

Autosomal linkage analysis for cannabis use behaviors in Australian adults

Arpana Agrawal^{a,*}, Katherine I. Morley^b, Narelle K. Hansell^b, Michele L. Pergadia^a,
Grant W. Montgomery^b, Dixie J. Statham^b, Richard D. Todd^{a,1}, Pamela A.F. Madden^{a,1},
Andrew C. Heath^{a,1}, John Whitfield^{b,1}, Nicholas G. Martin^{b,1}, Michael T. Lynskey^a

^a Washington University School of Medicine, Department of Psychiatry, 660 South Euclid Avenue, CB 8134, St. Louis, MO 63108, United States

^b Queensland Institute of Medical Research, Brisbane, Australia

Received 30 October 2007; received in revised form 15 April 2008; accepted 18 May 2008

Available online 7 July 2008

Abstract

Cannabis is the most commonly used illicit drug in developed and in developing nations. Twin studies have highlighted the role of genetic influences on early stages of cannabis use, such as a lifetime history of use, early-onset use and frequent use, however, we are not aware of any genomic studies that have examined these phenotypes. Using data on 2314 families consisting of 5600 adult Australian offspring and their parents, all of whom were scanned using 1399 unique autosomal markers, we conducted autosomal linkage analyses for lifetime history of cannabis initiation, early-onset cannabis use and frequency of use, using a variance components approach in the linkage package MERLIN. Suggestive evidence for linkage was found on chromosome 18 (LOD 2.14 for frequency of use, LOD 1.97 for initiation, at 90–97 cM) and also on chromosome 19 (LOD 1.92 for early-onset at 17 cM). These LOD scores did not meet genome-wide significance. Further replication of these linkage regions in other samples will be required, although these initial results suggest further targeted efforts on chromosome 18 may yield interesting candidate genes for early stages of cannabis-related behaviors.

© 2008 Elsevier Ireland Ltd. All rights reserved.

Keywords: Cannabis use; Linkage; Australian; MERLIN

1. Introduction

Cannabis remains the most widely used illicit substance in developed countries (Hall and Degenhardt, 2007). In the US, an estimated 14.6 million individuals aged 12 and older reported cannabis use in the past month, constituting over 72% of all illicit drug use across the US (SAMHSA, 2005). In Australia, 11.3 and 33.6% of individuals aged 14 and older report past 12 months and lifetime cannabis use, respectively (Australian Institute of Health and Welfare, 2005). Nearly 15% of European adults have used cannabis at least once in their lifetime (EMCDDA, 2006). In the United States, the estimated prevalence of cannabis use in the past 12 months ranges from 7 to 12% in those aged 12–25 years (SAMHSA, 2005). Lifetime cannabis use is reported by approximately 20% of the US adult population (Compton et al., 2004).

There is substantial evidence in favor of heritable factors influencing the liability to cannabis use behaviors. Across adult and adolescent twin samples, 17–67% of the variance in lifetime cannabis use has been attributed to genetic influences (Agrawal and Lynskey, 2006; Kendler and Prescott, 1998; Lynskey et al., 2003, 2006; McGue et al., 2000; Rhee et al., 2003). Other measures of cannabis use behavior, such as duration of cannabis use (heritability of 41–55%) (Lynskey et al., 2006), frequency of use (heritability of 55%) and early-onset use (heritability of 46–53%) are also genetically influenced (Lynskey et al., 2003, 2006).

Despite this evidence, linkage studies have rarely examined the genome for regions in linkage disequilibrium with cannabis use behaviors. Studies have isolated regions of the genome that possibly harbor genes associated with a vulnerability to poly-substance use disorders (Stallings et al., 2003, 2005; Uhl et al., 2001) and have also focused on linkage peaks for opiates (Gelernter et al., 2006) and cocaine (Gelernter et al., 2005), but rarely for cannabis. In fact, with the exception of two recent studies, first by Hopfer et al. (2006) which found evidence for linkage for cannabis dependence symptoms on chromosome 3

* Corresponding author. Tel.: +1 314 286 1778; fax: +1 314 286 2213.

E-mail address: arpana@matlock.wustl.edu (A. Agrawal).

¹ Principal investigator on a component project.

and 9 in an adolescent sample, and a second by Agrawal et al. (2008) reporting linkage to cannabis dependence symptoms from the Collaborative Study on the Genetics of Alcoholism on chromosome 14, we are not aware of other published reports of linkage signals for cannabis-related behaviors. Neither of the prior publications on cannabis-related phenotypes utilized data on early stages of cannabis-related behaviors, such as early-onset or frequency of use. This is an important shortcoming of the literature as numerous twin studies have demonstrated that early-onset cannabis use and frequent cannabis use are both associated with an increased likelihood of cannabis and other drug-related problems (Agrawal et al., 2008; Lessem et al., 2006; Lynskey et al., 2003, 2006). Therefore, in the present study, we performed genome-wide autosomal linkage analyses for (a) lifetime cannabis use, (b) early-onset cannabis use and (c) frequency of cannabis use in a sample of 5600 Australian adults.

