A large-scale genome-wide association study meta-analysis of cannabis use disorder

Citation for published version:
Johnson, EC, Demontis, D, Thorgeirsson, TE, Walters, R, Polimanti, R, Hatoum, AS, Sanchez-roije, S, Paul, SE, Wendt, FR, Clarke, T, Lai, D, Reginsson, GW, Zhou, H, He, J, Baranger, DAA, Gudbjartsson, DF, Wedow, R, Adkins, D, Adkins, A, Alexander, J, Bacanu, S, Bigdeli, TB, Boden, J, Brown, SA, Bucholz, K, Bybjerg-grauholm, J, Corley, RP, Degenhardt, L, Dick, D, Domingue, B, Fox, L, Goate, AM, Gordon, S, Hack, L, Hancock, D, Hartz, S, Hickie, IB, Hougaard, DM, Krauter, K, Lind, P, Mcclintick, J, Mcqueen, M, Meyers, J, Montgomery, G, Mors, O, Mortensen, PB, Nordinoff, M, Pearson, JF, Peterson, R, Reynolds, M, Rice, J, Runarsdottir, V, Saccone, N, Sherva, R, Silberg, JL, Tarter, R, Tyrfingsson, S, Wall, T, Wendt, B, Werge, T, Wetherill, L, Wright, MJ, Zeilers, S, Adams, MJ, Bierut, L, Boardman, J, Copeland, W, Farrer, L, Foroud, TM, Gillespie, NA, Grucza, RA, Harris, KM, Heath, A, Hesselbrock, V, Hewitt, J, Hopler, C, Horwood, J, Iacono, WG, Johnson, E, Kendler, K, Kennedy, M, Kranzler, H, Madden, PAF, Maes, H, Maher, B, Martin, NG, Mcque, M, McIntosh, AM, Medland, S, Nelson, EC, Porjesz, B, Riley, B, Stallings, M, Vanyukov, M, Vrieze, S, Davis, LK, Bogdan, R, Gelernter, J, Edenberg, H, Stefansson, K, Børglum, AD, Agrawal, A, Walters, R, Polimanti, R, Johnson, E, Mcclintick, J, Hatoum, A, He, J, Wendt, F, Zhou, H, Adkins, A, Aliev, F, Bacanu, S, Batzler, A, Bertelsen, S, Biernacka, J, Bigdeli, B, Chen, D, Choi, D, Copeland, W, Culverhouse, R, Dahmen, N, Degenhardt, L, Domngue, B, Elson, S, Frye, M, Gäbel, W, Hayward, C, Ising, M, Keyes, M, Kieler, F, Kramer, J, Kuperman, S, Lucea, S, Lynskey, M, Maier, W, Mann, K, Männistö, S, Müller-myhsk, B, Murray, A, Nurnberger, J, Palotie, A, Preuss, U, Räikkönen, K, Reynolds, M, Riding, M, Scherbaum, N, Schuckit, M, Soyka, M, Treutlein, J, Witt, S, Wodarz, N, Zill, P, Adkins, D, Boden, J, Boomsma, D, Bierut, L, Brown, S, Bucholz, K, Cichon, S, Costello, EJ, De Witt, H, Diazgranados, N, Dick, D, Eriksson, J, Farrer, L, Foroud, T, Gillespie, NA, Goate, A, Goldman, D, Grucza, R, Hancock, D, Harris, KM, Heath, A, Hesselbrock, V, Hewitt, J, Hopfer, C, Horwood, J, Iacono, W, Johnson, E, Kaprio, J, Karpyak, V, Kendler, K, Kranzler, H, Krauter, K, Lichtenstein, P, Lind, P, McGue, M, Mackillop, J, Madden, P, Maes, H, Magnusson, P, Martin, N, Medland, S, Montgomery, G, Nelson, E, Nöthen, M, Palmer, A, Pederson, N, Pennix, B, Porjesz, B, Rice, J, Rietschel, M, Riley, B, Rose, R, Rujescu, D, Shen, P, Silberg, J, Stallings, M, Tarter, R, Vanyukov, M, Vrieze, S, Wall, T, Whitfield, J, Zhao, H, Neale, B, Gelernter, J, Edenberg, H & Agrawal, A 2020, ‘A large-scale genome-wide association study meta-analysis of cannabis use disorder’, The Lancet Psychiatry, vol. 7, no. 12, pp. 1032-1045.
https://doi.org/10.1016/S2215-0366(20)30339-4

Digital Object Identifier (DOI):
10.1016/S2215-0366(20)30339-4

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
The Lancet Psychiatry
A large-scale genome-wide association study meta-analysis of cannabis use disorder

Emma C Johnson*, Ditte Demontis*, Thorgerir E Thorgerisson*, Raymond K Walters, Renato Polimanti, Alexander S Hatoum, Sandra Sanchez-Roige, Sarah E Paul, Frank R Wendt, Tondi-Kim Clarke, Dongbing Lai, Gunnar W Reginsson, Hang Zhou, June He, David A A Baranger, Daniel F Gudbjartsson, Robbee Wedow, Daniel E Adkins, Amy E Adkins, Jeffry Alexander, Silviu-Alin Bacanu, Tim B Bigdeli, Joseph Boden, Sandra A Brown, Kathleen K Bucholz, Jonas Bybjerg-Grauholm, Robin P Corley, Louisa Degenhardt, Danielle M Dick, Benjamin W Domingue, Louis Fox, Alison M Goate, Scott D Gordon, Laura M Hack, Dana B Hancock, Sarah M Hartz, Ian B Hickie, David M Hougaard, Kenneth Krauter, Penelope A Lind, Jeanette N McClintick, Matthew B McQueen, Jacquelyn L Meyers, Grant W Montgomery, Ole Mors, Preben B Mortensen, Merete Nordenstof, John F Pearson, Roseann E Peterson, Maureen D Reynolds, John P Rice, Valgerdur Runarsdottir, Nancy L Saccone, Richard Sherva, Judy L Silberg, Ralph E Tarter, Thorarinn Tyrfingsson, Tamara L Wall, Bradley T Webb, Thomas Werge, Leah Wetherill, Margaret J Wright, Stephanie Zellers, Mark J Adams, Laura J Bierut, Jason D Boardman, William E Copeland, Lindsay A Farer, Tatiana M Foroud, Nathan A Gillespie, Richard A Grucza, Kathleen Mullan Harris, Andrew C Heath, Victor Hesselbrock, John K Hewitt, Christian J Hopfer, John Horwood, William G Iacono, Eric O Johnson, Kenneth S Kendler, Martin A Kennedy, Henry R Kranzler, Pamela A F Madden, Hermine H Maes, Brion S Maher, Nicholas G Martin, Matthew McGuie, Andrew M McIntosh, Sarah E Medland, Elliott C Nelson, Bernice Poiriez, Brian P Riley, Michael C Stallings, Michael M Vanyukov, Scott Vrieze, the Psychiatric Genomics Consortium Sub stance Use Disorders Workgroup, Lea Davis, Ryan Bogdan, Joel Gelernter†, Howard J Edenberg†, Kari Stfanossen, Anders D Børglum†, Arpana Agrawal†

Summary

Background Variation in liability to cannabis use disorder has a strong genetic component (estimated twin and family heritability about 50–70%) and is associated with negative outcomes, including increased risk of psychopathology. The aim of the study was to conduct a large genome-wide association study (GWAS) to identify novel genetic variants associated with cannabis use disorder.

