1	Genome-wide association studies of lifetime and frequency cannabis use in 131,895
2	individuals
3	
4	Hayley H A Thorpe, PhD ¹ , Pierre Fontanillas, PhD ² , John J Meredith, BSc ³ , Mariela V Jennings,
5	BSc ³ , Renata B Cupertino, PhD ³ , Shreya Pakala, BSc ³ , 23andMe Research Team ² , Sarah L
6	Elson, PhD ² , Jibran Y Khokhar, PhD ¹ , Lea K Davis ^{4,5,6} , PhD, Emma C Johnson, PhD ⁷ , Abraham
7	A Palmer, PhD ^{3,8} , & Sandra Sanchez-Roige, PhD ^{3,5,8*}
8	
9	¹ Department of Anatomy and Cell Biology, Schulich School of Medicine and Dentistry, Western
10	University, London, ON, Canada.
11	² 23andMe, Inc., Sunnyvale, CA, USA
12	³ Department of Psychiatry, University of California San Diego, La Jolla, CA, USA
13	⁴ Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN,
14	USA
15	⁵ Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA
16	⁶ Department of Psychiatry and Behavioral Sciences, Vanderbilt University Medical Center,
17	Nashville, TN, USA
18	⁷ Department of Psychiatry, Washington University School of Medicine, St. Louis, MO USA
19	⁸ Institute for Genomic Medicine, University of California San Diego, La Jolla, CA, USA
20	Corresponding Author:
21	Sandra Sanchez-Roige, PhD
22	Department of Psychiatry, University of California San Diego
23	9500 Gilman Drive, La Jolla, CA, USA, 92093
24	Phone: +1-619-874-1426
25	Email: sanchezroige@ucsd.edu

26 Summary

27 Cannabis is one of the most widely used drugs globally. Decriminalization of cannabis is further 28 increasing cannabis consumption. We performed genome-wide association studies (GWASs) of 29 lifetime (N=131,895) and frequency (N=73,374) of cannabis use. Lifetime cannabis use GWAS 30 identified two loci, one near CADM2 (rs11922956, p=2.40E-11) and another near GRM3 31 (rs12673181, p=6.90E-09). Frequency of use GWAS identified one locus near CADM2 32 (rs4856591, p=8.10E-09; $r^2=0.76$ with rs11922956). Both traits were heritable and genetically 33 correlated with previous GWASs of lifetime use and cannabis use disorder (CUD), as well as 34 other substance use and cognitive traits. Polygenic scores (PGSs) for lifetime and frequency of 35 cannabis use associated cannabis use phenotypes in AllofUs participants. Phenome-wide 36 association study of lifetime cannabis use PGS in a hospital cohort replicated associations with 37 substance use and mood disorders, and uncovered associations with celiac and infectious diseases. This work demonstrates the value of GWASs of CUD transition risk factors. 38

- 39 Keywords: genome-wide association study, cannabis, addiction, genetic correlations, polygenic
- 40 score, phenome-wide association study, CADM2, GRM3

41 Introduction

42 Approximately 209 million people globally reported using cannabis in 2020¹. The number 43 of people who use cannabis regularly is expected to increase as cannabis is decriminalized in 44 many jurisdictions²⁻⁴. While people report using cannabis for medicinal purposes⁵, there is 45 increasing evidence that cannabis use has short- and long-term adverse consequences across psychiatric, cognitive, and physical health⁶⁻¹⁴. Up to 27% of those who use cannabis in their 46 47 lifetime are estimated to develop cannabis use disorder (CUD)¹⁵, in which cannabis use becomes problematic to an individual's intra- and interpersonal wellbeing¹⁶. However, it is currently unclear 48 49 what factors contribute most to the development of CUD, and thus of clinical interest to identify what makes an individual vulnerable to cannabis use and its negative effects. 50

Problematic cannabis use is estimated to be 51-78% heritable based on twin studies¹⁷⁻¹⁹ 51 52 and recent genome-wide association studies (GWASs) have implicated hundreds of loci 53 associated with CUD²⁰⁻²³. While CUD GWASs are of paramount importance, they come with three 54 major caveats. First, these studies only examine one extreme of the addiction spectrum and 55 neglect other substance-related behaviors and stages between substance initiation to substance use disorder (SUD) diagnosis (e.g., recreational use, escalating intake, dependence)²⁴. These 56 pre-addiction phenotypes²⁵ are thought to dictate an individual's progression to SUD²⁶⁻³² and are 57 heritable^{17,26,31,33}. However, aside from GWASs of lifetime cannabis use (having ever versus never 58 59 used cannabis in one's lifetime) using data from the International Cannabis Consortium (ICC) and other sources^{34,35}, which represents the opposing end of the addiction spectrum to CUD, the 60 genetics of other pre-addiction cannabis traits are understudied^{36,37}. Second, only a portion of 61 those engaging in frequent cannabis use seek treatment or have a CUD diagnosis^{38,39}. It is 62 therefore unlikely that CUD GWASs and downstream analyses fully characterize the genetics of 63 64 regular, potentially harmful cannabis use and its relationships with physical and mental health. 65 Third, curating case/control SUD GWASs are costly and laborious because they require individual

psychological assessments for both cases and controls. Pre-addiction phenotypes can be rapidly
 and inexpensively collected in large population-based cohorts via self-report questionnaires⁴⁰.

68 We collected data from 23andMe, Inc. research participants by asking if they had ever 69 used cannabis (N=131,895). Those who responded yes were asked a follow-up question about 70 the number of days they used cannabis in their heaviest use period (N=73,374) as a measure of 71 cannabis use frequency. We performed GWASs of lifetime and frequency of cannabis use, 72 followed by a battery of secondary analyses to compare biological, genetic, and phenotypic 73 associations. Because the frequency of cannabis use phenotype better distinguished between 74 light and heavy use, we hypothesized that the genetics of frequency of cannabis use would more 75 closely resemble CUD compared to lifetime cannabis use genetics.

77 Results

GWASs of Lifetime Cannabis Use and Frequency of Cannabis Use Uncover Associations with
 CADM2 and GRM3

Participant demographics are described in **Supplementary Table 1**. The cohort was 65.2% female with a mean age of 52.8±0.04 years old. Participant responses to surveys about lifetime and frequency of cannabis use are available in **Supplementary Table 2** and **Supplementary Fig. 1**.

For single nucleotide polymorphisms (**SNPs**) quality control, see **Supplementary Table 3.** Genomic control inflation factors for lifetime cannabis use (λ =1.08) and frequency of cannabis use (λ =1.03) suggested no substantial inflation due to population stratification for either GWAS. SNP-based heritability (*h2_{SNP}*) was 12.88%±0.97 for lifetime cannabis use, greater than the *h2_{SNP}* for lifetime cannabis use from the ICC (*h2_{SNP}*=6.63%±0.43)³⁴. *h2_{SNP}* for frequency of cannabis use was 4.12%±0.72 (**Supplementary Table 4**).

90 We identified two genome-wide significant (p<5.00E-08) loci for lifetime cannabis use on 91 chromosomes 3 and 7 (Fig 1A, Supplementary Fig. 2-3, Supplementary Table 5). The most 92 significant association was with rs11922956 (p=2.40E-11, chr3p12.1) located upstream the Cell adhesion molecule 2 gene (CADM2), replicating findings from previous lifetime use³⁴ and CUD^{22,23} 93 94 GWASs. CADM2 encodes a glycoprotein primarily expressed in the brain with functions in cell-95 cell adhesion, synaptic formation, excitatory neurotransmission, and energy homeostasis ^{41,42}. We 96 also found a novel association between lifetime cannabis use and rs12673181 (p=6.90E-09, 97 chr7q21.11), which is a SNP upstream of Metabotropic glutamate receptor 3 gene (GRM3) 98 encoding mGlu₃. mGlu₃ is an inhibitory group II receptor affecting a range of intracellular signaling cascades and cellular processes like glutamate neurotransmission and long-term plasticity 43. 99

- 100 Frequency of cannabis use GWAS identified one significant association with rs4856591
- 101 (p=8.10E-09, chr3p12.1; Fig1B, Supplementary Fig. 2, 5), which is near to CADM2 and is in
- 102 linkage disequilibrium (**LD**) with rs11922956 (r^2 =0.76, p<1.00E-04).



