Addiction Biology

HUMAN GENETIC STUDY

doi:10.1111/j.1369-1600.2011.00320.x

No association of candidate genes with cannabis use in a large sample of Australian twin families

Karin J. H. Verweij[1,2], Brendan P. Zietsch[1,2], Jimmy Z. Liu[1], Sarah E. Medland[1,2], Michael T. Lynskey[3], Pamela A. F. Madden[3], Arpana Agrawal[3], Grant W. Montgomery[1], Andrew C. Heath[3] & Nicholas G. Martin[1]

Genetic Epidemiology, Molecular Epidemiology, and Queensland Statistical Genetics Laboratories, Queensland Institute of Medical Research, Brisbane, Queensland, Australia[1], School of Psychology, University of Queensland, Brisbane, Queensland, Australia[2] and Department of Psychiatry, Washington University School of Medicine, St Louis, MO, USA[3]

ABSTRACT

While there is solid evidence that cannabis use is heritable, attempts to identify genetic influences at the molecular level have yielded mixed results. Here, a large twin family sample ($n = 7452$) was used to test for association between 10 previously reported candidate genes and lifetime frequency of cannabis use using a gene-based association test. None of the candidate genes reached even nominal significance ($P < 0.05$). The lack of replication may point to our limited understanding of the neurobiology of cannabis involvement and also to potential publication bias and false-positive findings in previous studies.

Keywords Association, cannabis, drug, genes, genetics, replication.

Correspondence to: Karin J.H. Verweij, Genetic Epidemiology, Queensland Institute of Medical Research, 300 Herston Road, Brisbane, Queensland 4006, Australia. E-mail: karin.verweij@qimr.edu.au

Twin and family studies have estimated the heritability of cannabis use phenotypes (lifetime use, frequency of use and abuse/dependence) at between 40–60% (Verweij *et al.* 2010) and have also shown substantial overlap in the genetic factors influencing earlier (experimental/regular use) and later (abuse/dependence) stages of cannabis use (Agrawal & Lynskey 2006). Identification of the specific genes contributing to cannabis use variation could improve the knowledge regarding the biological processes that underlie cannabis use, and potentially substance use in general. Linkage and candidate gene association studies have identified a handful of potentially important genetic variants, but different studies have yielded inconsistent results (for an overview, see Agrawal & Lynskey 2009). The present study tests for replication of previously identified candidate gene associations for cannabis use in a large population-based sample of Australian twin families.

For this study, we used data from 7452 Caucasian individuals (3334 males and 4118 females from 2595 independent families, mean age 43.2 ± 11.8 years) from the Australian Twin Registry (ATR) for whom we have both genotypic and phenotypic data. Individuals participated in various studies between 1996 and 2004. The ATR is a population-based sample, but a subset of the sample was ascertained based on large sibship size, or having a relative with nicotine or alcohol dependence. The effective sample size (i.e. correcting for non-independence of family members) was calculated to be 4312. Frequency of lifetime cannabis use used an open response format, but due to extreme skew was transformed by assigning values to seven bins (see Fig. 1).

DNA samples were collected in accordance with standard protocols and were genotyped in different waves on different Illumina platforms (Illumina 317 K, Illumina HumanCNV370-Quadv3, Illumina Human610-Quad, Illumina Inc., San diego, CA, USA). Imputation using MACH allowed combining data from different genotyping arrays and improved coverage of single-nucleotide polymorphisms (SNPs) in the candidate genes (see Medland *et al.* 2009).