2. Methods

2.1. Sample

The core of this community sample was drawn from the Australian Twin Registry (ATR) volunteer twin panel, which was formed in 1978–1979 and maintained by the Australian National Health and Medical Research Council. Data on these twin pairs and their family members, who were subsequently recruited (over a span of 25 years) for multiple projects draw from five telephone interview surveys:

- (1) An older volunteer cohort of male and female Australian twins born between 1940 and 1964 and interviewed via telephone in 1992–1994 (Heath et al., 2001; Heath and Martin, 1994). This sample consists of 2685 complete twin pairs and 519 twins whose co-twin did not participate (age range 27–89 years at interview).
- (2) Spouses or de facto partners of these older twins interviewed using a telephone diagnostic interview in 1996–1999 as part of a study on assortative mating for alcohol dependence (Agrawal et al., 2006; Grant et al., 2007), consisting of 3848 partners of twin participants (age range 19–90 years at interview).
- (3) A younger volunteer cohort of twin pairs (Twin89) born between 1964 and 1971 (Heath et al., 1997, 2001), interviewed via telephone in 1996–2000. The Twin89 sample (also called ‘Phase 2’ in Hansell et al., 2008) consists of 2765 twin pairs and 735 twins without co-twin participation (age range 24–36 years at interview).
- (4) Australian families with at least one current or former heavy smoker (NAG: defined as smoking at least a pack a day at peak lifetime use or 40 or more cigarettes in any 24 h period) (Madden, 2005) identified from (1), (2) or (3). See Saccone et al. (2007) for details—2540 individuals aged 21–82 years at interview had interview data.
- (5) Families ascertained using (1), (2) and (3) for a number of alcohol-related projects: (a) an unmatched case–control Candidate Gene Study; (b) a Big Sibships study, which included families with five or more full siblings, regardless of alcohol behaviors; (c) an Extreme Discordant and Concordant Sibships (EDAC) Study. These included 8072 individuals with interview data, aged 18–86 years at interview. Further details regarding data collection flow is also available in Hansell et al. (2008). Additional details on the twin, spouse, nicotine and alcohol samples are available in related publications (see for details, e.g. Heath et al., 1997; Heath and Martin, 1994; Knopik et al., 2004; Whitfield et al., 2004).

All data collection was approved by the Institutional Review Boards of the Washington University in St. Louis, USA and the Queensland Institute of Medical Research, Australia. The total dataset contains 13,200 individuals, however, genotypic data, which was collected only as a part of (4) and (5) was

available on 2352 families ($N=5704$ subjects, with 4032 twins, 527 parents, 377 full siblings and 2 half-siblings). The final dataset for our analyses only included 2314 families after excluding individuals with spurious or missing values for cannabis-related phenotypes (due to non-response) or missing genotypic data.

2.2. Genotypic data

A number of the studies outlined above included a microsatellite genome scan component, and thus genotypic data were available for the participants through their involvement in one or more of these genome scans. The scans included: (a) Gemini Genomics, 222 markers; (b) Sequana Therapeutics, 437 markers; (c) the Marshfield Clinic’s Mammalian Genotyping Service, 786 markers; (d) the University of Leiden, 430 markers; (e) the Australian Genome Research Facility, 394 markers; (f) the Finnish Genome Centre, 400 markers. Allowing for overlap between the various sets, there were in total 1461 unique markers, although no individual was genotyped for all of these. Note that additional markers in these datasets were available, but were excluded due to suspicions of unreliability given earlier comparisons of allele calls between the genome scans.

The genotypic data from all of these studies has been integrated to create a combined set of all available data for all individuals containing a total of 1461 unique markers after integration and cleaning (see Hansell et al., 2008 for details). Note that this is not a meta-analysis and that the samples are all drawn from the Australian Twin Registry.

Prior to integration, each genome scan was subjected to the following error checking process: (1) screening for Mendelian errors and exclusion of problematic markers; (2) confirmation of participants’ gender via X-chromosome markers; (3) checking of pedigree structures for each family using RELPAIR (Epstein et al., 2000) and GRR (Abecasis et al., 2001); (4) identification of unlikely recombination events using MERLIN (Abecasis et al., 2002). After identifying and correcting pedigree errors, the process was repeated. Data from all scans were then combined and checked again for problems with Mendelian inheritance and unlikely recombinants.

For the 5600 genotyped individuals analyzed here, there were a total of 1399 unique autosomal markers available. The median number of markers per individual was 392, but only 292 (5.2% of the genotyped sample) had less than 300 markers. Average marker heterozygosity was 75%.

2.3. Measures

2.3.1. Cannabis initiation. A dichotomous measure reflecting a lifetime history of cannabis use. If the participant had been interviewed more than once, any single report of a history of cannabis use was considered sufficient to code them as ‘affected’.

