>
<td>–</td>
<td>0.53*</td>
</tr>
<tr>
<td>Cannabis abuse/dependence</td>
<td>0.63*</td>
<td>0.40*</td>
<td>0.63*</td>
<td>–</td>
<td>0.37*</td>
</tr>
</tbody>
</table>

A, Additive genetic influences; C, shared environmental influences; E, non-shared environmental influences.

Data from $n = 9501$ individuals for conduct disorder, $n = 8470$ for nicotine dependence, $n = 8544$ for alcohol dependence, and $n = 6017$ for cannabis abuse/dependence.

*Significant at $p < 0.05$;

* $\Delta \chi^2 = 3.59$ when removed from the model, $p = 0.06$.

### Table 3. Standardized proportions of variance (95% confidence intervals) attributable to additive genetic, shared environmental, and non-shared environmental influences in the best-fitting quadrivariate model examining the genetic and environmental contributions to the co-morbidity between conduct disorder, nicotine dependence, alcohol dependence, and cannabis abuse/dependence in 9577 Australian twins

<table>
<thead>
<tr>
<th></th>
<th>Additive genetic</th>
<th>Shared environment</th>
<th>Non-shared environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conduct disorder</td>
<td>0.39 (0.31–0.48)</td>
<td>0.15 (0.07–0.22)</td>
<td>0.46 (0.39–0.53)</td>
</tr>
<tr>
<td>Nicotine dependence</td>
<td>0.62 (0.56–0.67)</td>
<td>–</td>
<td>0.38 (0.34–0.44)</td>
</tr>
<tr>
<td>Alcohol dependence</td>
<td>0.48 (0.41–0.55)</td>
<td>–</td>
<td>0.52 (0.45–0.59)</td>
</tr>
<tr>
<td>Cannabis abuse/dependence</td>
<td>0.65 (0.58–0.71)</td>
<td>–</td>
<td>0.35 (0.29–0.43)</td>
</tr>
</tbody>
</table>

All variance components significant at $p < 0.05$. 

However, Table 4. As shown in Table 3, familial influences accounted for substantial variance in all measures, with genetic factors explaining 39% of the variance in conduct disorder, and 62%, 48%, and 65% of the variance in ND, AD, and CAD, respectively. Shared environmental influences accounted for an additional 15% of the variance in conduct disorder. Genetic correlations were significant and moderate to substantial (ranging from 0.38–0.77; above the diagonal in Table 4), especially between conduct disorder and ND and CAD, where over 50% of the genetic influences were overlapping. Importantly, the genetic correlation across substances was entirely attributable to genes shared with conduct disorder. Non-shared environmental influences were predominantly measure specific, as shown by the modest correlations between individual-specific environmental factors (range: 0.03–0.28; below the diagonal in Table 4).
with AD and general substance use disorders (Krueger et al. 2002; Kendler et al. 2003b), but not all in the same analysis; (b) extended Hicks and colleagues analysis of adolescents (aged 16–19 years; Hicks et al. 2011), which could yield different patterns of overlap; (c) set to missing the data from individuals who had not initiated a substance and hence had unknown genetic liability. Our results are consistent with previous reports (Fu et al. 2002; Kendler et al. 2003b; Button et al. 2006; Hicks et al. 2011) showing that alcohol and drug dependence tend to share an appreciable proportion of their genetic liability with conduct and antisocial personality disorder, and further demonstrating that similar effects extend to ND and to substance use disorders contingent on initiation. The overall structure of genetic overlap (see Fig. 1) is also broadly consistent with numerous twin studies of adolescent and adult populations indexing a general liability factor underlying a host of externalizing disorders, including conduct disorder, substance use disorders and also impulsivity and non-substance-related disinhibition (Krueger et al. 2002; Hicks et al. 2011).