We performed a genome-wide association (GWA) analysis where SNPs ($n = 2\,380\,486$) across the entire genome were systematically tested for association with lifetime frequency of cannabis use. The analyses were performed in Merlin (Chen & Abecasis 2007), which

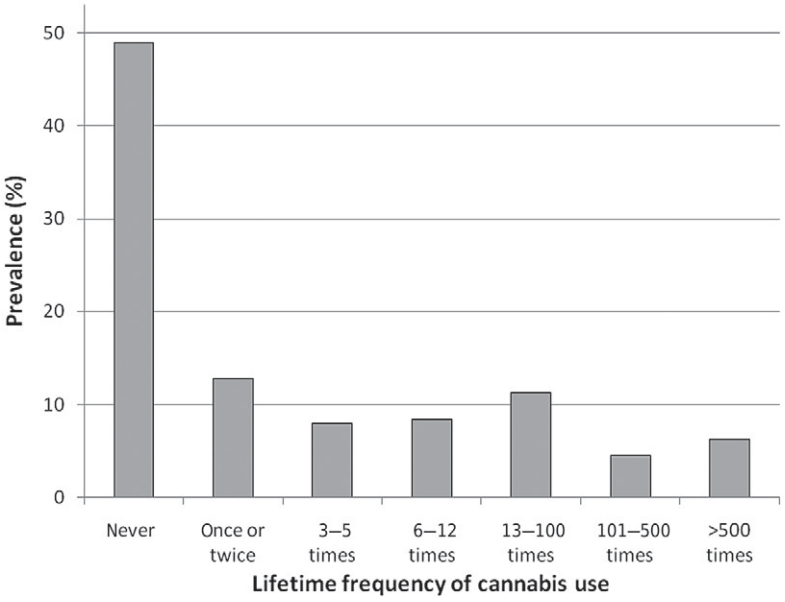

Figure 1 Prevalence of the different response categories for lifetime frequency of cannabis use

accounts for family relationships including MZ twins, and we included age and sex as covariates. Results of this GWA were then used to run a gene-based association test (VEGAS, Liu *et al.* 2010). VEGAS summarises evidence for association on a per gene basis ($n = 17\,591$ autosomal genes) by considering the *P*-value of all SNPs within genes [including ± 50 kb from the 5′ and 3′ untranslated region (UTR)], while accounting for linkage disequilibrium (LD) and number of SNPs per gene. As such, the gene-based test identifies genes that show more signal of association than expected by chance given their length and LD between the SNPs. We tested for association between lifetime frequency of cannabis use with 10 previously defined candidate genes for cannabis use disorders, as described by Agrawal & Lynskey (2009). In their review, they include genes that are posited to have specific influences on cannabis use because the genes relate to the action and metabolism of exogenous cannabinoids. They also describe non-specific candidate genes that potentially influence the biological basis of substance use in a more general way. For the latter cases, they included genes that encode the major neurotransmitter systems and have previously shown association with substance use. The selected candidate genes include genes that are thought to influence different stages of cannabis use. In total, we performed association analyses on 10 candidate genes—based on a Bonferroni correction for multiple testing, we declared the significance level to be $\alpha = 0.005$ (0.05/10).

The heritability of lifetime frequency of cannabis use was estimated in Merlin to be 45%, consistent with previous research (Verweij *et al.* 2010). Table 1 shows the results of our gene-based test for association, including information regarding the number of tagged SNPs per gene, the ranking of the gene (out of 17 591 genes), the number of SNPs within the gene with a *P*-value below 0.01 or 0.05, and the percentage of the variance of the gene that is covered by our SNPs. Results indicate that none of the candidate genes reached even nominal significance ($P < 0.05$), and none of them was within the top 1000 genes.

Because genes cover a lot of SNPs (up to 402 in our candidate genes), we also checked for association between specific SNPs (within our candidate genes) that have previously been found to be associated with cannabis use phenotypes (see Caspi *et al.* 2005; Agrawal & Lynskey 2009). Again, none of the candidate SNPs reached even nominal significance (see Table 2). Moreover, none of the other SNPs within the candidate genes reached Bonferroni corrected significance ($\alpha = 0.05$).

Empirical power estimation using simulated data showed 97%, 51%, and 21% power to detect a candidate gene or SNP that accounts for 0.5%, 0.2%, or 0.1% of the variation at a Bonferroni-corrected alpha level, respectively. Even with this power, we could not replicate association for any of the previously identified candidate genes or SNPs. This suggests the biology underlying cannabis involvement is highly complex, and that our understanding of the biochemical and addictive processes governing cannabis use is nascent. These findings could also indicate that previous associations were false-positives, and suggests that publication bias may misrepresent the actual association between the genetic variant and phenotype. This is not the first study demonstrating that candidate genes studies should be interpreted with caution—for example, Bosker *et al.* (2010) could not replicate most candidate genes previously associated with major depressive disorder.