2.3.2. Early-onset use. In all interviews, participants who had used cannabis at least once in their lifetime were asked the age at which they had first used cannabis. A dichotomous age-of-onset measure was coded to reflect age of initiation prior to 18 years of age, with lifetime abstainers coded as unaffected (Lynskey et al., 2006). Note that if age of initiation was recorded more than once (for, e.g., a participant from Twin89 being re-interviewed in NAG), then, to reduce recall bias, the value that represented the smallest difference between age at that interview and reported age of initiation was used to code the dichotomous measure. For example, about 32% of participants with two interview time points, reported the same age of initiation at both time points, with 26 and 42% of the participants reporting an older age-of-onset and a younger age-of-onset, respectively, at the second interview. Therefore, analyses were re-done using the lowest reported age of initiation. However, as the distribution for age-of-onset was not normally distributed, we did not utilize the continuous age-of-onset measure for linkage analyses. Even after transformation, considerable skewness and kurtosis remained, and as variance components methods are sensitive to distributional assumptions, we did not use this measure.

2.3.3. Frequency of cannabis use. An ordinal measure of cannabis use was created from the participant’s report of how often in their entire lifetime they had

used cannabis. Participants could respond with any number, but were also given the option of responding “Too many times to count” which was assumed to be greater than 950. These responses were then collapsed into a five category ordinal scale, defined separately for males and females. For males these categories were 1–2 times, 3–9, 10–49, 50–499, 500+; for females they were 1 time, 2–4 times, 5–11, 12–99, 100+. The largest number reported by the participant across all assessments was used. Frequency was also coded in all participants, including lifetime abstainers, who were coded as 0. Different thresholds were used to accommodate the variations in frequency of use reported by men and women.

2.4. Linkage analysis

Variance components linkage analyses were conducted on all phenotypes in MERLIN (Abecasis et al., 2002). The variance components approach allows for the incorporation of covariates for both dichotomous and continuous measures, and our analyses included gender (not for frequency of use) and age at interview for each measure (note that age at interview for the same individual could vary across phenotypes depending on the assessment that was used to extract the phenotypic value). The variance components approach allows for the inclusion of both members of a pair of monozygotic twins, identified by a unique value (“Z”)—while data on only one randomly selected twin is utilized in the linkage analysis, inclusion of phenotypic data on both members allows for greater precision in estimation of heritability for the variance components linkage model. We therefore allowed for data on both members of the MZ pair to be included in the pedigree, although identifying them ensured that only one twin’s genotype would be utilized. While this inclusion of MZ twins improves precision of heritability estimates, it is susceptible to confounds with shared environmental factors, which are uniformly shared by MZ twins (Kendler and Prescott, 1998; McGue et al., 2000; Rhee et al., 2003).

2.5. Empirical *p*-value calculations

Using the gene-dropping algorithm in MERLIN, the 2314 pedigrees were simulated 950 times using random seeds. The algorithm assigns random genotypic values to founders while retaining observed segregation flow and phenotypic assignment. Running variance components linkage analyses on these simulated datasets allows investigation of the frequency with which a LOD score exceeding a given threshold is likely to be observed in a given linkage run purely by chance alone. Formally, the empirical *p*-value for this threshold is calculated as $\bar{p} = r + 1/951$, where *r* was the number of times a simulated LOD score exceeded an observed LOD score value (North, Curtis, and Sham 2003).

3. Results

3.1. Sample characteristics

Of the 5600 individuals (from 2314 families), about 36% of the women and 49% of the men reported lifetime cannabis use (Table 1). About 12 and 18% of the women and men had initiated their cannabis use before 18 years of age. Nearly 7 and 10% of the women and men respectively had used cannabis 100 or more times. A nearly equal proportion had used it 1–2 times only. Analyses were conducted on 5122 non-founders (including 4373 sibling pairs from multiple sibling pedigrees) and 478 founders (parents with both genotypic and phenotypic data). Hence, average family size, not including parents was 2.4 but