Figure 1. Manhattan plots of A) lifetime cannabis use (*N*=131,895) and B) frequency of cannabis
 use (*N*=73,374). The horizontal line represents the significance threshold (*p*=5.00E-08). Nearest
 protein-coding genes (<1Mb) to significant loci (red dots) are labelled. For quantile-quantile plots
 and locus zoom plots, see Supplementary Fig. 2-4.

108 Secondary Analysis Identifies 40 Lifetime and 4 Frequency of Cannabis Use Genes

109 Mapping SNPs to genes via gene-based (i.e., MAGMA, H-MAGMA) and transcriptome-110 wide association study (TWAS; i.e., S-PrediXcan) analyses identified 40 unique genes associated 111 with lifetime cannabis use (Supplementary Tables 6-8), and 4 unique genes associated with 112 frequency of cannabis use (Supplementary Tables 9). None of the 4 genes associated with 113 frequency of cannabis use (i.e., MMS22L, DSCC1, CPSF7, RP11-51J9.6) were implicated in 114 lifetime cannabis use. The only gene to overlap across gene-based and TWAS analyses was 115 CADM2 (Supplementary Table 10). Of the 44 unique genes associated with lifetime and 116 frequency of cannabis use, 29 gene associations have not been previously associated with any 117 cannabis-related trait (**Supplementary Table 10**).

118 Gene-set and tissue-based enrichment analyses yielded no significant results 119 (**Supplementary Tables 11-12**).

Lifetime and Frequency of Cannabis Use Are Genetically Correlated with Psychiatric, Cognitive,
and Physical Health Traits

122 There were 115 traits genetically correlated (r_a) with lifetime cannabis use and 38 with 123 frequency of cannabis use after applying a 5% false discovery rate (FDR) correction (Fig. 2-3, 124 Supplementary Table 13). We identified 29 traits that were significantly genetically correlated 125 with both lifetime and frequency of cannabis use (10 anthropomorphic traits; 19 psychiatric traits), 126 which were usually consistent in their direction of effect, with exceptions for intelligence and 127 executive function (positively genetically correlated with frequency of use, negatively genetically 128 correlated with lifetime use), and tense/'highly strung' and delay discounting (negatively 129 genetically correlated with frequency of use, positively genetically correlated with lifetime use), as 130 we review below (Supplementary Fig. 5).

131 *Cannabis and Other Substance Use Traits.* The genetic correlation between lifetime and 132 frequency of cannabis use was moderate (r_g =0.54±0.08, p=1.89E-10), suggesting imperfect

133 genetic overlap between the two traits. We identified positive genetic correlations between CUD 134 and lifetime (r_0 =0.62±0.04, p=2.44E-59), as well as frequency of cannabis use (r_0 =0.45±0.07, 135 p=2.45E-10; Fig. 2). Compared to lifetime cannabis use from the ICC, our lifetime cannabis use 136 trait was more strongly genetically correlated with CUD (23andMe-CUD $r_0=0.62\pm0.04$, p=2.44E-137 59 vs. ICC-CUD r_q =0.48±0.04, p=4.30E-33). Positive genetic correlations with other aspects of 138 substance use (e.g., drug experimentation and lifetime cannabis use: $r_{o}=0.97\pm0.01$, p<1.35E-161; 139 frequency: $r_q=0.54\pm0.07$, p=5.45E-14) and misuse (e.g., Alcohol Use Disorder Identification Test 140 (AUDIT) problems and lifetime cannabis use: $r_a=0.46\pm0.06$, p=1.54E-15; frequency of cannabis 141 use: $r_q=0.30\pm0.10$, p=2.46E-03) were among the top genetic correlations for lifetime and 142 frequency of cannabis use (Fig. 3, Supplementary Table 13).



143

Figure 2. SNP-based heritability and genetic correlation analysis comparisons across cannabis-related traits. A) Genetic correlations and $h_{2\text{SNP}}$ across 23andMe lifetime cannabis use and frequency of cannabis use with ICC lifetime cannabis use ³⁴ and CUD from Levey *et al.* ²². $h_{2\text{SNP}}$ ±standard error shown in matrix diagonal (gray boxes), r_{g} ±standard error in off-diagonal (white boxes). Correlation coefficients shown in heatmap color, with *p* value underneath in black. **B)** CUD requires progression through multiple pre-addiction stages, including experimental use, regular use, compulsive/harmful use, dependence, cessation attempts, and relapse. Aside from

- 149 lifetime cannabis use as a proxy for experimental use and frequency of cannabis use as a proxy for regular use, which positively
- 150 genetically correlate with CUD, most of these stages have not been genetically explored with GWAS.

151 Psychiatric Disorders. Lifetime cannabis use was genetically correlated with 152 schizophrenia ($r_a=0.15\pm0.03$, p=7.33E-07); however, frequency of cannabis use was not 153 $(r_{o}=0.02\pm0.05, p=0.73)$. We also identified associations with other psychiatric traits and lifetime 154 cannabis use like ADHD ($r_{a}=0.31\pm0.05$, p=5.20E-12), depression ($r_{a}=0.22\pm0.04$, p=3.52E-10), 155 and cross-disorder ($r_0=0.30\pm0.05$, p=3.91E-10). We identified significant genetic correlations 156 between frequency of cannabis use and the psychiatric-related traits "depression possibly related 157 to stressful or traumatic events" (r_{q} =-0.54±0.16, p=9.22E-04), stress-related disorder 158 $(r_{o}=0.33\pm0.10, p=1.44E-03)$, and anxiety/panic attacks $(r_{o}=-0.38\pm0.14, p=6.06E-03)$, though only 159 stress-related disorder was also genetically correlated with lifetime cannabis use ($r_{q}=0.25\pm0.06$, 160 *p*=1.42E-04).

161 *Externalizing and Risk-Taking Traits.* Among the strongest associations for lifetime 162 cannabis use were positive genetic correlations with externalizing behavior (r_g =0.84±0.03, 163 p=5.65E-208), and two traits that were used to construct externalizing behavior⁴⁴: number of 164 sexual partners (r_g =0.69±0.03, p=6.16E-115) and age at first sex (reverse-coded; r_g =0.60±0.03, 165 p=1.08E-83). We found similar positive genetic correlations with frequency of cannabis use and 166 externalizing and risk-taking (externalizing: r_g =0.45±0.06, p=1.68E-15); number of sexual 167 partners: r_g =0.42±0.06, p=3.17E-12).

168 *Cognitive Traits.* We identified significant genetic correlations between lifetime cannabis 169 use and 11 cognitive and executive function-related traits; these included positive genetic 170 correlations with delay discounting (r_g =0.16±0.04, p=3.51E-04) and other impulsivity-related 171 measures (r_g =0.27±0.05 to 0.46±0.05, p=1.02E-22 to 3.20E-04), and negative genetic 172 correlations with childhood intelligence (r_g =-0.29±0.08, p=3.20E-04), educational years (r_g =-173 0.17±0.03, p=1.84E-07), common executive function (r_g =-0.13±0.03, p=3.63E-05), and 174 intelligence (r_g =-0.12±0.03, p=3.04E-05).