Methods To conduct this GWAS meta-analysis of cannabis use disorder and identify associations with genetic loci, we used samples from the Psychiatric Genomics Consortium Sub stance Use Disorders working group, iPSYCH, and deCODE (20,916 case samples, 363,116 control samples in total), contrasting cannabis use disorder cases with controls. To examine the genetic overlap between cannabis use disorder and 22 traits of interest (chosen because of previously published phenotypic correlations [eg, psychiatric disorders] or hypothesised associations [eg, chronotype] with cannabis use disorder), we used linkage disequilibrium score regression to calculate genetic correlations.

Findings We identified two genome-wide significant loci: a novel chromosome 7 locus (FOXP2, lead single-nucleotide polymorphism [SNP] rs7783012; odds ratio [OR] 1·11, 95% CI 1·07–1·15, p=1·9×10−⁹) and the previously identified chromosome 8 locus (near CHRNA2 and EPHX2, lead SNP rs4732724; OR 0·89, 95% CI 0·86–0·93, p=6·4×10−²¹), but they showed significantly different genetic correlations with 12 of the 22 traits we tested, suggesting at least partially different genetic underpinnings of cannabis use and cannabis use disorder. Cannabis use disorder was positively genetically correlated with other psychopathology, including ADHD, major depression, and schizophrenia.

Interpretation These findings support the theory that cannabis use disorder has shared genetic liability with other psychopathology, and there is a distinction between genetic liability to cannabis use and cannabis use disorder.

Funding National Institute of Mental Health; National Institute on Alcohol Abuse and Alcoholism; National Institute on Drug Abuse; Center for Genomics and Personalized Medicine and the Centre for Integrative Sequencing: The European Commission, Horizon 2020; National Institute of Child Health and Human Development; Health Research Council of New Zealand; National Institute on Aging; Wellcome Trust Case Control Consortium; UK Research and Innovation Medical Research Council (UKRI MRC); The Brain & Behavior Research Foundation; National Institute on Deafness and Other Communication Disorders; Substance Abuse and Mental Health Services Administration (SAMHSA); National Institute of Biomedical Imaging and Bioengineering; National Health and Medical Research Council (NHMRC) Australia; Tobacco-Related Disease Research Program of the University of California; Families for Borderline Personality Disorder Research (Beth and Rob Elliott) 2018 NARSAD Young Investigator Grant; The National Child Health Research Foundation (Cure Kids); The Canterbury Medical Research Foundation; The New Zealand Lottery Grants Board; The University of Otago; The Carney Centre for Pharmacogenomics; The James Hume Bequest Fund; National Institutes of Health: Genes, Environment and Health Initiative; National Institutes of Health; National Cancer Institute; The William T Grant Foundation; Australian...
Research in context

Evidence before this study
Cannabis use disorder is heritable (50–70% according to twin and family studies), yet identification of genomic variants associated with cannabis use disorder from genomic-wide association studies (GWASs) remains sparse. We surveyed all peer-reviewed journal publications in English on GWASs of cannabis use disorder or cannabis dependence using Google Scholar and PubMed, published between Jan 1, 1990, and April 1, 2020. Search terms included “cannabis dependence”, “cannabis use disorder”, “marijuana dependence”, “marijuana abuse”, “marijuana use disorder”, and “GWAS”. The most promising finding to date is a variant that is a cis-eQTL for CHRNA2 (Demontis and colleagues), which was replicated in an independent dataset for cannabis use disorder. Independently, GWAS of cannabis use have identified multiple genetic risk loci; however, the extent to which the genetics of cannabis use correlates with liability to cannabis use disorder has not been determined. Although GWASs of cannabis use have been studied in the context of a variety of psychiatric and psychosocial correlates, it is expected that some divergent associations will be seen when looking at cannabis use disorder. Previous studies have drawn causal links between cannabis exposure and brain volume, but the relationship between genetic liability to cannabis use disorder and brain volume in individuals naive to cannabis has not yet been studied.

Added value of this study
Our study is the current largest GWAS of cannabis use disorder and the first to include a transancestral component. We found a novel risk locus on chromosome 7. The lead risk variant at this locus is an eQTL for FOXF2—a gene previously implicated in risk-taking behaviours. Contrasting cannabis use and cannabis use disorder, we found that increased liability for cannabis use disorder is genetically correlated with low educational attainment, early age at first birth, and higher body-mass index, traits that show opposite directions of association with lifetime cannabis use ever-use. We also found that genetic liability for cannabis use disorder is associated with increased risk of mental health problems, infectious diseases, and respiratory illnesses in a large independent sample. Finally, we found a significant association between increased polygenic liability for cannabis use disorder and low social mobility in cannabis-naïve children, suggesting a potential role of cannabis-related genetic predisposition in early brain development.

Implications of all the available evidence
Cannabis use disorder is a psychiatric illness that is genetically associated with many negative outcomes (including increased risk of psychiatric disorders and respiratory illnesses). Lifetime cannabis use and cannabis use disorder show at least partially divergent genetic influences and associations with relevant traits. Given increasingly permissive cannabis laws and positive perceptions of its safety, the recognition that cannabis use disorder is a serious psychiatric illness should spur prevention and treatment efforts.
European ancestry | African ancestry | European ancestry—case-control individuals

| Cases | Controls | Cases | Controls | Cases | Controls |
|-------|----------|-------|----------|-------|----------|
| CADS  | 397      | 699   | 59       | 35    | 397      |
| CHDS  | 201      | 420   |          |       | 201      |
| FSCD  | 226      | 314   | 199      | 401   | 226      |
| COGEND Nico | 306 | 607   | 154      | 313   | 306      |
| COGEND SAGE | 228 | 830   | 79       | 187   | 228      |
| GEDI-GSMS | 81   | 491   |          |       | 81       |
| ADA   |          | 1000  | 811      |       |          |
| Total | 2397     | 3814  | 1491     | 1767  | 2397     |

Family-based studies

| Add health | 722 | 4071 |   |   |   |
| PGC studies total | 8277 | 23497 | 3848 | 5897 | 5289 |
| IPSYCH | 2758 | 53266 | 0 | 0 | 2758 | 53266 |
| deCODE | 6033 | 280396 | 0 | 0 | 6033 | 280396 |
| Total (European ancestry) | 17068 | 357219 | 0 | 0 | 14080 | 343726 |
| Total (Tranancesral) | 30516 cases | 33416 controls |   |   |   |

Summary statistics

Table 1: Numbers of cases and controls in meta-analysis

Methods

Samples

We performed a GWAS of 20 samples in total: 18 from the Psychiatric Genomics Consortium Substance Use Disorders working group (European ancestry 8277 cases, 23497 controls; African ancestry 3848 cases, 5897 controls), one IPSYCH sample (European ancestry 2758 cases, 53266 controls), and one deCODE sample (European ancestry 6033 cases, 280396 controls; table 1; appendix pp 2–8).

This study was approved by the institutional review board at Washington University School of Medicine and was done in accordance with all relevant ethical regulations. Investigators for each contributing study obtained informed consent from their participants and received ethics approvals from their respective review boards in accordance with applicable regulations. Personal identifiers associated with phenotypic information and samples from deCODE were encrypted using a third-party encryption system. The IPSYCH group used pseudonymised unique identifications.