Such results have key implications for gene-finding efforts such as genome-wide association studies (GWAS). In particular, they imply that, as least when examining initiators, phenotypes that aggregate across various substance use disorders (e.g. Wetherill et al. 2015) or examine genetic variants that influence multiple substances (McGue et al. 2013; Vrieze et al. 2014) may effectively be capturing genetic variation shared with liability to conduct disorder. Thus, approaches of aggregating across substances or of using conduct disorder and other aspects of behavioral disinhibition as proxies may be particularly useful in increasing power to detect genetic variants shared across substances, particularly in adolescent cohorts where subjects may not have fully passed through the period of greatest risk for onset of substance use disorders (Iacono et al. 1999; Rose et al. 2004; Button et al. 2006; Hicks et al. 2011). However, at least one study reports that polygenic scores created for individual substance-related measures and behavioral disinhibition are only modestly ($r = 0.03–0.07$) correlated with each other (Vrieze et al. 2013).

Developmental changes in the degree to which genetic factors contribute to substance involvement itself also should be considered, although there is disagreement about the pattern of change. For instance, research by Vrieze et al. (2012) suggests that the genetic contribution to the overlap across substances may decline from adolescence into adulthood, while research by Kendler and colleagues suggests that the genetic contribution to overlap across substance use remains steady, even increases modestly, across that time span (Kendler et al. 2008). There is also evidence that the correlation between conduct disorder and substance involvement is more attributable to shared environmental factors during adolescence (Rose et al. 2004). Thus, our finding of strong genetic overlap may be specific to the age of our cohort (young adults) and may not generalize to other developmental settings.

Despite genetic overlap, the statistically significant role of substance-specific genetic influences cannot be ignored. In fact, a significant proportion of the genetic variance in ND (40%), AD (73%), and CAD (47%) was substance-specific. The finding of substance-specific genetic influences corroborates previous research, although the relatively higher proportion of specific genetic influences on AD v. ND is in contrast to one prior study that found ND to be most significantly influenced by specific genetic factors (Kendler et al. 2007). The magnitude of these specific genetic factors underscores that when studying the genomic underpinnings of individual substance use disorders, variants are not exclusively expected to reflect a general liability to conduct disorder or externalizing problems. Most notably, for AD and ND, the most robustly validated single nucleotide polymorphisms identified via GWAS have been substance specific. For alcohol,

### Table 4. Genetic (above diagonal) and non-shared environmental correlations* between conduct disorder, nicotine dependence, alcohol dependence, and cannabis abuse/dependence from the best-fitting model examining the genetic and environmental contributions to the co-morbidity between conduct disorder, alcohol dependence, nicotine dependence and cannabis abuse/dependence in 9577 Australian twins

<table>
<thead>
<tr>
<th></th>
<th>Conduct disorder</th>
<th>Nicotine dependence</th>
<th>Alcohol dependence</th>
<th>Cannabis abuse/dependence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conduct disorder</td>
<td>–</td>
<td>0.77* (0.69–0.87)</td>
<td>0.52* (0.44–0.61)</td>
<td>0.72* (0.64–0.81)</td>
</tr>
<tr>
<td>Nicotine dependence</td>
<td>0.03 (–0.10 to 0.14)</td>
<td>–</td>
<td>0.40* (0.33–0.48)</td>
<td>0.56* (0.48–0.64)</td>
</tr>
<tr>
<td>Alcohol dependence</td>
<td>0.28* (0.19 to 0.38)</td>
<td>0.27* (0.17–0.35)</td>
<td>–</td>
<td>0.38* (0.31–0.45)</td>
</tr>
<tr>
<td>Cannabis abuse/dependence</td>
<td>0.17* (0.05 to 0.31)</td>
<td>0.18* (0.06–0.31)</td>
<td>0.28* (0.17–0.39)</td>
<td>–</td>
</tr>
</tbody>
</table>

*Indicates $p < 0.05$.