Table 1 Association results of the ten candidate genes with frequency of cannabis use.

Gene	*Chromosome*	*Start position gene*[a]	*End position gene*[a]	*Number of tagged SNPs*	*P-value*	*Rank*	*SNPs at* $\alpha < 0.05$/ $\alpha < 0.01$	*% of gene variance covered*
CNR1	6	88906303	88911775	145	0.45	7814	10/3	96%
CNR2	1	24073046	24112404	15	0.06	1089	0/0	86%
FAAH	1	46632525	46652107	90	0.96	16851	0/0	97%
MGLL	3	128890598	129024741	170	0.25	4342	13/1	97%
TRPV1	17	3415489	3459454	116	0.31	5394	6/1	95%
GPR55	2	231480286	231498185	88	0.95	16591	0/0	91%
GABRA2	4	45946338	46086813	145	0.47	8179	0/0	98%
DRD2	11	112785526	112851211	197	0.89	15389	0/0	100%
OPRM1	6	154402135	154609693	402	0.37	6352	39/19	98%
COMT	22	18309308	18336530	131	0.17	2956	17/8	98%

[a]Note that the analyses included ±50 kb from the gene border. CNR1/CNR2 = cannabinoid receptors 1 and 2; COMT = catecholamine-o-methyl transferase; DRD2 = dopamine receptor D2; FAAH = fatty acid amide hydrolase; GABRA2 = gamma-amino butyric acid; GPR55 = orphan cannabinoid receptor; MGLL = monoglyceride lipase; OPRM1 = Mu-opiod receptor 1; SNP = single-nucleotide polymorphisms; TRPV1 = transient receptor potential vanilloid 1.

Table 2 Association results of the candidate SNPs as reported by Agrawal & Lynskey (2009) with frequency of cannabis use.

SNP	*Gene*	*P-value*
rs2023239	CNR1	0.39
rs806379	CNR1	0.88
rs1535255	CNR1	Not tagged
rs806380	CNR1	0.78
rs6454674	CNR1	0.39
rs806368	CNR1	0.50
rs12720071	CNR1	0.94
rs806379	CNR1	0.88
rs2501432	CNR2	Not tagged
rs324420	FAAH	0.58
rs279858	GABRA2	0.27
rs1799971	OPRM1	0.12
rs4680	COMT	0.31

CNR1/CNR2 = cannabinoid receptors 1 and 2; COMT = catecholamine-o-methyl transferase; FAAH = Fatty acid amide hydrolase; GABRA2 = Gamma-amino butyric acid; OPRM1 = Mu-opiod receptor 1; SNP = single-nucleotide polymorphisms.

The present study had several limitations. Firstly, the phenotype (lifetime frequency of use) is not the ideal replication phenotype for some of the genes, which were originally associated with late stages of cannabis use (abuse). However, some of the candidate genes are thought to be involved in earlier stages of cannabis use, and twin studies show the different stages of cannabis involvement are largely influenced by the same genetic factors (Agrawal & Lynskey 2006), so we would still expect replication. Secondly, the non-normal distribution of our phenotype decreases statistical power, while also elevating the chance of false-positives. However, the first problem would be heavily outweighed by the much larger sample than previous studies, and the latter is not an issue for our findings as there were no positive findings.

While our findings indicate that none of the 10 candidate genes explain a substantial portion of the variance in frequency of lifetime cannabis use, this does not imply that lifetime cannabis use is not heritable. Other, unidentified common genetic variants may play a role in cannabis use, and the variants may be numerous with individually small effects. We therefore need to look systematically across the whole genome for these variants—that is, a large-scale GWA study is needed. Just such a study is now underway, incorporating multiple datasets from around the world.