Measures

Psychiatric Genomics Consortium cases met criteria for a lifetime diagnosis of DSM-IV (or DSM-III-R) cannabis abuse or dependence derived from clinician ratings or semi-structured interviews. Cases from the IPSYCH sample had ICD-10 codes of F12.1 (cannabis abuse) or F12.2 (cannabis dependence), or both in the Danish Psychiatric Central Research Register; the remaining individuals in the sample were used as controls. Cases in the deCODE sample met criteria for lifetime DSM-III-R or DSM-IV cannabis abuse or dependence or DSM-5 cannabis use disorder according to diagnoses made at the National Center of Addiction Medicine in Iceland, whereas controls were derived from the general population of Iceland (appendix pp 2–3). Exposure data were not available for some large groups (eg, IPSYCH and deCODE); therefore, controls were defined regardless of lifetime cannabis exposure across all datasets.

Genotyping: quality control and imputation

For the Psychiatric Genomics Consortium, standard procedures for GWAS quality control and imputation were applied using the Picopili pipeline for case-control groups and the Picopili pipeline for family-based samples. Briefly, variants in each group were filtered for call rate (<5% missingness), followed by individual-level filtering for call rate (<2% missingness) and heterozygosity ([F̂]) <0.20). If available, chromosome X variants were checked to ensure concordance between genotype sex and reported sex. Variants were then filtered more stringently: variants with more than 2% missingness, differential missingness between cases and controls greater than 2%, invariant markers, and those departing from Hardy-Weinberg equilibrium in cases (p<1×10^-6) or controls (p<1×10^-9) were removed (appendix pp 8–10).

Principal components analysis was done on a stringently quality-control set of variants using EIGENSOFT to exclude population outliers, infer ancestry among the retained individuals (using the 1000 Genomes Phase 3 cosmopolitan reference panel), and derive ancestry-specific principal components for inclusion in analyses (appendix p 10). Sample and variant quality-controlled procedures, including filters for call rate, heterozygosity, and departure from Hardy-Weinberg equilibrium, were done within each ancestry group in each sample. Each group was phased using SHAPEIT and imputed using IMPUTE2 to the 1000 Genomes Phase 3 cosmopolitan reference panel (appendix pp 10–11).
individuals were removed and individuals who were cryptically related across groups were excluded from all but one group (appendix p 12). Single-nucleotide polymorphisms (SNPs) were filtered for INFO score of more than 0·8 and minor allele frequency of at least 0·01 before meta-analysis (appendix pp 13–14).

Quality control of iPSYCH data mirrored the process implemented in the Psychiatric Genomics Consortium, with minor deviations in thresholds for exclusion (appendix p 9). As for deCODE, samples were assayed with several Illumina arrays at deCODE genetics. SNPs with low call rate (<95%), significant deviation from Hardy-Weinberg equilibrium (p<0·001), and excessive inheritance error rates (>0·01) were excluded. We did variant imputation on the basis of the IMPUTE HMM model and long-range phasing. Variants were further filtered for imputation INFO score more than 0·8 and minor allele frequency at least 1% inclusion in meta-analysis.

Statistical analysis
We did separate association analyses for each sample (ie, 18 individual samples from Psychiatric Genomics Consortium, iPSYCH, and deCODE) by ancestry. For the eight case-control studies from the Psychiatric Genomics Consortium, imputed dosages were analysed using logistic regression models, implemented in the Ricopili pipeline.11 For family-based samples of the Psychiatric Genomics Consortium, we did association analyses with imputed best-guess genotypes using generalised estimating equations for complex pedigrees, in the Picopili pipeline.7 For calculation of SNP heritability and genetic correlations, subsets of genetically unrelated individuals were selected from each family-based sample from the Psychiatric Genomics Consortium and analysed using logistic regression through Picopili (5289 cases, 10 004 controls). These results were then meta-analysed along with the case-control groups. Psychiatric Genomics Consortium covariates included sex and five to ten within-ancestry principal components to account for population stratification (appendix pp 12–13). Because age was not available in all samples, it was not included as a covariate in the Psychiatric Genomics Consortium analyses. Sensitivity analyses in one representative sample showed this to have no impact on study-specific findings.

In the iPSYCH cohort, logistic regression was done with imputed dosages, covarying for five ancestral principal components, data processing waves, and the presence of another psychiatric disorder (because iPSYCH was established to study major psychiatric disorders, cases of cannabis use disorder and controls include comorbidity).7 Adding sex as a covariate to iPSYCH analyses has been shown not to alter findings.19 deCODE data were analysed using logistic regression of imputed dosage data with sex, age, and county of origin as covariates.21 To account for inflation due to population stratification and relatedness, test statistics were divided by an inflation factor estimated from linkage disequilibrium score regression (LDSR; appendix p 13).

Effective sample size-weighted meta-analyses across case-control and family-based samples within ancestry were done using METAL (appendix pp 13–14).20 First, summary statistics of case-control and family-based samples from the Psychiatric Genomics Consortium were combined and weighted by the effective sample size, because effect sizes from case-control logistic regression analyses and family-based analyses using generalised estimating equations and logistic mixed models are not directly comparable. Then the Psychiatric Genomics Consortium results were meta-analysed with those from the iPSYCH and deCODE samples (between-sample genetic correlations [r] 0·66–0·70). The summary statistics were filtered such that an SNP had to be present in at least two of the three contributing GWASs (deCODE, iPSYCH, and the Psychiatric Genomics Consortium).

We also did a meta-analysis that excluded related individuals from the family-based samples of the Psychiatric Genomics Consortium, using an inverse variance-weighted scheme, to generate summary statistics that produced effect sizes for use in follow-up analyses (140 080 cases, 343 726 controls). A transancestral meta-analysis using METAL21 combined results across the European and African ancestry cohorts, comprising 20·916 individuals with cannabis use disorder (17·068 from European ancestry, 3848 from African ancestry) and 363·116 controls (357·219 from European ancestry, 5897 African ancestry; appendix pp 13–14, 17). Conditional analyses were done in GCTA-COJO24 by conditioning the meta-analysis summary statistics on the lead variants of genome-wide significance.