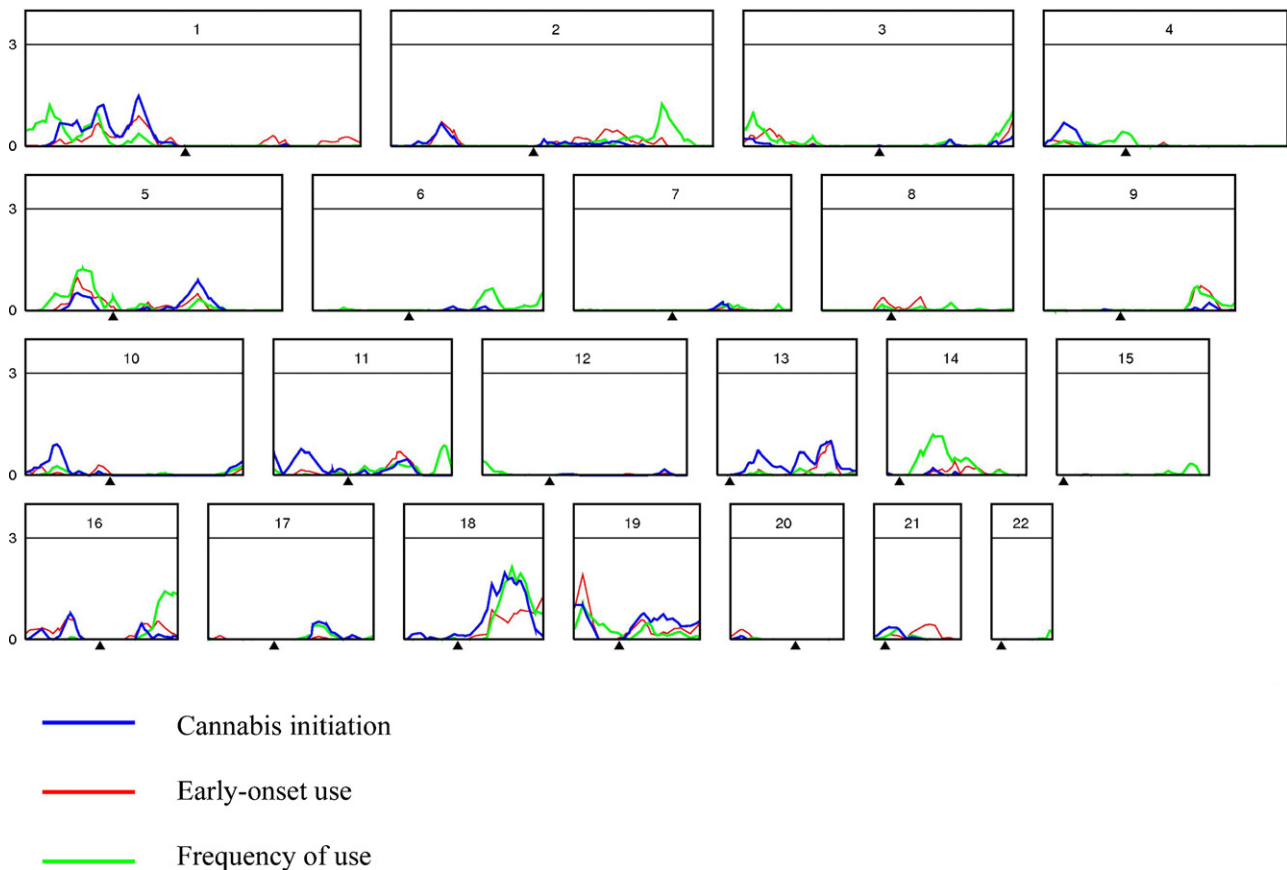


Fig. 1. Linkage peaks (LOD score of *y*-axis and DeCode centiMorgan position on *x*-axis) for cannabis initiation (blue), early-onset (red) and frequency of use (green) for 22 autosomal markers. The *x*-axis for each chromosome is scaled for chromosomal length (i.e. *x*-axis is longest for chromosome 1 and shortest for chromosome 22), and triangles on the *x*-axis represent the position of the centromere. LOD = 3.0 is denoted in each plot.

Table 1
Prevalence of cannabis use behaviors in the 5600 adult Australian participants

Phenotype	Males	Females
Cannabis initiation	49.2	35.7
Early-onset use		
Never used	50.8	64.3
18+	30.8	23.2
≤17	18.4	12.4
Frequency of use		
Never used	50.8	64.3
1–2/1 time (s)	9.6	6.4
3–9/2–4	10.4	7.9
10–49/5–11	10.3	6.7
50–499/12–99	8.8	7.4
500+/100+	9.5	7.2

Data on 2314 families. Average family size (including parents) is 4.22 with up to 12 members per family.

was 4.2 when including parents who may have been lacking phenotypic and/or genotypic data.

3.2. Variance components linkage analyses

The results from the variance components linkage analyses are graphically presented in Fig. 1. Elevations in then LOD score were noted on chromosome 1, 18 and 19. For cannabis initiation as well as for frequency of use, elevations on chromosome 18 (LOD score of 1.97 and 2.14, respectively) were noted (see Table 2). The maximum LOD for early-onset was 1.92 on chromosome 19. A modest peak (LOD 1.48 at 100 cM) was also noted for initiation on chromosome 1. These LOD scores did not meet genome-wide significance at $\alpha=0.05$. The empirical p -value of observing a LOD of 2.0 was 0.65.

4. Discussion

Applying a variance components linkage method to three related definitions of cannabis use, initiation, early-onset and frequency, we identified a region on chromosome 18 that possibly confers biological vulnerability to early stages of cannabis involvement. This region has not previously been identified in relationship to symptoms of cannabis dependence, although a region approximately 40 cM upstream to our maximum LOD score was identified by Gelernter et al. (2005) for linkage to moderate cocaine and opioid abuse, a linkage peak predominantly observed in African-American participants.