Acknowledgements

We would like to thank the twins and their families registered at the Australian Twin Registry for their participation. We also thank Dixie Statham (sample collection); Lisa Bowdler, Steven Crooks (DNA processing); Scott Gordon (QC genotype data); David Smyth, Harry Beeby, and Daniel Park (IT support). Funding was provided by the Australian National Health and Medical Research Council (241944, 339462, 389927, 389875, 389891, 389892, 389938, 442915, 442981, 496739, 552485, 552498), the Australian Research Council (A7960034, A79906588, A79801419, DP0770096, DP0212016, DP0343921), the FP-5 GenomEUtwin Project QLG2-CT-2002-01254), and the U.S. National Institutes of Health (NIH grants AA07535, AA10248, AA13320, AA13321, AA13326, AA14041, DA12854, MH66206). A portion of the genotyping on which this study was based (Illumina 370 K scans) was carried out at the Center for Inherited Disease Research, Baltimore, through an access award to our late colleague Dr Richard Todd (Psychiatry, Washington University School of Medicine, St Louis). Statistical analyses were carried out on the Genetic Cluster Computer, which is financially supported

by the Netherlands Scientific Organization (NWO 480-05-003). K.J.H.V. is supported by an ANZ Trustees PhD scholarship in Medical Research. BPZ is supported by a UQ Postdoctoral Fellowship.

Authors Contribution

KJHV was responsible for the study concept and design of the study. NGM, ACH, PAFM, and GWM contributed to the data acquisition. KJHV performed the data analysis and interpretation of findings. SEM, and JZL assisted with data analysis. KJHV drafted the manuscript. BPZ, MTL, and AA, provided critical revision of the manuscript for important intellectual content. All authors critically reviewed content and approved final version for publication.

References

Agrawal A, Lynskey MT (2006) The genetic epidemiology of cannabis use, abuse and dependence. Addiction 101:801–812.

Agrawal A, Lynskey MT (2009) Candidate genes for cannabis use disorders: findings, challenges and directions. Addiction 104:518–532.

Bosker FJ, Hartman CA, Nolte IM, Prins BP, Terpstra P, Posthuma D, van Veen T, Willemsen G, Derijk RH, de Geus EJ, Hoogendijk WJ, Sullivan PF, Penninx BW, Boomsma DI, Snieder H, Nolen WA (2010) Poor replication of candidate genes for major depressive disorder using genome-wide association data. Mol Psychiatry. Epub ahead of print.

Caspi A, Moffitt TE, Cannon M, McClay J, Murray R, Harrington H, Taylor A, Arseneault L, Williams B, Braithwaite A, Poulton R, Craig IW (2005) Moderation of the effect of adolescent-onset cannabis use on adult psychosis by a functional polymorphism in the catechol-O-methyltransferase gene: longitudinal evidence of a gene X environment interaction. Biol Psychiatry 57:1117–1127.

Chen WM, Abecasis GR (2007) Family-based association tests for genomewide association scans. Am J Hum Genet 81:913–926.

Liu JZ, McRae AF, Nyholt DR, Medland SE, Wray NR, Brown KM, AMFS Investigators, Hayward NK, Montgomery GW, Visscher PM, Martin NG, Macgregor S (2010) A versatile gene-based test for genome-wide association studies. Am J Hum Genet 87:139–145.

Medland SE, Nyholt DR, Painter JN, McEvoy BP, McRae AF, Zhu G, Gordon SD, Ferreira MAR, Wright MJ, Henders AK, Campbell MJ, Duffy DL, Hansell NK, Macgregor S, Slutske WS, Heath AC, Montgomery GW, Martin NG (2009) Common variants in the trichohyalin gene are associated with straight hair in Europeans. Am J Hum Genet 85:750–755.

Verweij KJH, Zietsch BP, Lynskey MT, Medland SE, Neale MC, Martin NG, Boomsma DI, Vink JM (2010) Genetic and environmental influences on cannabis use initiation and problematic use: a meta-analysis of twin studies. Addiction 105:417–430.