The FUMA web-based platform25 version 1.3.5e was used for visualisation and annotation, and MAGMA26 was used within the FUMA framework to do gene-based association analyses, with SNPs assigned to genes on the basis of physical position (appendix pp 14–15). We also used Hi-C coupled MAGMA to assign non-coding SNPs (intergenic and intronic) to genes on the basis of their chromatin interactions (exonic and promoter SNPs are still assigned to genes on the basis of their genomic location; appendix p 15).27 Pathway analyses were done using PASCAL to test canonical pathways in the European ancestry sample.28 All variants within all genes were tested, using default settings, with the structure of linkage disequilibrium estimated using the 1000 Genomes European sample as a reference. We used S-PrediXcan29 to examine gene expression differences associated with case-control status, using our summary statistics of cannabis use disorder and transcriptome data from the PredictDB Data Repository for 11 brain regions, liver tissue, whole blood, and two types of adipose tissue. We included these tissues because cannabis use disorder is a psychiatric disorder and of Psychology (A E Adkins PhD, Prof D M Dick PhD), and Department of Human & Molecular Genetics (Prof D M Dick, J L Silberg PhD, H H Maes PhD), College Behavioral and Emotional Health Institute (A E Adkins), Virginia Commonwealth University Alcohol Research Center (J Alexander PhD, S A Bacaro PhD, B T Webb PhD), Prof B P Riley PhD), and Department of Psychiatry, Virginia Institute for Psychiatric and Behavioral Genetics (R E Peterson PhD, J L Silberg, N A Gillespie PhD, Prof K S Kendler MD, H H Maes), Virginia Commonwealth University, Richmond, VA, USA; Department of Psychiatry and Behavioral Sciences, SUNY Downstate Health Sciences University, Brooklyn, NY, USA (T B Bigdeli PhD, J L Meyers PhD, Prof B Foroogh PhD); Department of Psychological Medicine (Prof J Blunden PhD, Prof J Horwood MSc), Biostatistics and Computational Biology Unit (J F Pearson PhD), and Department of Pathology and Biomedical Science (J F Pearson, Prof M A Kennedy PhD), University of Otago, Christchurch, New Zealand; Department for Congenital Disorders, Center for Neonatal Screening, Statens Serum Institut, Copenhagen, Denmark (J Bybjerg-Grauholm, D M Hougaard); Department of Molecular, Cellular, and Developmental Biology (Prof K Krauter PhD), Department of Integrative Physiology (M B McQueen ScD), and Institute for Behavioral Genetics, (R P Corley PhD, Prof J Hewitt PhD, Prof M C Stallings PhD), University of Colorado Boulder, Boulder, CO, USA (Prof K Krauter); National Drug and Alcohol Research Centre, University of New South Wales, Sydney, NSW, Australia (Prof L Degenhardt PhD); Stanford University Graduate School of Education (B W Domingue PhD), and Department of Psychiatry and Behavioral Sciences (M J Hack MD), Stanford University, Stanford, CA, USA; Department of Neuroscience, Ichan School of Medicine at
Mount Sinai, New York, NY, USA (Prof A M Goate DPhil); QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia (Prof S E Medland PhD); Indiana University School of Medicine, Indianapolis, IN, USA (Prof R D Higgins MD); University Hospital, Copenhagen, Denmark (Prof M Nordentoft M D); Institute for Biostatistics and Epidemiology Division, RTI International, Research Triangle Park, Durham, NC, USA (Prof O Mors PhD); Mount Sinai, New York, NY, USA (Prof M M Vanyukov PhD); and RTI International, Research Triangle Park, Durham, NC, USA (Prof M M Vanyukov PhD).

PRS for cannabis use disorder were computed using PRS-CS40 for each of the 66,915 genotyped individuals of European descent in BioVU (appendix pp 16–17). Genotyping and quality control of this sample have been described elsewhere.41 A logistic regression model was fitted to each of 1335 case or control phenotypes that had at least 100 cases to estimate the odds of each diagnosis given the PRS for cannabis use disorder, after adjustment for sex, median age of the longitudinal electronic health record measurements, and the top ten ancestral principal controls. To explore whether pleiotropic effects of the PRS for cannabis use disorder were mediated by smoking behaviours, we did two phenotype-wide association study (PheWAS) sensitivity analyses: a PheWAS on summary statistics of cannabis use disorder that had been conditioned on the top smoking initiation loci using mtCOJO,34 and a PheWAS using a diagnosis of tobacco use disorders as an additional covariate in the regression model, which is a conservative over-correction given the extremely high comorbidity expected between cannabis use disorder and tobacco use disorder. We used a Bonferroni-corrected phenome-wide significance threshold of 3.74×10⁻⁸; this is overly conservative because it incorrectly assumes independence between phenotypes. PheWAS analyses were run using the PheWAS R package, version 0.12.42

Data from the Adolescent Brain Cognitive Development Study (Registered; ABCD study)43 (data release 2.0.1) were used to test the association of PRS for cannabis use disorder with brain structure among 4539 cannabis-naïve children (through self-reporting or hair toxicology) of European ancestry (mean age 9.93 years [SD 0.63], 2125 [47%] were girls; appendix p 17). Total bilateral white matter volume, grey matter volume, and intracranial volume were estimated using FreeSurfer 5.3.43 PRS from the cannabis use disorder GWAS were generated at nine p value thresholds (ie, p<0·0001, p<0·001, p<0·01, p<0·10, p<0·20, p<0·30, p<0·40, p<0·50, and p<1·00), as were PRS for cannabis use disorder.1 Linear mixed models were used to include scanner (for imaging analyses) and family as nested random effects, done using the lme4 package in R, version 3.6.0. All analyses included as fixed effect covariates the first 20 ancestral principal components, age, sex, age by sex, parents combined income, caregiver education, genotyping batch, caregiver’s marital status, prenatal cannabis exposure before and after knowledge of pregnancy, and twin status. Multiple testing within each brain structure phenotype was accounted for by applying random field theory correction across p value thresholds, as this method directly models the overlap across the different PRS thresholds and corrects for the statistical dependence among them.44

Role of the funding source
The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full...
access to all of the data and the final responsibility to submit for publication.

**Results**

We identified two genome-wide significant loci in the transancestral meta-analysis of cannabis use disorder (African and European ancestries, 20916 cases, 363116 controls; appendix pp 17, 20). These loci were significant in the European ancestry meta-analysis but did not reach significance in the much smaller African ancestry analysis (17068 cases, 357219 controls vs 3848 cases, 5897 controls; table 2). The lead SNPs were rs7432724 on chromosome 8 (ptransancestral=2.64×10⁻⁹, pEUR=0.98×10⁻³, p=1.84×10⁻⁷, Z=5.97, ptransancestral=2.43×10⁻⁹) and rs7783012 on chromosome 7 (ptransancestral=3.28×10⁻⁷, pEUR=0.98×10⁻³, p=1.84×10⁻⁷, Z=5.97, ptransancestral=2.43×10⁻⁹).

**Table 2:** Association statistics for the lead genome-wide significant SNPs across each of the three primary samples (deCODE, iPSYCH, PGC) in the European ancestry and transancestral meta-analyses

| Chromosome | Position | SNP         | Effect allele | deCODE | deCODE p value | iPYSCH | iPYSCH p value | PGC EUR | PGC EUR p value | PGC EUR | PGC EUR Z score | EPG EUR | EPG EUR p value | Trans-ancestral | Trans-ancestral p value |
|------------|----------|-------------|---------------|---------|----------------|--------|---------------|---------|----------------|---------|----------------|---------|----------------|-----------------|----------------------|
| 7          | 114116881| rs7783012   | A             | 1.10    | (0.03)         | 2.58×10⁻⁷ | (0.03)        | 1.11    | 9.56×10⁻⁴      | 3.47    | 5.22×10⁻⁹      | 1.31    | 1.84×10⁻⁹      | 5.97    | 2.43×10⁻⁹          |
| 8          | 27432062 | rs4732724   | C             | 0.90    | (0.03)         | 3.03×10⁻⁴ | (0.03)        | 0.84    | 0.61×10⁻¹      | -1.91   | 0.056          | 0.89    | 6.46×10⁻⁹      | -5.95   | 2.64×10⁻⁹          |

comp=complete meta-analysis (including related individuals and summary statistic cohorts); EUR-European ancestry OR; OR=odds ratio. PGC=Psychiatric Genomics Consortium. SNP=single nucleotide polymorphism. unrel=unrelated genotyped meta-analysis. *Complete deCODE, iPSYCH, and PGC EUR meta-analysis (excluding related individuals and summary statistic cohorts in the PGC). †Transancestral meta-analysis with deCODE, iPYSCH, and PGC samples (including related individuals and summary statistic cohorts). ‡SNP was only present in half of the PGC samples.