For frequency of cannabis use, we identified positive genetic correlations with intelligence (r_g =0.40±0.05, p=4.18E-14) and common executive function (r_g =0.34±0.06, p=7.86E-09). There was also a negative genetic correlation with delay discounting (r_g =-0.23±0.07, p=1.62E-03), indicating those who use cannabis more frequently may devalue delayed rewards. Consistent with lifetime cannabis use, we found positive genetic correlation with the impulsivity-related measure perseverance (r_g =0.28±0.09, p=1.48E-03).

Physical Health Traits. We identified genetic correlations between lifetime cannabis use and 17 physical health traits, including chronic pain (r_g =0.21±0.04, p=5.59E-09), back pain (r_g =0.22±0.05, p=2.19E-06), and coronary artery disease with angina (r_g =0.17±0.04, p=2.59E-05). For frequency of cannabis use, there was a positive genetic correlation with diabetes (r_g =0.20±0.07, p=5.96E-03) and a negative genetic correlation with irritable bowel syndrome (r_g =-0.27±0.10, p=6.55E-03).



Figure 3. Comparison of genetic correlations across anthropometric (light gray), health (medium gray), and psychiatric (dark gray) traits between lifetime cannabis use (lanes 1 and 2) and frequency of cannabis use (lanes 3 and 4). Lanes 1 and 3 show r_g values calculated by LDSC, and lanes 2 and 4 show FDR-corrected *p* values. Only traits for which at least one cohort was FDR-significant are displayed. For a full list of correlations and trait names, see **Supplementary Table 13**. *reverse coded traits.

194 Lifetime and Frequency of Cannabis Use Polygenic Scores Associate with Cannabis Use 195 Phenotypes

196 Lifetime and frequency PGS associations with cannabis use traits in All of Us (AoU) were 197 considered in single (i.e., models only incorporating lifetime or frequency of cannabis use PGS 198 as variables) and joint (i.e., models incorporating lifetime and frequency of cannabis use PGS as 199 variables) PGS models (Supplementary Tables 14-16). In the joint-PGS model simultaneously 200 accounting for lifetime and frequency PGS in the European cohort, based on genetic similarity 201 (see **Methods**), lifetime cannabis use PGS associated with lifetime cannabis use (β =0.19±0.01, 202 p < 2.00E-16), daily cannabis use ($\beta = 0.09 \pm 0.03$, p = 5.09E-04), and problematic cannabis use 203 (β =0.22±0.02, p<2.00E-16; **Table 1, Supplementary Table 16**). Frequency of cannabis use PGS 204 was associated with lifetime cannabis use (β =0.06±0.01, p<2.00E-16), and nominally associated 205 with problematic cannabis use (β =0.06±0.03, p=0.01), which did not survive multiple testing correction. Lifetime and frequency PGSs were estimated to explain 0.31-1.52% of the phenotypic 206 207 variance in cannabis use traits (Fig. 4). In the African cohort, based on genetic similarity (see 208 **Methods**), lifetime cannabis use was predicted by the lifetime PGS (β =0.08±0.01, p=2.76E-12) 209 and the frequency PGS (β =0.04±0.01, p=1.88E-04), which contributed an estimated 0.20% to 210 phenotypic variance. In both populations, phenotypic variance was primarily attributable to the 211 lifetime cannabis use PGS versus the frequency of cannabis use PGS.

In all models, age was a significant negative predictor and being a male was a significant
 positive predictor of problematic, daily, and lifetime cannabis use (Supplementary Tables 14 17).

Table 1. Joint-PGS regression analysis associating lifetime cannabis use PGS, frequency of cannabis use PGS, and select covariates
 with lifetime, daily, and problematic cannabis use in AoU cohorts. Bold PGS results are significant following Bonferroni correction
 (*p*<8.33E-03). For full analysis variables, see Supplementary Table 16.

		European Cohort							African Cohort									
	Lifetime Use (<i>N</i> _{case} =64,711, <i>N</i> _{control} =49,595)			Lifetime Use Daily Use (N _{case} =64,711, (N _{case} =1,411, N _{control} =49,595) N _{control} =28,112)		Problematic Use (<i>N</i> _{case} =1,825, <i>N</i> _{control} =118,704)		Lifetime Use (<i>N</i> _{case} =26,064, <i>N</i> _{control} =21,610)		Daily Use (<i>N</i> _{case} =2,483, <i>N</i> _{control} =11,718)		Problematic Use (<i>N</i> _{case} =2,315, <i>N</i> _{control} =50,262)						
Variable	β	StdErr	р	β	StdErr	p	β	StdErr	p	β	StdErr	р	β	StdErr	p	β	StdErr	р
Lifetime PGS	0.19	0.01	<2.00E-16	0.09	0.03	5.09E-04	0.22	0.02	<2.00E-16	0.08	0.01	2.76E-12	-0.02	0.03	0.52	3.62E-03	0.02	0.88
Frequency PGS	0.06	0.01	<2.00E-16	-0.03	0.03	0.38	0.06	0.03	0.01	0.04	0.01	1.88E-04	0.03	0.03	0.23	0.01	0.03	0.61



Figure 4. Percent proportion of lifetime, daily, and problematic cannabis use variance attributable to lifetime cannabis use PGS, frequency of cannabis use PGS, or both (joint-PGS) in European and African AoU cohorts. Bonferroni-corrected significance of PGS contribution for single- and joint-PGS models (see **Table 1, Supplementary Tables 15-16**) shown above data label in its corresponding legend color (n.s. p>0.05, *p<0.05, *p<0.01, **p<0.001).

Lifetime Cannabis Use Polygenic Score Associates with Psychiatric and Infectious Disease
 Diagnoses

225 Our phenome- and laboratory-wide association studies (PheWAS/LabWAS) uncovered 226 15 FDR-significant PheWAS associations and 9 FDR-significant LabWAS associations with 227 lifetime cannabis use in the BioVU European cohort, as described below (Fig. 5; Supplementary 228 Tables 19-20). When CUD was included as a covariate, 8 PheWAS and 6 LabWAS associations 229 remained, as we discuss below. Tobacco smoking is prevalent among cannabis users⁴⁵; 4 230 PheWAS and 4 LabWAS associations persisted when adjusting for tobacco use disorder (**TUD**), 231 and 1 PheWAS and 5 LabWAS associations persisted when CUD and TUD were jointly included 232 as covariates, described further below. We found no significant associations with cannabis use 233 frequency in the European cohort, nor any significant associations for lifetime or frequency of 234 cannabis use in the African cohort.

235 Psychiatric Disorders. Our PheWAS identified positive associations between lifetime 236 cannabis use PGS and seven psychiatric disorders: TUD (β =0.09±0.01, p=2.44E-15), substance 237 addiction and disorders (β =0.14±0.02, p=8.56E-13), CUD (β =0.21±0.03, p=1.24E-10), alcohol-238 related disorders (β =0.10±0.02, p=2.43E-05), mood disorder (β =0.05±0.01, p=3.38E-07), two 239 anxiety traits (anxiety disorders: β =0.05±0.01, p=8.85E-06; anxiety disorder: β =0.04±0.01, 240 p=2.55E-04), depression ($\beta=0.05\pm0.01$, p=1.73E-05), bipolar ($\beta=0.09\pm0.02$, p=1.59E-04), and 241 suicide ideation or attempt (β =0.12±0.03, p=2.64E-04). TUD, substance addiction and disorders, 242 and mood disorders persisted following adjustment for CUD, only substance addiction and 243 disorders persisted following control for TUD, and no psychiatric disorders were significant 244 following control for both CUD and TUD. We did not find evidence of an association with 245 schizophrenia (β =0.02±0.06, p=0.68), schizophrenia and other psychotic disorders (β =0.03±0.03, 246 p=0.29), or psychosis ($\beta=0.08\pm0.04$, p=0.07).