A number of interesting candidate genes with possible influences on onset and frequency of cannabis use are encompassed

by our linkage peak on chromosome 18. These include a family of 10 “serpins” or serine-peptidase inhibitor genes that cluster proximal to our linkage peak (*SERPINB2*, *SERPINB3*, *SERPINB4*, *SERPINB5*, *SERPINB7*, *SERPINB8*, *SERPINB10*, *SERPINB11*, *SERPINB12*, *SERPINB13*). Also in the region is *VPS4B* (vacuolar protein sorting 4), *DOK6* (docking protein 6) as well as two cadherin genes (*CDH7*, *CDH19*) which are involved in cell adhesion. These families of genes (serpins, cadherins and vacuolar protein sorting genes) have been identified in prior genome-wide association studies of nicotine dependence and of substance abuse vulnerability (Bierut et al., 2007; Liu et al., 2006), although not the candidates on chromosome 18, with the exception of *DOK6*, which was identified by Liu et al. (2006) in their genome-wide association study of substance abuse vulnerability.

Nusbaum et al. (2005) have characterized all 243 genes residing on chromosome 18. They noted that the serpin and cadherin families of genes are overrepresented on this chromosome, despite the fact that this chromosome has limited protein-coding genes and represents less than 3% of the coding content of the genome. The genes in our linkage peak are involved in a variety of regulatory activities. The serpins are involved in cancers, independently and via interactions with cadherins impacting metastasis of carcinomas. The cadherins are involved in cell adhesion which is vital to the development of neuronal connectivity and memory processes. Additionally, *VPS4B*, is part of the endosomal sorting complex and is involved in cellular trafficking.

There are certain caveats to our current study: first, our LOD scores did not satisfy empirical thresholds set for genome-wide significance. Second, we did not have sufficient power in our sample to exclude lifetime abstainers of cannabis use when conducting linkage analyses for early-onset or frequency of use—hence, linkage findings for these phenotypes may, to some extent, reflect linkage to cannabis initiation. The correlations between frequency of cannabis use and other externalizing behaviors, such as antisocial behavior, conduct disorder, and experimentation with other illicit drugs implies that some of our linkage peaks may generalize to these aspects of the externalizing spectrum as well. Third, we did not recover the LOD of 1.92 on chromosome 19 for age-of-onset when conducting the analyses with the lowest reported age of initiation—therefore, the potential that this finding is a false positive exists and this result should be interpreted with some caution. On the other hand, using a regression-based approach to re-do analyses with the frequency of use measure successfully identified the LOD on chromosome 18 (LOD = 1.6 with regression-based approach). Fourth, our sample consists of Australian adults of

Table 2
Results of multipoint linkage analyses for cannabis use behaviors in Australian adults

	Heritability (%)	Chromosome	Position (cM)	LOD score	Observed p -value
Cannabis	51.6	1	99.7	1.48	0.005
Initiation		18	90.4	1.97	0.0013
Early-onset	51.8	19	16.7	1.92	0.0015
Frequency	56.4	18	96.8	2.14	0.0009

predominantly Anglo-Celtic ancestry and our linkage finding may not be generalizable to other ethnic or racial groups. Fifth, all genotyped individuals did not have data on other aspects of cannabis-related behaviors (including abuse and dependence problems). Sixth, the lower prevalence of cannabis use in women of this combined sample may reflect cohort trends—some female participants were from the older Australian cohort where prevalence of cannabis use was only 19%. Finally, advances in genotyping technology might enhance the definition of linkage findings in this sample.

The key implication of our study is the importance of including cannabis use and quantitative measures of frequency in future linkage analyses. Although environmental factors are more critical during these early stages (accounting for nearly 50% of population variation in cannabis use), than in later stages, a major proportion of individual differences in cannabis use behaviors are also due to genetic predisposition, which may reflect a general vulnerability to polydrug use as well as other impulse-disinhibited behaviors. For instance, our linkage peak for cannabis initiation may encompass genetic influences on several stages of cannabis-related behavior as well as other correlated behaviors, such as polydrug involvement, all of which share genetic underpinnings (Kendler et al., 2003a,b). Our frequency measure may be associated with developing a vulnerability to problems with cannabis and LOD scores associated with this phenotype may reflect genetic influences on both early stages (as seen on chromosome 18) as well as later stages of cannabis involvement. We cannot, with these analyses, be certain however, that the LODs observed with the frequency of use measure are more specific to a later stage of cannabis involvement—this is particularly challenging also as a preponderance of twin models suggest considerable overlap between genetic factors on cannabis use and abuse/dependence (Agrawal et al., 2005; Young et al., 2006). Finally, early-onset cannabis use may also reflect genetic risk for polysubstance use, a propensity for other impulse-disinhibited behaviors as well as risk for abuse/dependence—multiple twin analyses have demonstrated a robust association between early age of initiation of cannabis use and subsequent problems with other illicit drugs and related behavioral problems (Lynskey et al., 2003,2006; Agrawal et al., 2004; Lessem et al., 2006). Thus, while our LOD scores are modest, they will be useful in weighting findings from other independent linkage studies, both of cannabis-related behaviors as well as correlated phenotypes of substance use and misuse and of conduct problems. Therefore, replication of our linkage findings, in conjunction with gene association studies will greatly improve our current understanding of genetic risk on multiple stages of cannabis use behaviors—a phenotype currently underutilized in genomic research.