**Figure 1:** Manhattan plot of the European ancestry-only genome-wide meta-analysis
The gene-wise association analysis of European ancestry summary statistics identified three significant genes (\( \alpha = 2.66 \times 10^{-9} \)): FOX P2 (\( p = 7.31 \times 10^{-9} \)), PDE4B (\( p = 6.66 \times 10^{-7} \)), and ENO4 (\( p = 3.51 \times 10^{-9} \); appendix p 17, 27). No pathways were significant (appendix p 18). Three genes, NAT6 (amygdala, cortex, frontal cortex), HYAL3 (both adipozyse, whole blood, cerebellum, frontal cortex, hippocampus, nucleus accumbens, and spinal cord), and IFRD2 (cerebellum) were significantly related to cannabis use disorder through genetically regulated gene expression (appendix pp 18, 28). Connecting SNPs to genes via chromatin interaction data revealed significant associations in adult brain tissue (ten genes), fetal brain tissues (12 genes), iPSC-derived astrocytes (11 genes), and iPSC-derived neurons (eight genes); these genes included HYAL3, ENO4, CHRNA2, and FOXP2 (appendix pp 18, 29).

The SNP-heritability (\( h^2_{SNP} \)) for cannabis use disorder was 0.067–0.121 (SE 0.006–0.011) on the liability scale, depending on the estimated population prevalence and \( h^2_{SNP}0.02 \) (SE 0.002) on the raw scale. Cannabis use disorder showed significant \( r_g \) with 16 of the 23 studied phenotypes, for which the strongest relationships were observed with smoking initiation (\( r_g 0.66, p = 3.20 \times 10^{-8} \)), Townsend Deprivation Index (a measure of regional poverty; \( r_g 0.58, p = 3.30 \times 10^{-7} \), educational attainment (\( r_g 0.39, p = 6.70 \times 10^{-5} \)), and age at which first child is born (\( r_g 0.49; p = 5.40 \times 10^{-6} \); figure 2, appendix p 18). Thus, increased risk of cannabis use disorder is genetically correlated with increased liability for smoking initiation, living in an area of high material poverty, having children at an early age, and low levels of educational attainment. Liability to cannabis use disorder was also positively genetically correlated with alcohol use, nicotine dependence, psychiatric disorders (eg, AD H D, schizophrenia, major depression), and body-mass index (BMI).

The \( r_g \) between cannabis use and cannabis use disorder was 0.50 (SE 0.05, \( p = 1.50 \times 10^{-7} \)). Of the eight genome-wide significant SNPs associated with cannabis use, only four had \( p < 0.05 \) in the meta-analysis of cannabis use disorder (modest sample overlap between the two studies: genetic covariance intercept 0.014 (SE 0.005). Conditioning the summary statistics of cannabis use disorder for loci associated with cannabis use neither substantially modified the effect sizes of the genome-wide significant loci (rs4732724, \( \beta = 0.11, SE 0.02, p = 8.25 \times 10^{-6} \); rs7783012, \( \beta = 0.10, SE 0.02, p = 2.62 \times 10^{-9} \)) nor identified additional novel loci (appendix p 18). The heritability of cannabis use disorder adjusted for cannabis use loci (using \( mtCOJO \)) was 0.095 (SE 0.01) on the liability scale (estimated population prevalence 8.5%).

The \( r_g \)s with cannabis use disorder and cannabis use were significantly different for 12 of the 22 traits compared (figure 2, appendix p 18). Cannabis use and cannabis use disorder were positively genetically correlated with liability to smoking initiation, schizophrenia, major depressive disorder, risk tolerance, and the Townsend Deprivation Index. Cannabis use was positively genetically correlated with educational achievement and later age at birth of first child, and negatively with BMI. In contrast, cannabis use disorder was genetically correlated with low education attainment, early age at birth of first child, and high BMI. Liability to cannabis use disorder was genetically correlated with nicotine dependence (\( r_g 0.48, p = 1.35 \times 10^{-9} \)), whereas the genetic correlation of this trait with cannabis use was not significant (\( p = 0.44 \)); conditioning the \( r_g \) of cannabis use disorder on cannabis use loci (with \( p < 0.001 \)) made little difference in the magnitude of the \( r_g \) (appendix p 18).

We found no evidence of genetically causal relationships between liability to cannabis use disorder and to any of the most highly correlated traits (ie, educational attainment, age at first birth, Townsend Deprivation...
Figure 3: PheWAS associations between polygenic risk for CUD and phenotypes in the BioVU biobank

The 46 phenotypes shown are significantly associated with CUD (p<3.74 x 10^-5, corrected for 1335 phenotypes tested). CUD=cannabis use disorder. PheWAS=phenotype-wide association study. NOS=not otherwise specified. SIRS=systemic inflammatory response syndrome.
The PRS for cannabis use disorder were significantly associated with reduced total white matter volume in cannabis-naïve children from the ABCD Study (standardised β = –0.05; p = 0.001), explaining up to 0.18% of the variance in white matter volume at the most predictive threshold of p < 0.5 (appendix p 18). Children in the highest quartile of PRS, on average, had a white matter volume that was 1% lower than those in the lowest quartile. Results remained significant when including intracranial volume as a covariate (standardised β = –0.08; p = 0.01) and when excluding 2482 (54%) of 4539 who used any substance (standardised β = –0.05; p = 0.001), or when excluding 2482 (54%) of 4539 who used any substance or were prenatally exposed to any substance (standardised β = –0.05; p = 0.03). The PRS for cannabis use were not significantly correlated with white matter volume (figure 3). After correction for multiple testing, there was no association between PRS for cannabis use disorder or cannabis use and grey matter volume (all p > 0.01; appendix p 18).

**Discussion**

This GWAS meta-analysis confirmed one previously identified locus on chromosome 8 as associated with cannabis use disorder and identified a new locus on chromosome 7. The lead variant at the chromosome 7 locus (rs7783012) is a cis-eQTL for FOXP2 expression in brain and adipose tissue. FOXP2 was also significantly implicated in gene-based tests that incorporated information about chromatin interactions in iPSC-derived astrocytes (appendix p 29). rs7783012 has also been associated with measures related to externalising behaviours (eg, ADHD, age at first sexual intercourse, generalised risk tolerance) and with educational attainment. 

FOXP2 is essential to synaptic plasticity and has been implicated in the normal development of speech and language acquisition. However, because of the prominence of the protein product of FOXP2 as a regulator of numerous genes, indirect pathways of vulnerability beyond risk-taking are also possible.

Individual SNPs on chromosome 8 are eQTLs for CHRNA2 and EPHX2, extending previous work by...
Cannabis use frequency in the UK Biobank was genetically correlated with cannabis use disorder as well, but, similarly to other psychiatric and behavioural traits, the PRS for cannabis use disorder explained only a small proportion of variance in cannabis use frequency (R² 0·04%).

Some previous cross-sectional studies have linked differences in grey matter volume with cannabis use and dependence, but, however, a large mega-analysis did not find reductions in regional or global volumes in cannabis-dependent adults compared with controls. In our study, the association between PRS for cannabis use disorder and white matter volume persisted in the subset of children who were not exposed to any substance, including prenatally. This finding suggests that polygenic liability for cannabis use disorder might index differences in white matter volume in the developing brain, independently of the onset of substance use behaviours. Still, the association between PRS for cannabis use disorder and white matter was small (R² 0·15–0·18%), and additional studies are needed to confirm this association.