Infectious Diseases. We found significant positive associations between lifetime cannabis use and infectious diseases, including human immunodeficiency virus (**HIV**) disease (β =0.21±0.04, *p*=1.14E-07), symptomatic HIV infection (β =0.21±0.04, *p*=1.26E-07), viral hepatitis C (β =0.13±0.03, *p*=3.99E-06), and viral hepatitis (β =0.11±0.03, *p*=6.17E-06). All associations persisted following control for CUD, both HIV associations persisted following control for TUD, but no infectious disease associations persisted following control for both CUD and TUD.

253 *Other Diagnoses.* Lifetime cannabis use PGS was negatively associated with one 254 digestive trait, celiac disease (β =-0.34±0.05, *p*=1.55E-11). This association persisted with 255 following control for CUD, TUD, and combined CUD and TUD.

Blood Laboratory Biomarkers. LabWAS revealed associations with lifetime cannabis use across four blood biomarkers: mean corpuscular hemoglobin (**MCH**; β =0.02±3.53E-03, *p*=1.60E-07), carbon dioxide serum/plasma (β =-0.02±3.47E-03, *p*=1.92E-06), MCH concentration (β =0.02±3.85E-03, *p*=9.41E-05), and mean corpuscular volume (β =0.01±3.53E-03, *p*=7.77E-04). Following CUD adjustment, all but mean corpuscular volume remained significant; following adjustment for TUD alone or alongside CUD, carbon dioxide serum/plasma and MCH remained significant.

263 *Immune Laboratory Biomarkers.* Two immune biomarkers, leukocytes in blood 264 (β =0.02±3.51E-03SE, *p*=2.778E-09) and complement C4 in serum or plasma (β =0.06±0.02, 265 *p*=6.84E-05), were positively associated with lifetime cannabis use. Both remained significant 266 following control with TUD and CUD independently or together.

267 Other Laboratory Biomarkers. The kidney biomarker creatinine in blood (β =-0.02±3.90E-268 03, *p*=1.02E-04), endocrine biomarker parathyrin intact in serum or plasma (β =-0.04±0.01, 269 *p*=1.25E-03), and the metabolic biomarker calcitriol in serum and plasma (β =-0.02±0.01, 270 *p*=1.37E-03) were negatively associated with lifetime cannabis use; none were significant

- following control for TUD, but creatinine in blood remained significant when CUD, and when CUD
- and TUD were used as covariates.



Figure 5. Forest plot of FDR-significant phenome associations with lifetime cannabis use PGS unconditioned (UC), or with adjustment for cannabis use disorder (CUD), tobacco use disorder (TUD), or both (CUD-TUD). A) PheWAS results. B) LabWAS results. For full trait information, see Supplementary Tables 19-20.

277 Discussion

278 This study contributes to the growing body of cannabis use genetics literature by providing 279 new GWASs of 131,895 individuals of European genetic similarity assessed for lifetime cannabis 280 use and, for the first time, 73,374 individuals assessed for frequency of cannabis use. Both 281 GWASs replicated the robust associations with variants nearby CADM2 and lifetime cannabis use 282 GWAS identified one novel locus near GRM3. We found that lifetime and frequency of cannabis 283 use reliably genetically correlated with substance use-related traits, including CUD, and PGSs for 284 both traits associated with cannabis use phenotypes in AoU. Polygenic analysis of lifetime 285 cannabis use also revealed positive associations with substance use and mood disorders consistent with the literature, and novel phenotypic associations with anxiety disorders, infectious 286 287 diseases, and red blood cell biomarkers. Overall, these results support the value of cannabis use 288 phenotypes spanning the addiction spectrum in the exploration of genetic factors influencing 289 cannabis use vulnerability and health risk.

290 Pre-addiction phenotypes are intended to capture prodromal symptoms of SUD; lifetime 291 and frequency of cannabis use are heritable risk factors for CUD development^{15,27,29,30,32,34}. They 292 can be easily self-reported in large cohorts, making them attractive targets for GWAS. Lifetime 293 cannabis use captures both experimental/occasional and heavy use; despite the simplicity of this 294 phenotype, we uncovered multiple novel genetic associations with lifetime cannabis use (i.e., 295 GRM3 locus, genetic correlations, polygenic associations), and found it reliably associated with 296 CUD and multiple other important traits. Although frequency of use may better account for regular 297 cannabis use, this trait did not associate with CUD to a greater degree compared to lifetime 298 cannabis use ($r_q=0.45\pm0.07$ vs. 0.62±0.04), potentially due to lower power (N=73,374 vs. 299 131,895). However, lifetime and frequency of cannabis use was genetically correlated with each 300 other and their associations with other complex traits were almost always directionally consistent. 301 This included positive genetic correlations between lifetime and frequency of use with other

302 substance use, misuse, and behavioral traits thought to be substance use risk factors like 303 externalizing, impulsivity, and risk-taking^{21,44,46-50}, consistent with ICC lifetime cannabis use genetic correlations³⁴ and reports of a general "addiction risk factor" or externalizing factor 304 305 accounting for genetic overlap across substances^{23,44,51-53}. We previously demonstrated that consumption and problematic use phenotypes (i.e., alcohol^{24,54,55}, tobacco⁵⁶) are correlated but 306 307 non-identical traits; this is likely true for cannabis. Future multivariate analyses incorporating 308 lifetime, frequency, and other cannabis use GWASs (e.g., CUD, dependence, craving, etc.) could 309 effectively boost locus discovery, identify novel relationships between CUD behaviors and health, 310 and parse genomic factors pertaining to the stages of CUD³⁶, as we and others have previously demonstrated for other substance use traits^{23,54,55,57-59}. 311

312 One of our most notable findings was a novel association between lifetime cannabis use 313 and rs12673181, which is located upstream of the GRM3 gene that encodes the group II inhibitory 314 glutamate receptor mGlu₃. There are no known associations with this or other *GRM3* SNPs with 315 cannabis-related traits, and while GWASs implicate GRM3 variants in other substance use (i.e., 316 alcohol, smoking)⁶⁰, schizophrenia⁶¹⁻⁶⁵, neuroticism^{66,67}, educational attainment⁶⁸, and other 317 phenotypes⁶⁹⁻⁷¹, these variants are not in LD with rs12673181. Recent studies also suggest that 318 mGlu₃ potentiates activity of mGlu₅⁷², which has also garnered attention for its potential role in 319 addictive-like behaviors and endocannabinoid synthesis^{73,74}. While rs12673181 lies upstream of 320 GRM3, it is not a known expression quantitative trait locus (eQTL) of GRM3 (Supplementary 321 **Table 5**)⁷⁵. Further functional work, especially pertaining to the regulation of *GRM3*, is required to 322 characterize its association with cannabis use vulnerability.

Through multiple lines of evidence, we found lifetime and frequency of cannabis use associated with the *CADM2* gene, replicating prior GWASs of lifetime cannabis use and CUD^{23,34}. Other GWASs have found an association between SNPs in *CADM2* and other substance use traits^{23,48,51,60,76-91}, risk-taking^{76,85,89,92-94}, impulsivity⁴⁸, and externalizing behaviors⁴⁴.