* Genetic correlations are above the diagonal; non-shared environmental correlations are below the diagonal.
rs1229984 in the alcohol dehydrogenase (ADH1B) gene is associated with accelerated conversion of ethanol to acetaldehyde, thus protecting against alcoholism via a specific metabolic pathway (Thomasson et al. 1993). This functional polymorphism has been identified at genome-wide significant levels in two independent studies (Bierut et al. 2012; Gelernter et al. 2014). For tobacco smoking, a similar metabolic variant in the cytochrome P450 A6 (CYP2A6) gene has been identified in one large meta-analysis (Thorgeirsson et al. 2010). However, the most widely replicated findings are for tobacco smoking and rs16969968 in the cholinergic nicotinic receptor a5 subunit (CHRNA5) (Liu et al. 2010; Thorgeirsson et al. 2010; Tobacco and Genetics Consortium, 2010) and other variants in the family of genes encoding nicotinic receptor subunits, which have been repeatedly identified at genome-wide significant levels. It is worth noting that the large GWAS meta-analyses of smoking that identified rs16969968 relied on nicotine consumption measures such as cigarettes per day (not dependence) and primarily included samples in which the co-morbid effects of conduct disorder are likely to be minimal. These substance-specific findings have also arisen in a sample with a high degree of drug-related co-morbidity (e.g. Rice et al. 2012) but in that study, when studying one substance, diagnoses of other substance use disorders were accounted for, thus likely refining the ability to home in on substance-specific results. Future studies may also wish to control for conduct disorder to further enhance chances of identifying drug-specific variants.

Some limitations of the current study are worth noting. First, this is a sample of adult Caucasian Australians and findings may not generalize to other ethnic groups. Second, in order to test for the role of substance initiation on the genetic overlap with substance use disorder, we set to missing the substance-use initiation on the genetic overlap with substance use disorder; an alternate alternative would have been to restrict the analyses to those with more substantial exposure, as we did with alcohol. It is possible that the genetic overlap of conduct disorder with ND and CAD might change with a different substance use threshold (e.g. if conduct disorder is associated more with the transition to ‘regular use’ than with the transition to ‘dependence’). Fourth, the ordering of variables in our Cholesky model reflects the most commonly reported order of onset, but there are individuals who had a different order of onset. Although our variance components are applicable regardless of the order of onset, some of the parameter estimates in the Cholesky model would differ if a different variable order was used. Fifth, there is some evidence that the genetic structure of alcohol problems is not attributable to a single factor; in twin analyses examining the seven individual symptoms of DSM-IV AD, Kendler et al. (2012) found that the individual symptoms were best explained using three distinct genetic factors. However, analyzing the individual symptoms for AD, ND and CAD was beyond the scope of our analysis. Sixth, our CAD diagnoses included only a subset of DSM-IV symptoms (1+ of two abuse symptoms or 2+ of four dependence symptoms), which may have affected our prevalence and/or heritability. This was done because only this subset of items was included in the 1996–2000 cohort assessment. However, Lynskey and colleagues (2002) have reported that a diagnosis of CAD based on these six items reasonably approximates the population prevalence and does not attenuate estimates of heritability. Finally, all of our diagnoses are drawn from adult retrospective reports which may have been subject to recall bias.

In this era of big data, several efforts are underway to mega- and meta-analyze genome-wide data for substance use disorders. Our study indicates that looking across substances may provide a unique set of genetic variants that relate to a general vulnerability to conduct problems (at least among substance initiators), whereas co-varying for co-morbid conduct disorder may isolate substance specific findings. Importantly, our study documents the enormous value of twin studies in providing a paradigm for development of substance-related phenotypes for gene-identification efforts.

**Supplementary material**

For supplementary material accompanying this paper visit http://dx.doi.org/10.1017/S0033291715001518.

**Acknowledgements**

This study was supported by grant funding from the National Institute on Drug Abuse (NIDA) grants DA23668 (A.A.), DA32573 (A.A.), DA18267 (M.T.L.) and T32DA007313 (L.F.) as well as from the National Institute of Alcohol Abuse and Alcoholism (NIAAA) including AA11998, AA07728 and AA13221 to A.C.
H. Data collection for sample 2 was also facilitated through access to the Australian Twin Registry, a national resource supported by an Enabling Grant (ID 628911) from the National Health & Medical Research Council. Funding agencies did not participate in design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript. The authors thank the Australian Twin Registry and the twins themselves for participating in this research.

Declaration of Interest
None.

References


Neale MC (2004). *Statistical Modeling with Mx*. Department of Psychiatry, Box #980710: Richmond, VA 23298.