Some limitations are noteworthy. Our African ancestry sample was under-powered; more data are needed, particularly in light of potential disparities that result from a majority of genetic studies focusing on European-ancestry populations. We had little or no information regarding comorbid psychiatric disorders for the majority of PGC samples; however, we did conditional analyses to account for these and it made little difference. Information regarding lifetime cannabis exposure and the potency of cannabis used was scarce. Our estimates of genome-wide SNP-h² were far lower than the h² estimated from twin and family studies (0·07–0·12 vs 0·5–0·7). This discrepancy between pedigree-estimated heritability and SNP-heritability is common across essentially all substance use disorders, and might be due to low power, some heritability residing in variants too rare to be included in our GWAS, and insufficient coverage of optimal common-variant genomic coverage in available microarray data even after imputation. An additional limitation is that we did not do formal Mendelian randomisation analysis. To do this analysis, we would have needed to remove sample overlap between our cannabis use disorder GWAS and the other GWASs of interest, which would have greatly decreased our statistical power. However, after doing latent causal variable analyses, an approach related to mendelian randomisation that can account for sample overlap among the input GWAS, there was no significant evidence of causal relationships between liability to cannabis use disorder and to any of the top genetically correlated traits: educational attainment, age at first birth, Townsend Deprivation Index, smoking initiation, or ADHD (appendix p 16). Overall, estimates of genetic overlap might also be sensitive to sample characteristics, such as older volunteers in the UK Biobank cohort and some younger registry-based cohorts in our GWAS.
addition, imbalance between cases and controls could have affected our findings, although we did not observe substantial genetic heterogeneity (appendix pp 23–24). In conclusion, our findings provide further evidence that cannabis use disorder is a serious, psychiatric illness with genetic and neurobiological influences that diverge at least partially from cannabis use. From a public health perspective, the recognition that cannabis use disorder is a serious form of psychopathology should spur efforts to identify and aid at-risk individuals in the face of escalating cannabis use worldwide, especially among adolescents.

Contributors

SAB, DD, AD, JHo, GW, MN, AA, RB, JHE, IBH, JG, ECJ, TET, and RKW designed the study. MJ, A, JBD, AMG, AMM, AMM, D, RL, BTW, RET, S-A, VB, BSM, AFE, KX, WEC, ADB, FWR, SEM, JG-B, LD, NGM, DBH, JHo, RCP, LW, MAK, DMD, JBR, MDR, MCS, MMV, BPR, LBJ, LF, JB, KKB, JG, NAG, DFG, KMH, SMH, ACH, JJKH, IBH, DMH, WGI, EJO, KJ, PAFM, HHH, MM, MM, GW, OM, PEM, EFP, JNP, BP, VR, NLS, RJS, JLS, TET, TT, SV, TLR, TW, MJW, and RW collected and interpreted the data. DAA, AMP, RR, ASH, DD, RD, RET, BSM, FFR, SEM, TBB, SEP, LHM, DBH, IJK, MAK, DEA, SZ, JA, GW, AA, RB, SS-S, RK-C, BWD, NAD, SDG, DFH, SMH, JHe, ECJ, EJO, RE, PAL, GW, CMT, RKW, and RW analysed the data. MJ, DAA, DL, RP, ASH, DD, S-A, DBH, DEA, GWR, AA, RB, SS-S, JDB, LKD, JHE, JG, DFG, Ho, JHe, ECJ, NLS, TET, RKW, and RW gave the statistical input. ASH, DFB, FWR, SEP, LJM, AA, RB, T-R, KCE, JHE, JG, ECJ, TET, and RKW wrote the manuscript. SAM, AMG, HRR, DL, LAF, RP, ASH, DD, RET, WEC, ADB, SEP, DBH, JHo, RLW, MAK, DMD, MCS, MVW, GW, AA, RB, JNM, AS-S, R, JDB, KKB, LKD, JHE, JG, ECJ, EJO, NLS, TET, RKW, and HZ edited the manuscript. AMM, KSK, SEM, LD, DMD, JLM, BFR, AA, RB, JHE, JG, NAG, ACH, EOJ, HHHM, ECN, and TET provided phenotype expertise. CJH, AMM, LAF, ADK, KSK, HNM, AA, RB, ALD, KKD, JHE, TMF, JG, KMH, KXH, WGI, HHM, ECRM, BP, KS, TSV, TW, and MJW supervised the study.