327 Supporting the genetic correlation observed across cannabis GWAS data, PGSs for 328 lifetime and, to a lesser degree, frequency of cannabis use, associated with phenotypes across 329 the CUD progression spectrum (i.e., lifetime, daily, and problematic use). More variance was 330 explained by lifetime (0.29-1.40%) rather than frequency of use PGS (0.12-0.19%), and together 331 they explained up to 1.6% of phenotypic variance. This is on par with recent substance use PGS 332 analyses⁹⁵⁻⁹⁹, including by Hodgson et al. ¹⁰⁰, who estimated that ICC lifetime cannabis use PGS 333 predicted 0.82% of variance in lifetime cannabis use and 1.2% of variance in continued cannabis 334 use in UK Biobank participants. Although it is improbable that cannabis use PGS alone will be sufficient for clinical utility¹⁰¹, lifetime and frequency of cannabis use PGS could be useful for 335 models predicting problematic cannabis use risk. 336

337 Largely consistent with the genetic correlations we observed. PheWAS uncovered positive 338 associations between lifetime cannabis use PGS with substance use, depression, anxiety, bipolar. and suicidality in the BioVU cohort (N<66,917). To our knowledge, the positive 339 340 associations with HIV and hepatitis diagnoses, negative association with celiac disease, and 341 mixed associations with multiple blood and immune laboratory biomarkers are novel. Our findings 342 complement a recent PheWAS conducted in the Yale-Penn sample (N<10,610), which is a cohort 343 deeply phenotyped for psychiatric disorder diagnoses and related diagnostic criteria. That study 344 found ICC lifetime cannabis use PGS positively associated with CUD, as well as traits related to other substance use (e.g., alcohol, tobacco, sedatives, stimulants) and depression¹⁰². That many 345 346 of these relationships disappear when controlling for CUD in our PheWAS and in the Kember et al. ¹⁰² study, as well as when controlling for TUD in our study, supports the hypothesis that these 347 348 associations are mediated by regular cannabis and tobacco use rather than genetic liability for lifetime cannabis use. Furthermore, like others¹⁰², we found minimal evidence of a relationship 349 350 between lifetime cannabis use genetics, schizophrenia, and psychosis (aside from bipolar), 351 despite the genetic relationship between cannabis use and psychosis being the subject of intense

interest¹⁰³⁻¹⁰⁶following observations of their apparent bidirectional phenotypic relationship¹⁰⁷.
Epidemiological evidence supports a link between cannabis and heavy or high potency cannabis
use¹⁰⁸⁻¹¹⁰. Identifying more robust variant associations, especially for frequency of cannabis use,
will aid future causal inference analyses that can resolve the role of cannabis genetics in health.

Our results were consistent with the prior cannabis use GWAS by ICC³⁴. Lifetime cannabis 356 357 use measured in US-based 23andMe research participants was genetically correlated with the 358 same trait examined in the ICC cohort, which is composed of participants across North America, 359 Europe, and Australia³⁴. Both lifetime cannabis use datasets were genetically correlated with 360 CUD, but the magnitude of this association was stronger in the 23andMe dataset compared to 361 ICC ($r_q=0.62$ vs. 0.48) despite our smaller sample size. Heritability estimates for our lifetime 362 cannabis use trait was also higher (12.88% vs. 6.63%). Heritability may decrease when meta-363 analyzing cohorts, possibly due to cohort-specific environmental/geocultural differences that could exist surrounding cannabis use¹¹¹⁻¹¹³. Furthermore, while we found consistent positive 364 correlations with psychiatric disorders, including schizophrenia^{21,34,49}, ADHD^{21,34,47,49}, bipolar 365 disorder^{34,49}, and depression^{21,49} between 23andMe and ICC lifetime cannabis use, we also 366 367 observed that the genetic correlation with educational attainment was negative with 23 and Me and positive with ICC lifetime cannabis use³⁴. Interestingly, while most genetic correlations between 368 369 lifetime and frequency of cannabis use were also mostly in agreement, lifetime cannabis use 370 negatively genetically correlated with intelligence and common executive function and positively 371 genetically correlated with delay discounting, while we saw the inverse with frequency of use. This 372 is not entirely unprecedented, as relationships between cannabis use and cognitive traits can be 373 paradoxical, especially among those with psychiatric disorders, such as those with psychosis who 374 use cannabis exhibiting greater cognitive abilities than those who do not¹¹⁴. In sum, although most 375 associations were consistent, the differences we observed in trait heritability and patterns of 376 genetic correlations suggest some disunity between 23andMe and ICC lifetime cannabis use

377 cohorts, as well as lifetime and frequency of cannabis use data, which will warrant careful378 consideration before attempting to meta-analyzing GWAS data.

379 There are several limitations to our study. The legal status of cannabis use differs across 380 countries and even US states and has been changing over the last several decades. Thus, for 381 some of our older subjects, both lifetime and frequency of use could be reflecting use decades 382 ago, whereas others are referencing more recent use. Most studies suggest that legalizing 383 recreational cannabis use increases lifetime and frequency of use rates¹¹⁵, which may have 384 impacted our findings in complex ways that depend on which location a given participant was in 385 at the time of their use. In addition, frequency of cannabis determined by the number of use days 386 over a 30-day window does not accurately reflect lifetime use intensity because it does not 387 account for the duration of regular use or use quantity. These characteristics are important to 388 CUD trajectory and other health and wellbeing relationships¹¹⁶⁻¹¹⁹. Lifetime and frequency of 389 cannabis use GWASs also relied on self-reported data. Cannabis use is most common during 390 adolescence and young adulthood¹²⁰, but participants in this study averaged in their 50s and could 391 have been at greater risk for recall bias regarding cannabis use in early life¹²¹. Socioeconomic 392 variables are also associated with cannabis use rates^{122,123}, and the on-average higher 393 socioeconomic status of 23 and Me research participants may have influenced our findings³⁶. 394 Finally, GWASs were conducted using genomic information from individuals of genetically 395 predicted European ancestry. While we extended our polygenic analyses to African cohorts, 396 cross-population transferability of PGS is suboptimal compared to investigations where discovery 397 and target populations are ancestrally aligned^{124,125}. This, along with lower sample numbers, may 398 explain why we observed fewer associations in African versus European cohorts. Due to sample 399 size constraints, we also did not explore associations in other ancestral groups, further limiting 400 the generalizability of our results.

This project showcases the utility of pre-addiction phenotypes in cannabis use genomic discovery. Lifetime and frequency of cannabis use genetically associated with CUD and other SUD, alongside concerning health and psychiatric problems. Increasing sample size and investigating other heritable, diverse phenotypes (e.g., drug responsivity, craving, withdrawal; **Figure 2B**) will be integral to further our understanding of CUD vulnerability and the health consequences of cannabis use.

408 Acknowledgements

409 We would like to thank the research participants and employees of 23andMe for making 410 this work possible. Participants provided informed consent and volunteered to participate in the 411 research online, under a protocol approved by the external AAHRPP-accredited Institutional 412 Review Board (IRB), Ethical & Independent (E&I) Review Services. As of 2022, E&I Review 413 Services is part of Salus IRB (https://www.versiticlinicaltrials.org/salusirb). The following members 414 of the 23andMe Research Team contributed to this study: Stella Aslibekyan, Adam Auton, 415 Elizabeth Babalola, Robert K. Bell, Jessica Bielenberg, Katarzyna Bryc, Emily Bullis, Daniella 416 Coker, Gabriel Cuellar Partida, Devika Dhamija, Sayantan Das, Teresa Filshtein, Kipper Fletez-417 Brant, Will Freyman, Karl Heilbron, Pooja M. Gandhi, Karl Heilbron, Barry Hicks, David A. Hinds, 418 Ethan M. Jewett, Yunxuan Jiang, Katelyn Kukar, Keng-Han Lin, Maya Lowe, Jey C. McCreight, 419 Matthew H. McIntyre, Steven J. Micheletti, Meghan E. Moreno, Joanna L. Mountain, Priyanka 420 Nandakumar, Elizabeth S. Noblin, Jared O'Connell, Aaron A. Petrakovitz, G. David Poznik, 421 Morgan Schumacher, Anjali J. Shastri, Janie F. Shelton, Jingchunzi Shi, Suyash Shringarpure, 422 Vinh Tran, Joyce Y. Tung, Xin Wang, Wei Wang, Catherine H. Weldon, Peter Wilton, Alejandro 423 Hernandez, Corinna Wong, Christophe Toukam Tchakouté.