Conflict of interest

None.

Acknowledgements

We thank Gu Zhu for assistance with plotting software.

Role of funding source: this study is supported by NIH grants DA12854 (to P.A.F.M.), AA07728, AA07580 and AA13321 (to A.C.H.), AA13320 (to R.D.T.), DA18267 & DA18660 (to M.T.L.) and DA023668 (to A.A.) and DA019951 (to M.L.P.) and grants from the Australian National Health and Medical Research Council. Funding agencies did not participate in research and data analysis.

Contributors: all authors read, revised and approved this submission. Agrawal contributed for analysis, writing initial draft and revisions. Morley contributed for analysis, statistical and bioinformatics expertise. Hansell provided us statistical expertise. Pergadia provided us phenotypic expertise, quality control for phenotypic data on several component projects. Martin worked as the supervisor and principal investigator on component project and provided statistical expertise. Statham worked as the study co-ordinator. Todd provided us statistical and phenotypic expertise and worked as the principal investigator on component project. Madden provided us phenotypic expertise and worked as the principal investigator on component project. Heath provided us phenotypic expertise and worked as the principal investigator on component project. Whitfield provided us phenotypic expertise and worked as the principal investigator on component project. Montgomery was involved in genotyping supervision and statistical and bioinformatics expertise. Lynskey helped us in on-site supervision, conceptualization of phenotypes, critical revisions and phenotypic expertise.

References

- Abecasis, G.R., Cardon, L.R., Cookson, W.O., Sham, P.C., Cherny, S.S., 2001. Association analysis in a variance components framework. *Genet. Epidemiol.* 21 (Suppl. 1) (S341-6, S341-S346).
- Abecasis, G.R., Cherny, S.S., Cookson, W.O., Cardon, L.R., 2002. Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. *Nat. Genet.* 30, 97–101.
- Agrawal, A., Lynskey, M., 2006. The genetic epidemiology of cannabis use, abuse and dependence: a review. *Addiction* 101, 801–812.
- Agrawal, A., Heath, A.C., Grant, J.D., Pergadia, M.L., Statham, D.J., Bucholz, K.K., Martin, N.G., Madden, P.A., 2006. Assortative mating for cigarette smoking and for alcohol consumption in female Australian twins and their spouses. *Behav. Genet.* 36, 553–566.
- Agrawal, A., Hinrichs, A., Dunn, G., Bertelsen, S., Dick, D., Saccone, S.F., Saccone, N.L., Gruza, R., Wang, J.C., Cloninger, C.R., Edenberg, H., Foroud, T., Hesselbrock, V., Kramer, J., Bucholz, K., Kuperman, S., Nurnberger, J., Porjesz, B., Schuckit, M., Goate, A., Bierut, L., 2008. Linkage scan for quantitative traits identifies new regions of interest for substance dependence in the Collaborative Study on the Genetics of Alcoholism (COGA) Sample. *Drug Alcohol Depend.* 93, 12–20.
- Agrawal, A., Neale, M., Jacobson, K., Prescott, C.A., Kendler, K.S., 2005. Illicit drug use and abuse/dependence: modeling of two-stage variables using the CCC approach. *Addict. Behav.* 30, 1043–1048.
- Agrawal, A., Neale, M.C., Prescott, C.A., Kendler, K.S., 2004. A twin study of early cannabis use and subsequent use and abuse/dependence of other illicit drugs. *Psychol. Med.* 34, 1227–1237.
- Australian Institute of Health and Welfare, 2005. National Drug Strategy Household Survey: First Results, Technical Report, Canberra (AUST): Australian Institute of Health and Welfare, Drug Statistics Series No. 13.
- Bierut, L.J., Madden, P.A., Breslau, N., Johnson, E.O., Hatsukami, D., Pomerleau, O.F., Swan, G.E., Rutter, J., Bertelsen, S., Fox, L., Fugman, D., Goate, A.M., Hinrichs, A.L., Konvicka, K., Martin, N.G., Montgomery, G.W., Saccone, N.L., Saccone, S.F., Wang, J.C., Chase, G.A., Rice, J.P., Ballinger,