Declaration of interests

TW has acted as an advisor to Lundbeck A/S. IJB and the spouse of NLS are listed as inventors on Issued US Patent 8 080 371, Markers for Addiction, covering the use of particular SNPs in determining the diagnosis, prognosis, and treatment of addiction. AMM has received research support from Eli Lilly, Janssen, Pfizer, and the Sackler Foundation. HRK is a member of the American Society of Clinical Psychopharmacology’s Alcohol Clinical Trials Initiative, which was supported in the last 3 years by AbbVie, Ethypharm, Indivior, Lilly, Lundbeck, Otsuka, Pfizer, Arbor, and Amygdala Neurosciences. HRK and JG are named as inventors on PCT patent application number 15/878 640, entitled Genotype-guided dosing of opioid agonists, filed Jan 24, 2018. LD reports untied educational grant funding by the following grants: MH109532 (AA, HJE, JG); F32AA027435 (ECJ); K02DA032573 (AA); U101MH110934–01 (ADB); ST32DA007261 (AH); T32MH180895 (DAAB); NIH F32 MH112058 (FRW); Center for Genomics and Personalized Medicine and the Centre for Integrative Sequencing, ISF, Aarhus University, Denmark (ADB); the European Commission, Horizon 2020, grant number 667302 (ADB); R01-DA034076 (TET); R01-HD060726 (BWD, KMH, JDB, MMB); Health Research Council of New Zealand 16/600 (JB, JHo, MAF, JFP, JFF); Health Research Council of New Zealand 1/(2) (JB, JHo); AAAA27827, AG661622 (RB); U24DA004147 (SAB); R01HD099365 (ADB); R01DA032082 (WE); DA01015; DA038504 (RKP); K20 AA018755 (DMD); P01-HD001921; R01-HD073342 (KMH1); R01 DA008763; DA025899; DA024413; DA016977; H1M1; Wellcome Trust (00636/Z/14-Z, 210677/Z/19/Z), UKRI MRC (MC_PC_1709), MRC/MR/S5831831 (AMM); R01MH111384, The Brain & Behavior Research Foundation NARSAD grant 28632 (REP); R21 DA047527; R21 DC008098 (RKP); SAMHSA Grant # H79TR001668 (MRR); R01DA026911, R21DA046971 (NLS); R01HD050735, U54EB020403 (subaward), NHMRC Australia (496682; 1009604) (MJW); Tobacco-Related Disease Research Program of the University of California Grant Number T2KCRF265, Families for Borderline Personality Disorder Research (Beth and Rob Elliott) 2018 NARSAD Young Investigator Grant #26767 (SS-R). The IPSYCH team was supported by grants from the Lundbeck Foundation (R165–2013–15120, R102-A9118, R115–2014–270, and R248–2017–2003), the European Commission (Horizon 2020, grant number 667302), and the universities and university hospitals of Aarhus and Copenhagen, Denmark. The Danish National Biobank resource was supported by the Novo Nordisk Foundation. Data handling and analysis on the GenoMEKP HPC facility was provided by the Center for Genomics and Personalized Medicine and the Centre for Integrative Sequencing, ISF, Aarhus University, Denmark (grant to ADB). Work at deCODE genetics was partly funded by the NIH (R01–DA04076). The BioVU dataset(s) used for the PhENAS analyses described were obtained from Vanderbilt University Medical Center’s BioVU, which is supported by numerous sources: institutional funding, private agencies, and federal grants. These grants include the NIH funded Shared Instrumentation Grant 1SR0025141, and Clinical Translational Sciences Award Grant ULTTR000243, ULTR000445, and ULTR024795. Genomic data are also supported by investigator-led projects that include 1U1HG004798, 1R01NS032830, R2C2GM093638, P50GM113505, U101HG006378, U101HG006592, and R01HD074711, and additional funding sources https://www.vanderbilt.edu/ bioscience/funding/. Dr Degenhardt is funded by the NIH (R01 DA147480) and the Australian National Health and Medical Research Council Research (1135991). Data used in the preparation of this article were obtained from the Adolescent Brain Cognitive Development (ABCD) Study, held in the NIH Data Archive. The ABCD Study is supported by the NIH and additional federal partners under award numbers U01DA040122, U01DA040128, U01DA041048, U01DA041089, U01DA041106, U01DA041120, U01DA041334, U01DA04148, U01DA04156, U01DA04174, U24DA041123, U24DA041347, U01DA041093, and U01DA04125. A listing of participating sites and a complete listing of the study investigators can be found online. ABCD consortium investigators designed and implemented the study or provided data, or both, but did not necessarily participate in analysis or writing of this study. This manuscript reflects the views of the authors and might not reflect the opinions or views of the NIH or ABCD consortium investigators. The ABCD data repository grows and changes over time. The ABCD data used in this study came from the NIMH Data Archive 2.0.1. Funding support for the Comorbidity and Trauma Study (dbGAP accession number, phs002735.v1.p3) was provided by NIDA (R01 DA17305); GWAS genotyping services at the Center for Inherited Disease Research at The Johns Hopkins University were supported by NIH (contract N01-HG-64010). Funding support for the Center for Education and Drug Abuse Research (dbGAP accession number, phs001649.v1.x1) was provided by NIDA (R01 DA023065). The Christchurch Health and Development Study has been supported by funding from the Health Research Council of New Zealand, the National Child Health Foundation (Cure Kids), the Canterbury Medical Research Foundation, the New Zealand Lottery Grants Board, the University of Otago, the Carney Centre for
Pharmacogenomics, the James Hume Bequest Fund, US NIH grant MH077887, and NIDA grant, a developmental model of gene-environment interplay in substance processes (U01 DA013424; R01 DA024411) 2007–12. The Collaborative Study on the Genetics of Alcoholism (COGA; Principal Investigators B Porjesz, V Hesselbrock, and T Foroud; Scientific Director, A Agrawal; Translational Director, D Dick) includes 11 different centres: University of Connecticut (V Hesselbrock); Indiana University (H J Edenberg, T Foroud; J Nurnberger Jr, Y Liu); University of Iowa (S Kuperman, J Kranzler); SUNY Downstate (B Porjesz, J Meyers, C Kamarajan, A Pandey); Washington University in St Louis (L Bierut, J Rice, K Bucholz, A Agrawal); University of California at San Diego (M Schuckit); Rutgers University (J Tischfeld, A Brooks, R Hart); The Children’s Hospital of Philadelphia, University of Pennsylvania (L Almasy); Virginia Commonwealth University (D Dick; J Salvation); Icahn School of Medicine at Mount Sinai (A Gato, M Kapoor, P Slesinger); and Howard University (D Scott). Other COGA collaborators include L Bauer (University of Connecticut); L Wetherill, X Xuei, D Lai, S O’Connor, M Plawecki, S Lourens (Indiana University); L Aton (University of Iowa); G Chan (University of Iowa; University of Connecticut); D B Chorlian, J Zhang, S Kinreich, G Pandey (SUNY Downstate); M Chao (Icahn School of Medicine at Mount Sinai); A Anokhin, V McCutcheon, S Saconne (Washington University); F Aliev, P Barr (Virginia Commonwealth University); and H Chin, and A Parsian (NC Institute on Addictions Collaborators). We continue to be inspired by our memories of Henri Begleiter and Theodore Reich, founding Principal Investigator and Co-Principal Investigator of COGA, and also owe a debt of gratitude to other past organisers of COGA, including Ting-Kai Li, P Michael Conneally, Raymond Crowe, and Wendy Reich for their crucial contributions. This national collaborative study is supported by N01 DA000158 (U01 DA000158). Support for the Study of Addiction: Genetics and Environment (SAGE) was provided through the NIAAA, National Institute on Alcohol Abuse and Alcoholism (NIAAA) Staff Collaborators. We would like to acknowledge all participating groups of the International Cannabis Consor...
5 Pasman JA, Verweij JK/H, Gerring Z, et al. GWAS of lifetime cannabis use reveals new risk loci, genetic overlap with psychiatric traits, and a causal influence of schizophrenia. Nat Neurosci 2018; 21: 1611–70.
6 Gillespie NA, Neale MC, Jacobson K, Kendler KS. Modeling the genetic and environmental association between peer group deviance and cannabis use in male twins. Addiction 2009; 104: 420–29.
7 Walters RK, Polimanti R, Johnson EC, et al. Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. Nat Neurosci 2018; 21: 1656–69.
8 Kranzler HR, Zhou H, Kember RL, et al. Genome-wide association study of alcohol consumption and use disorder in 274,424 individuals from multiple populations. Nat Commun 2019; 10: 1099.