424 We would also like to thank The Externalizing Consortium for sharing the GWAS summary 425 statistics of externalizing. The Externalizing Consortium: Principal Investigators: Danielle M. Dick, 426 Philipp Koellinger, K. Paige Harden, Abraham A. Palmer. Lead Analysts: Richard Karlsson Linnér, Travis T. Mallard, Peter B. Barr, Sandra Sanchez-Roige. Significant Contributors: Irwin D. 427 428 Waldman. The Externalizing Consortium has been supported by the National Institute on Alcohol 429 Abuse and Alcoholism (R01AA015416 -administrative supplement), and the National Institute on 430 Drug Abuse (R01DA050721). Additional funding for investigator effort has been provided by 431 K02AA018755, U10AA008401, P50AA022537, as well as a European Research Council Consolidator Grant (647648 EdGe to Koellinger). The content is solely the responsibility of the 432

authors and does not necessarily represent the official views of the above funding bodies. The
Externalizing Consortium would like to thank the following groups for making the research
possible: 23andMe, Add Health, Vanderbilt University Medical Center's BioVU, Collaborative
Study on the Genetics of Alcoholism (COGA), the Psychiatric Genomics Consortium's Substance
Use Disorders working group, UK10K Consortium, UK Biobank, and Philadelphia
Neurodevelopmental Cohort.

439 We gratefully acknowledge All of Us participants for their contributions, without whom this 440 research would not have been possible. We also thank the National Institutes of Health's All of 441 Us Research Program for making available the participant data examined in this study. The All of 442 Us Research Program is supported by the National Institutes of Health, Office of the Director: 443 Regional Medical Centers: 1 OT2 OD026549; 1 OT2 OD026554; 1 OT2 OD026557; 1 OT2 444 OD026556; 1 OT2 OD026550; 1 OT2 OD 026552; 1 OT2 OD026553; 1 OT2 OD026548; 1 OT2 445 OD026551; 1 OT2 OD026555; IAA #: AOD 16037; Federally Qualified Health Centers: HHSN 446 263201600085U; Data and Research Center: 5 U2C OD023196; Biobank: 1 U24 OD023121; The 447 Participant Center: U24 OD023176; Participant Technology Systems Center: 1 U24 OD023163; 448 Communications and Engagement: 3 OT2 OD023205; 3 OT2 OD023206; and Community 449 Partners: 1 OT2 OD025277; 3 OT2 OD025315; 1 OT2 OD025337; 1 OT2 OD025276.

450 Data Availability

We will provide 23andMe summary statistics for the top 10,000 SNPs upon publication. 23andMe GWAS summary statistics will be made available through 23andMe to qualified researchers under an agreement with 23andMe that protects the privacy of the 23andMe participants. Please visit <u>https://research.23andme.com/collaborate/#dataset-access/</u> for more information and to apply to access the data.

456 We will share the Jupyter notebooks used for PGS analysis in AoU with registered All of 457 Us researchers upon request.

458 Author Contributions

459 SSR and AAP conceived the idea. PF and SLE contributed formal analyses and curation 460 of 23andMe data. HHAT contributed to formal analyses, investigation, and data visualization. 461 contributed to formal data analysis and data visualization. JJM, MVJ, RBC, and SP contributed to 462 formal analyses. HHAT and SSR wrote the manuscript. All authors reviewed and edited the 463 manuscript.

464 Funding Sources

SR is funded through the National Institute on Drug Abuse (NIDA DP1DA054394) and the
Tobacco-Related Disease Research Program (T32IR5226). HHAT is funded by a Canadian
Institute of Health Research (CIHR) Postdoctoral Fellowship (#491556). JYK is funded by a CIHR
Canada Research Chair in Translational Neuropsychopharmacology. PF, the 23andMe Research

469 Team, and SLE are employed by 23andMe, Inc. AAP is supported by NIDA (P50DA037844).

470 **Declaration of Interests**

471 PF, the 23andMe Research Team, and S.L.E. are employed by and hold stock or stock
472 options in 23andMe, Inc. The remaining authors have nothing to disclose.

474 <u>Methods</u>

475 Participants and GWASs

476 Lifetime and frequency of cannabis use GWASs were conducted in male and female 23andMe 477 research participants of European genetic similarity, as previously described⁴⁸. Ancestry falls 478 along a spectrum^{126,127}; individuals were only included in the analysis if they had >97% European 479 genetic similarity (see Supplementary Methods), as determined through local ancestry 480 analysis¹²⁸. Participants provided informed consent and volunteered to participate in research 481 online under a protocol approved by the external AAHRPP-accredited Institutional Review Board 482 (IRB), Ethical & Independent (E&I) Review Services. As of 2022, E&I Review Services is part of 483 Salus IRB (https://www.versiticlinicaltrials.org/salusirb). During 4 months in 2015 and 14 months 484 between 2018 to 2020, participants completed a questionnaire surveying a range of personal and 485 behavior characteristics. Included in this survey were questions on lifetime substance use and 486 substance use frequency. Specifically, "Yes" or "No" responses to the question "Have you ever in 487 your life used marijuana?" were collected as a measure of lifetime cannabis use. If participants 488 answered "Yes", they were prompted to answer the question "How many days did you use 489 marijuana during your heaviest 30 days?" as a measure of frequency of cannabis use. 490 Participants could respond between 0 and 30 days.

491 For lifetime cannabis use and frequency of cannabis use, 23andMe conducted GWASs of 492 up to 33,419,581 imputed genetic variants using linear regression and assuming an additive 493 genetic model. Samples were genotyped on one of five genotyping platforms. The V1 and V2 494 platforms were variants of the Illumina HumanHap550 + BeadChip, including about 25,000 495 custom SNPs selected by 23andMe, with a total of ~560,000 SNPs. The V3 platform was based 496 on the Illumina OmniExpress + BeadChip, with custom content to improve the overlap with our V2 497 array, with a total of ~950,000 SNPs. The V4 platform is a fully custom array, including a lower 498 redundancy subset of V2 and V3 SNPs with additional coverage of lower-frequency coding

499 variation, and ~570.000 SNPs. The v5 platform, in current use, is an Illumina Infinium Global Screening Array (~640,000 variants) supplemented with ~50,000 variants of custom content. 500 501 Samples that failed to reach 98.5% call rate were excluded from the study. We excluded SNPs of 502 low genotyping quality, including those that failed a Mendelian transmission test in trios or with 503 large allele frequency discrepancies compared to European 1000 Genomes reference data, failed 504 Hardy-Weinberg Equilibrium testing, failed batch effects testing, or had a call rate <90%, as well 505 as SNPs with a minor allele frequency<0.1% and imputed variants with low imputation quality or 506 with evidence of batch effects (Supplementary Table 3). Model covariates included age, sex, 507 the first 5 genetic principal components (PCs), and indicator variables for genotype platforms (see 508 Supplementary Methods for additional details). Unrelated participants categorized as of 509 European ancestry were included in the GWASs (lifetime cannabis use N=131.895; frequency of 510 cannabis use N=73,374; Durand et al., 2014). For full details on genotyping and GWASs, see

511 Supplementary Methods.

512 Functional annotation and gene-Based Analyses

513 Functional annotation. Using the web-based platform Functional Mapping and Annotation 514 of Genome-Wide Association Studies (FUMA v1.3.8), SNPs were annotated based on ANNOVAR 515 categories, Combined Annotation Dependent Depletion scores, RegulomeDB scores, eQTLs, 516 and chromatin state predicted by ChromHMM. Novel SNPs were identified as those not in LD 517 $(r^2 < 0.10)$ or within ±1Mb of GWAS-significant SNPs uncovered by other GWASs of cannabis use 518 traits (e.g., initiation, age of onset, CUD) sourced from the literature^{20-23,34,35,51,129-133} and from the 519 EBI GWAS Catalog (https://www.ebi.ac.uk/gwas/). Novel genes were identified as those not identified by gene-based analyses in other cannabis-related studies^{22,34,35,103,134-138} or with 520 521 start/stop positions within ±1Mb of previously uncovered GWAS-significant SNPs.