- D.G., 2007. Novel genes identified in a high-density genome wide association study for nicotine dependence. *Hum. Mol. Genet.* 16, 24–35.
- Compton, W.M., Grant, B.F., Colliver, J.D., Glantz, M.D., Stinson, F.S., 2004. Prevalence of marijuana use disorders in the United States: 1991–1992 and 2001–2002. *JAMA* 291, 2114–2121.
- EMCDDA, 2006. Annual report 2006: the state of the drugs problem in Europe. European Monitoring Center for Drugs and Drug Addiction, Lisbon.
- Epstein, M.P., Duren, W.L., Boehnke, M., 2000. Improved inference of relationship for pairs of individuals. *Am. J. Hum. Genet.* 67, 1219–1231.
- Gelernter, J., Panhuysen, C., Weiss, R., Brady, K., Hesselbrock, V., Rounsaville, B., Poling, J., Wilcox, M., Farrer, L., Kranzler, H.R., 2005. Genomewide linkage scan for cocaine dependence and related traits: significant linkages for a cocaine-related trait and cocaine-induced paranoia. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 136, 45–52.
- Gelernter, J., Panhuysen, C., Wilcox, M., Hesselbrock, V., Rounsaville, B., Poling, J., Weiss, R., Sonne, S., Zhao, H., Farrer, L., Kranzler, H.R., 2006. Genomewide linkage scan for opioid dependence and related traits. *Am. J. Hum. Genet.* 78, 759–769.
- Grant, J.D., Heath, A.C., Bucholz, K.K., Madden, P.A., Agrawal, A., Statham, D.J., Martin, N.G., 2007. Spousal concordance for alcohol dependence: evidence for assortative mating or spousal interaction effects? *Alcohol. Clin. Exp. Res.* 31, 717–728.
- Hall, W., Degenhardt, L., 2007. Prevalence and correlates of cannabis use in developed and developing countries. *Curr. Opin. Psychiatry* 20, 393–397.
- Hansell, N.K., Agrawal, A., Whitfield, J.B., Morley, K., Zhu, G., Lind, P., Pergadia, M.L., Madden, P., Todd, R.D., Heath, A., Martin, N.G., 2008. Test-retest reliability and heritability of telephone interview measures of alcohol consumption and dependence. *Twin Res. Hum. Genet.* 11, 287–305.
- Heath, A.C., Bucholz, K.K., Madden, P.A., Dinwiddie, S.H., Slutske, W.S., Bierut, L.J., Statham, D.J., Dunne, M.P., Whitfield, J.B., Martin, N.G., 1997. Genetic and environmental contributions to alcohol dependence risk in a national twin sample: consistency of findings in women and men. *Psychol. Med.* 27, 1381–1396.
- Heath, A.C., Howells, W., Kirk, K.M., Madden, P.A., Bucholz, K.K., Nelson, E.C., Slutske, W.S., Statham, D.J., Martin, N.G., 2001. Predictors of non-response to a questionnaire survey of a volunteer twin panel: findings from the Australian 1989 twin cohort. *Twin Res.* 4, 73–80.
- Heath, A.C., Martin, N.G., 1994. Genetic influences on alcohol consumption patterns and problem drinking: results from the Australian NH&MRC twin panel follow-up survey. *Ann. NY Acad. Sci.* 708, 72–85.
- Hopfer, C.J., Lessem, J.M., Hartman, C.A., Stallings, M.C., Cherny, S.S., Corley, R.P., Hewitt, J.K., Krauter, K.S., Mikulich-Gilbertson, S.K., Rhee, S.H., Smolen, A., Young, S.E., Crowley, T.J., 2006. A genome-wide scan for loci influencing adolescent cannabis dependence symptoms: evidence for linkage on chromosomes 3 and 9. *Drug. Alcohol. Depend.*
- Kendler, K.S., Prescott, C.A., 1998. Cannabis use, abuse, and dependence in a population-based sample of female twins. *Am. J. Psychiatry* 155, 1016–1022.
- Kendler, K.S., Jacobson, K.C., Prescott, C.A., Neale, M.C., 2003a. Specificity of genetic and environmental risk factors for use and abuse/dependence of cannabis, cocaine, hallucinogens, sedatives, stimulants, and opiates in male twins. *Am. J. Psychiatry* 160, 687–695.
- Kendler, K.S., Prescott, C.A., Myers, J., Neale, M.C., 2003b. The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. *Arch. Gen. Psychiatry* 60, 929–937.
- Knopik, V.S., Heath, A.C., Madden, P.A.F., Bucholz, K.K., Slutske, W., Nelson, E., Statham, D., Whitfield, J.B., Martin, N.G., 2004. Genetic effects on alcohol dependence risk: re-evaluating the importance of psychiatric and other heritable risk factors. *Psychol. Med.* 34, 1519–1530.