9 Pedersen CB, Byeberg-Graulholm J, Pedersen MG, et al. The iPSYCH2012 case–coloroquet sample: new directions for unravelling genetic and environmental architectures of severe mental disorders. Mol Psychiatry 2018; 23: 6–14.
10 Gulcher JR, Kristjansson K, Gudbjartsson H, Stefansson K. Protection of privacy by third-party encryption in genetic research in Iceland. Eur J Hum Genet 2000; 8: 739–42.
11 American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th edn. Arlington, VA: American Psychiatric Publishing, 2000.
12 Mors O, Perto GP, Mortensen PB. The Danish psychiatric central research register. Scand J Public Health 2011; 39: 54–57.
13 Lam M, Awasthi S, Watson HJ, et al. RICOPLI: Rapid Imputation for ConSortias PipeLine. Bioinformatics 2020; 36: 930–33.
14 Galinsky KJ, Bhatia G, Loh P-R, et al. Fast Principal-Component Analysis reveals convergent evolution of ADH1B in Europe and East Asia. Am J Hum Genet 2016; 98: 456–72.
15 Price AL, Patterson NJ, Plenge RM, et al. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006; 38: 904–09.
16 1000 Genomes Project Consortium. A global reference for human genetic variation. Nature 2015; 526: 68–74.
17 Delaneau O, Marchini J, Zagury J-F. A linear complexity phasing method for thousands of genomes. Nat Methods 2011; 9: 79–81.
18 Hancock DB, Levy JL, Gaddis NC, et al. Assessment of genotype imputation performance using 1000 Genomes in African American studies. PLoS One 2012; 7: e50610.
19 Gudbjartsson DF, Helgason H, Gudjonsson SA, et al. Large-scale whole-genome sequencing of the Icelandic population. Nat Genet 2015; 47: 435–44.
20 Demontis, D, Walters RK, Martin J, et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. Nat Neurosci 2019; 51: 63–75.
21 Price AL, Helgason A, Palsson S, et al. The impact of divergence time on the nature of population structure: an example from Iceland. PLoS Genet 2009; 5:e1000505.
22 Bulik-Sullivan BK, Loh P-H, Hilary K Finucane, et al. LD core regression distinguishes confounding from polygenic in genome-wide association studies. Nat Genet 2015; 47: 291–95.
23 Wilker CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010; 26: 2190–91.
24 Yang J, Ferreira T, Morris AP, et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat Genet 2012; 44: 569–73.
25 Watanabe K, Taskesen E, van Bochove A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun 2017; 8: 1826.
26 de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput Biol 2015; 11: e1004219.
27 Sey NHA, Hu B, Mali W, et al. A computational tool (H-MAGMA) for improved prediction of brain-disorder risk genes by incorporating brain chromatin interaction profiles. Nat Neurosci 2020; 23: 583–91.
28 Lamparter D, Marbach D, Rueedi R, Kutuklik Z, Bergmann S. Fast and rigorous computation of gene and pathway scores from SNP-based summary statistics. PLoS Comput Biol 2016; 12: e1004774.
29 Barbeira AN, Dickinson SP, Bonazzola R, et al. Exploring the phenotypic consequences of tissue-specific gene expression variation inferred from GWAS summary statistics. Nat Commun 2018; 9: 1825.
30 Kreuz DS, Axelrod J, Delta-9-tetrahydrocannabinol: localization in body fat. Science 1973; 179: 391–93.
31 GTEX Consortium. The Genotype-Tissue Expression (GTEX) project. Nat Genet 2013; 45: 380–85.
32 Bulik-Sullivan B, Finucane HK, Anttila V, et al. An atlas of genetic correlations across human diseases and traits. Nat Genet 2015; 47: 1236–46.
33 Stimson FS, Ruan WJ, Pickering R, Grant BF. Cannabis use disorders in the USA: prevalence, correlates and co-morbidity. Psychol Med 2006; 36: 1467–60.
34 O’Connor LJ, Price AL. Distinguishing genetic correlation from causation across 52 diseases and complex traits. Nat Genet 2018; 50: 1728–34.
35 Zhu Z, Zheng Z, Zhang F, et al. Causal associations between risk factors and common diseases inferred from GWAS summary data. Nat Genet 2018; 50: 224.
36 Liu M, Yang J, Wedow R, et al. Association studies of up to 1-2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. Nat Genet 2019; 51: 237–44.
37 Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 168 schizophrenia-associated genetic loci. Nature 2016; 541: 421–27.
38 Choi SW, O’Reilly PF. PRSice-2: polygenic risk score software for biobank-scale data. GigaScience 2019; 8: giz082.
39 Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 2005; 102: 15545–50.
40 Ge T, Chen C-Y, Ni Y, Feng Y-CA, Smoller JW. Polygenic prediction via Bayesian regression and continuous shrinkage priors. Nat Commun 2019; 10: 1–10.
41 Dennis J, Sealock J, Levinson RT, et al. Genetic risk for major depressive disorder and loneliness in sex-specific associations with coronary artery disease. Mol Psychiatry 2019; published online Dec. 3. https://doi.org/10.1038/s41380-019-0614-y.
42 Carroll RJ, Bastarache L, Denny JC. PRSice: data analysis and plotting tools for phenotype-wide association studies. Bioinformatics 2011; 27, 1728–34.
43 Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 168 schizophrenia-associated genetic loci. Nature 2016; 541: 421–27.
44 Nichols TE. Multiple testing corrections, nonparametric methods, and random field theory. Neuroimage 2012; 62: 774–81.
45 Morris R, Carstairs V. Which deprivation? A comparison of selected deprivation indexes. J Public Health Med 1991; 13: 318–26.
46 Lee JJ, Wedow R, Ayuso Ocakay A, et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1-1 million individuals. Nat Genet 2018; 50: 1112–21.
47 Hancock DB, Guo Y, Reginsson GW, et al. Genome-wide association study across European and African American ancestries identifies a SNP in DNM1L1 contributing to nicotine dependence. Mol Psychiatry 2018; 23: 1911–19.
48 Wray NR, Ripke S, Matheisen M, et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. Nat Genet 2018; 50: 668–81.
49 Spielhofer EK, Wilker CJ, Berndt SI, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat Genet 2010; 42: 937–48.
50 Watanabe K, Stringer S, Frei O, et al. A global overview of pleiotropy and genetic architecture in complex traits. Nat Genet 2019; 51: 1339–48.
51 Linnér RK, Birolli P, Kong E, et al. Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. Nat Genet 2019; 51: 245–57.
52 Lai CSL, Gerrelli D, Monaco AP, Fisher SE, Copp AJ. FOXP2 expression during brain development coincides with adult sites of pathology in a severe speech and language disorder. Brain 2003; 126: 2455–62.
53 Pardinas AF, Holmans P, Pocklington AJ, et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. Nat Genet 2018; 50: 381–89.
55 Fowler IL, Carr VJ, Carter NT, Lewin TJ. Patterns of current and lifetime substance use in schizophrenia. Schizophr Bull 1998; 24: 443–55.
56 McDougall DR, Watson JE, Abdeen AA, et al. Anti-inflammatory ω-3 endocannabinoid epoxides. Proc Natl Acad Sci USA 2017; 114: E6034–43.
57 Snider NT, Nast JA, Tesmer LA, Helenberg PF. A cytochrome P450-derived epoxygenated metabolite of anandamide is a potent cannabinoid receptor 2-selective agonist. Mol Pharmacol 2009; 75: 965–72.
58 Wagner K, Inceoglu B, Hammock BD. Soluble epoxide hydrolase inhibition, epoxygenated fatty acids and nociception. Prostaglandins Other Lipid Mediat 2011; 96: 76–83.
59 Dudbridge F, Allen RJ, Sheehan NA, et al. Adjustment for index event bias in genome-wide association studies of subsequent events. Nat Commun 2019; 10: 1561.
60 Lewis CM, Vassos E. Polygenic risk scores: from research tools to clinical instruments. Genome Med 2020; 12: 44.
61 Battistella G, Fornari E, Annoni J-M, et al. Long-term effects of cannabis on brain structure. Neuropsychopharmacology 2014; 39: 2041–48.
62 Mackey S, Allgaier N, Chaaran B, et al. Mega-analysis of gray matter volume in substance dependence: general and substance-specific regional effects. Am J Psychiatry 2019; 176: 119–28.
63 Peterson RE, Kuchenhaecker K, Walters RK, et al. Genome-wide association studies in ancestrally diverse populations: opportunities, methods, pitfalls, and recommendations. Cell 2019; 179: 589–603.
64 Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. Nat Genet 2019; 51: 584–93.
65 Smith GD, Ebrahim S. Mendelian randomization: prospects, potentials, and limitations. Int J Epidemiol 2004; 33: 30–42.
66 Fry A, Littlejohns TJ, Sudlow C, et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. Am J Epidemiol 2017; 186: 1026–34.