522 <u>MAGMA gene-based and pathway analyses</u>. We used Multi-marker Analysis of GenoMic 523 Annotation (**MAGMA**, v1.08, Ensembl build v92), which is included in FUMA, to annotate SNPs

524 to protein-coding genes. LD was estimated using the 1000 Genomes European reference sample, 525 and significance determined following Bonferroni correction (p < 2.53E-06). Gene-set analysis was 526 conducted on 10.678 gene-sets and Gene Ontology terms curated from the Molecular Signatures 527 Database (MsigDB v7.0). Tissue-specific gene expression profiles were assessed in 54 tissue 528 types and 30 general tissue types with average gene expression in each tissue used as a 529 covariate. Using Genome-Tissue Expression (GTEx, v8) RNA-seq data, gene expression values 530 were log₂ transformed from the average Reads Per Kilobase Million (max value=50) per tissue. 531 Significance was determined following Bonferroni correction (p < 9.26E-04 for 54 tissue types; 532 p<1.67E-03 for 30 general tissue types).

533 <u>*H-MAGMA*</u>. We incorporated lifetime and frequency of cannabis use GWAS data with 534 chromatin interaction profiles from human brain tissue using Hi-C coupled MAGMA (**H-MAGMA**; 535 Sey et al., 2020). H-MAGMA assigns non-coding SNPs to genes based on chromatin interactions 536 from fetal brain, adult brain, midbrain neurons, cortical neurons, iPSC-derived neurons, and iPSC-537 derived astrocytes datasets (<u>https://github.com/thewonlab/H-MAGMA</u>). Exonic and promoter 538 SNPs were assigned to genes based on physical position¹³⁹. We applied a Bonferroni correction 539 based on the total number of gene-tissue pairs tested (p<9.42E-07 to 9.45E-07).

540 S-PrediXcan. We performed a transcriptome-wide association study using S-PrediXcan 541 (v0.7.5) to identify eQTL-linked genes associated with lifetime and frequency of cannabis use¹⁴⁰. 542 S-PrediXcan uses genetic information to predict gene expression levels in various tissues and 543 tests if eQTLs correlate with lifetime or frequency of cannabis use across 49 bodily tissues 544 (N_{aenes}=1,619 to 9,949). S-PrediXcan uses precomputed tissue weights from the GTEx project 545 database (https://www.gtexportal.org/) as the reference transcriptome dataset via Elastic net 546 models. As input data, we included summary statistics, transcriptome tissue data, and covariance 547 matrices of the SNPs within each gene model (HapMap SNP set available at the PredictDB Data

548 Repository)¹⁴⁰ from all available tissues. We applied Bonferroni correction for each tissue type 549 (p<3.09E-05 to 5.03E-06).

LDSC heritability and genetic correlations across health, psychiatric, and anthropomorphic traits
Linkage Disequilibrium Score regression (LDSC; https://github.com/bulik/ldsc) was used
to calculate h2_{SNP} and genetic correlations¹⁴¹. h2_{SNP} was calculated from pre-computed LD scores
("eur_w_ld_chr/"). r_g were calculated between lifetime or frequency of cannabis use with 292 other
traits across 22 health, psychiatric, and lifestyle categories (Supplementary Methods). We
applied a 5% FDR correction to account for multiple testing.

556 Polygenic score analyses

PGS of lifetime, daily, and problematic cannabis use in AoU. We tested the associations 557 558 between lifetime or frequency of cannabis use PGSs with cannabis traits available for AoU 559 participants clustering within a European or African genetic ancestry panel (for details, see All of Us Research Program Genomics Investigators ¹⁴²). AoU is a diverse health database currently 560 561 including survey responses, physical measurements, genotyping data, and electronic health records (EHR) for over 400,000 individuals living in the United States^{142,143}. Using survey and 562 563 EHR data, participants were assigned binary identifiers for lifetime cannabis use (concept id: 564 1585636), daily cannabis use among those who reported cannabis use in their lifetime (concept 565 id: 1585650), and problematic cannabis use (concept ids: 434327, 440387, 440996, 433452, 566 437838, 4323639, 4103419, 435231, 434019, 434328; Supplementary Methods).

We calculated PGSs in male or female participants who had available short-read whole genome sequencing data and applicable cannabis use data. We used the Allele Count/Allele Frequency (**ACAF**) threshold SNP callset curated by AoU, which includes SNPs of MAF > 1% or allele counts over 100 for each ancestral subpopulation. Using PRS-CS "auto" v1.1.0, the SNP set was filtered to biallelic SNPs present in the HapMap3 European ancestry set and SNPs were weighted. Lifetime and frequency of cannabis use PGSs were created from 782,975 weighted

573 SNPs using the allelic-scoring function, score, in PLINK (v1.9; Ge et al., 2019). The base R 574 function *glm* was used to fit logistic regression models for each cannabis use trait using PGS(s), 575 as well as the additional covariates of age, sex, and the first 10 global PCs provided by AoU. 576 Models included single PGS models (lifetime or frequency PGS + additional covariates), a joint-577 PGS model (lifetime PGS + frequency PGS + additional covariates), and a null model (additional 578 covariates only). For the joint-PGS model, Bonferroni correction was applied for two tests (lifetime 579 PGS and frequency PGS) and three outcomes (lifetime, daily, and problematic cannabis use) for 580 a total of N=6 comparisons (p<8.33E-03); single PGS models were corrected for one test and three outcomes (N=3, p<1.67E-02). Joint-PGS liability scale R^2 values were calculated as 581 previously described by Lee et al. ¹⁴⁵ using the NagelkerkeR2 function in the R package fmsb 582 (v0.7.6) and the estimated prevalence of cannabis use traits in US adults (Supplementary 583 584 **Methods**). PGS ΔR^2 was calculated by subtracting R^2 calculated from models including PGS from the R^2 of the null model. 585

586 Phenome- and Laboratory-wide association analyses in a hospital cohort (BioVU). We 587 tested associations between lifetime or frequency of cannabis use PGSs and medical condition 588 liability from hospital-based cohorts using data from the Vanderbilt University Medical Center 589 (**VUMC**; IRB #160302, #172020, #190418)¹⁴⁶. The BioVU cohort, a subset of VUMC biobank 590 participants (N=72,821), provided genotyping data and EHR containing clinical data and laboratory-assessed biomarkers^{144,146,147}. For each unrelated European (*N*=66,917) and African 591 592 (N=12,383) BioVU participant based on genetic similarity, we computed lifetime and frequency of 593 cannabis use PGSs using the PRS-CS v1.1.0¹⁴⁴.

594 For PheWAS, we fitted a logistic regression model to each case/control disease 595 phenotypes ("phecodes") to estimate the log odds of each diagnosis given lifetime cannabis 596 use/frequency of cannabis use PGS, while adjusting for sex, median age of the longitudinal EHR, 597 and the first 10 PCs with the PheWAS v0.12 R package¹⁴⁴. At least two International Disease

598 Classification (**ICD**) codes mapping to a PheWAS disease category (Phecode Map 1.2; 599 <u>https://phewascatalog.org/phecodes</u>) and a minimum of 100 cases were required for phecode 600 inclusion. We also conducted additional sensitivity analyses using TUD (phecode 318) and CUD 601 (see **Supplementary Table 12** for CUD ICD codes) as covariates to examine if SUD mediated 602 associations with cannabis PGSs. We calculated the 5% FDR for all associations performed 603 (*N*=1,405).