- Lessem, J.M., Hopfer, C.J., Haberstick, B.C., Timberlake, D., Ehringer, M.A., Smolen, A., Hewitt, J.K., 2006. Relationship between adolescent marijuana use and young adult illicit drug use. *Behav. Genet.*
- Liu, Q.R., Drgon, T., Johnson, C., Walther, D., Hess, J., Uhl, G.R., 2006. Addiction molecular genetics: 639,401 SNP whole genome association identifies many “cell adhesion” genes. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 141, 918–925.
- Lynskey, M.T., Grant, J.D., Nelson, E.C., Bucholz, K.K., Madden, P.A., Statham, D.J., Martin, N.G., Heath, A.C., 2006. Duration of cannabis use—a novel phenotype? *Addict Behav.* 31, 984–994.
- Lynskey, M.T., Heath, A.C., Bucholz, K.K., Slutske, W.S., Madden, P.A., Nelson, E.C., Statham, D.J., Martin, N.G., 2003. Escalation of drug use in early-onset cannabis users vs co-twin controls. *JAMA* 289, 427–433.
- Madden, P.A., 2005. Genetic Vulnerability to Nicotine Dependence. *Am. J. Med. Genet. Part B.*
- McGue, M., Elkins, I., Iacono, W.G., 2000. Genetic and environmental influences on adolescent substance use and abuse. *Am. J. Med. Genet.* 96, 671–677.
- Nusbaum, C., Zody, M.C., Borowsky, M.L., Kamal, M., Kodira, C.D., Taylor, T.D., Whittaker, C.A., Chang, J.L., Cuomo, C.A., Dewar, K., FitzGerald, M.G., Yang, X., Abouelleil, A., Allen, N.R., Anderson, S., Bloom, T., Bugalter, B., Butler, J., Cook, A., DeCaprio, D., Engels, R., Garber, M., Gnirke, A., Hafez, N., Hall, J.L., Norman, C.H., Itoh, T., Jaffe, D.B., Kuroki, Y., Lehoczky, J., Lui, A., Macdonald, P., Mauceli, E., Mikkelsen, T.S., Naylor, J.W., Nicol, R., Nguyen, C., Noguchi, H., O’Leary, S.B., O’Neill, K., Piqani, B., Smith, C.L., Talamas, J.A., Topham, K., Totoki, Y., Toyoda, A., Wain, H.M., Young, S.K., Zeng, Q., Zimmer, A.R., Fujiiyama, A., Hattori, M., Birren, B.W., Sakaki, Y., Lander, E.S., 2005. DNA sequence and analysis of human chromosome 18. *Nature* 437, 551–555.
- Rhee, S.H., Hewitt, J.K., Young, S.E., Corley, R.P., Crowley, T.J., Stallings, M.C., 2003. Genetic and environmental influences on substance initiation, use, and problem use in adolescents. *Arch. Gen. Psychiatry* 60, 1256–1264.
- Saccone, S.F., Pergadia, M.L., Loukola, A., Broms, U., Montgomery, G.W., Wang, J.C., Agrawal, A., Dick, D.M., Heath, A.C., Todorov, A.A., Maunu, H., Heikkila, K., Morley, K.I., Rice, J.P., Todd, R.D., Kaprio, J., Peltonen, L., Martin, N.G., Goate, A.M., Madden, P.A., 2007. Genetic linkage to chromosome 22q12 for a heavy smoking quantitative trait in two independent samples. *Am. J. Hum. Genet.* 80, 856–866.
- Stallings, M.C., Corley, R.P., Dennehey, B., Hewitt, J.K., Krauter, K.S., Lessem, J.M., Mikulich-Gilbertson, S.K., Rhee, S.H., Smolen, A., Young, S.E., Crowley, T.J., 2005. A genome-wide search for quantitative trait loci that influence antisocial drug dependence in adolescence. *Arch. Gen. Psychiatry* 62, 1042–1051.
- Stallings, M.C., Corley, R.P., Hewitt, J.K., Krauter, K.S., Lessem, J.M., Mikulich, S.K., Rhee, S.H., Smolen, A., Young, S.E., Crowley, T.J., 2003. A genome-wide search for quantitative trait loci influencing substance dependence vulnerability in adolescence. *Drug. Alcohol. Depend.* 70, 295–307.
- Substance Abuse and Mental Health Services Administration (SAMHSA). 2004 National Survey on Drug Use and Health. 2005.
- Uhl, G.R., Liu, Q.R., Walther, D., Hess, J., Naiman, D., 2001. Polysubstance abuse-vulnerability genes: genome scans for association, using 1004 subjects and 1494 single-nucleotide polymorphisms. *Am. J. Hum. Genet.* 69, 1290–1300.
- Young, S.E., Rhee, S.H., Stallings, M.C., Corley, R.P., Hewitt, J.K., 2006. Genetic and environmental vulnerabilities underlying adolescent substance use and problem use: general or specific? *Behav. Genet.* 36, 603–615.
- Whitfield, J.B., Zhu, G., Madden, P.A.F., Neale, M.C., Heath, A.C., Martin, N.G., 2004. The genetics of alcohol intake and of alcohol dependence. *Alcohol. Clin. Exp. Res.* 28, 1153–1160.