604 For LabWAS, we implemented the pipeline established by Dennis et al. 147. LabWAS 605 associates PGS with laboratory biomarkers (i.e., measurements) evaluated in BioVU participants. 606 LabWAS uses the median, inverse normal quantile transformed age-adjusted values from the 607 QualityLab pipeline in a linear regression to determine the association between lifetime or 608 frequency of cannabis use PGSs with 314 phenotypes. We controlled for the same covariates as 609 for the PheWAS analyses, excluding median age because the pipeline corrects for age using 610 cubic splines with 4 knots. We applied 5% FDR correction across all LabWAS associations 611 performed (N=314).

612

All results are presented as the mean±standard error unless otherwise specified.

614 **References**

- 615 1. United Nations Office on Drug and Crime (2022). Booklet 3 Drug market trends of616 Cannabis and Opioids.
- 2. Pacek, L.R., Mauro, P.M., and Martins, S.S. (2015). Perceived risk of regular cannabis
- 618 use in the United States from 2002 to 2012: differences by sex, age, and race/ethnicity.
- 619 Drug Alcohol Depend *149*, 232-244. 10.1016/j.drugalcdep.2015.02.009.
- 620 3. Statistics Canada (2013). Mental and substance use disorders in Canada.
- 4. Rotermann, M. (2021). Looking back from 2020, how cannabis use and related
- behaviours changed in Canada. Health Rep 32, 3-14. 10.25318/82-003-
- 623 x202100400001-eng.
- 5. Whiting, P.F., Wolff, R.F., Deshpande, S., Di Nisio, M., Duffy, S., Hernandez, A.V.,
- 625 Keurentjes, J.C., Lang, S., Misso, K., Ryder, S., et al. (2015). Cannabinoids for Medical
- 626 Use: A Systematic Review and Meta-analysis. JAMA *313*, 2456-2473.
- 627 10.1001/jama.2015.6358.
- 628 6. Bellocchio, L., Inchingolo, A.D., Inchingolo, A.M., Lorusso, F., Malcangi, G., Santacroce,
- 629 L., Scarano, A., Bordea, I.R., Hazballa, D., D'Oria, M.T., et al. (2021). Cannabinoids
- 630 Drugs and Oral Health-From Recreational Side-Effects to Medicinal Purposes: A
- 631 Systematic Review. Int J Mol Sci 22. 10.3390/ijms22158329.
- 632 7. Diep, C., Tian, C., Vachhani, K., Won, C., Wijeysundera, D.N., Clarke, H., Singh, M., and
- 633 Ladha, K.S. (2022). Recent cannabis use and nightly sleep duration in adults: a
- 634 population analysis of the NHANES from 2005 to 2018. Reg Anesth Pain Med 47, 100-
- 635 104. 10.1136/rapm-2021-103161.
- 8. Hasin, D.S., Kerridge, B.T., Saha, T.D., Huang, B., Pickering, R., Smith, S.M., Jung, J.,
 Zhang, H., and Grant, B.F. (2016). Prevalence and Correlates of DSM-5 Cannabis Use

638	Disorder, 2012-2013	Findings from the National E	pidemiologic Surve	on Alcohol and
-----	---------------------	------------------------------	--------------------	----------------

- 639 Related Conditions-III. Am J Psychiatry *173*, 588-599. 10.1176/appi.ajp.2015.15070907.
- 640 9. Hayley, A.C., Stough, C., and Downey, L.A. (2017). DSM-5 cannabis use disorder,
- 641 substance use and DSM-5 specific substance-use disorders: Evaluating comorbidity in a
- 642 population-based sample. Eur Neuropsychopharmacol *27*, 732-743.
- 643 10.1016/j.euroneuro.2017.06.004.
- Howard, J., and Osborne, J. (2020). Cannabis and work: Need for more research. Am J
 Ind Med *63*, 963-972. 10.1002/ajim.23170.
- 11. Keen, L., 2nd, Turner, A.D., George, L., and Lawrence, K. (2022). Cannabis use
- 647 disorder severity and sleep quality among undergraduates attending a Historically Black
- 648 University. Addict Behav *134*, 107414. 10.1016/j.addbeh.2022.107414.
- 12. Lo, J.O., Hedges, J.C., and Girardi, G. (2022). Impact of cannabinoids on pregnancy,
- reproductive health, and offspring outcomes. Am J Obstet Gynecol 227, 571-581.
- 651 10.1016/j.ajog.2022.05.056.
- 13. Pacek, L.R., Herrmann, E.S., Smith, M.T., and Vandrey, R. (2017). Sleep continuity,

architecture and quality among treatment-seeking cannabis users: An in-home,

- unattended polysomnographic study. Exp Clin Psychopharmacol 25, 295-302.
- 655 10.1037/pha0000126.
- 14. Page, R.L., 2nd, Allen, L.A., Kloner, R.A., Carriker, C.R., Martel, C., Morris, A.A., Piano,
- 657 M.R., Rana, J.S., Saucedo, J.F., American Heart Association Clinical Pharmacology, C.,
- 658 et al. (2020). Medical Marijuana, Recreational Cannabis, and Cardiovascular Health: A
- 659 Scientific Statement From the American Heart Association. Circulation *14*2, e131-e152.
- 660 10.1161/CIR.00000000000883.

- 15. Feingold, D., Livne, O., Rehm, J., and Lev-Ran, S. (2020). Probability and correlates of
- 662 transition from cannabis use to DSM-5 cannabis use disorder: Results from a large-scale
- 663 nationally representative study. Drug Alcohol Rev 39, 142-151. 10.1111/dar.13031.
- 16. American Psychiatric Association (2013). Substance-Related and Addictive Disorders. In
- 665 Diagnostic and Statistical Manual of Mental Disorders, (https://doi-
- 666 org.subzero.lib.uoguelph.ca/10.1176/appi.books.9780890425596.dsm16).
- 17. Hines, L.A., Morley, K.I., Rijsdijk, F., Strang, J., Agrawal, A., Nelson, E.C., Statham, D.,
- 668 Martin, N.G., and Lynskey, M.T. (2018). Overlap of heritable influences between
- 669 cannabis use disorder, frequency of use and opportunity to use cannabis: trivariate twin
- 670 modelling and implications for genetic design. Psychol Med *48*, 2786-2793.
- 671 10.1017/S0033291718000478.
- 18. Kendler, K.S., Ohlsson, H., Maes, H.H., Sundquist, K., Lichtenstein, P., and Sundquist,
- 573 J. (2015). A population-based Swedish Twin and Sibling Study of cannabis, stimulant

and sedative abuse in men. Drug Alcohol Depend *149*, 49-54.

- 675 10.1016/j.drugalcdep.2015.01.016.
- 19. Verweij, K.J., Zietsch, B.P., Lynskey, M.T., Medland, S.E., Neale, M.C., Martin, N.G.,
- Boomsma, D.I., and Vink, J.M. (2010). Genetic and environmental influences on
- 678 cannabis use initiation and problematic use: a meta-analysis of twin studies. Addiction
- 679 *105*, 417-430. 10.1111/j.1360-0443.2009.02831.x.
- 680 20. Demontis, D., Rajagopal, V.M., Thorgeirsson, T.E., Als, T.D., Grove, J., Leppala, K.,
- 681 Gudbjartsson, D.F., Pallesen, J., Hjorthoj, C., Reginsson, G.W., et al. (2019). Genome-
- 682 wide association study implicates CHRNA2 in cannabis use disorder. Nat Neurosci 22,
- 683 1066-1074. 10.1038/s41593-019-0416-1.
- 684 21. Johnson, E.C., Demontis, D., Thorgeirsson, T.E., Walters, R.K., Polimanti, R., Hatoum,
- A.S., Sanchez-Roige, S., Paul, S.E., Wendt, F.R., Clarke, T.K., et al. (2020). A large-

686		scale genome-wide association study meta-analysis of cannabis use disorder. Lancet
687		Psychiatry 7, 1032-1045. 10.1016/S2215-0366(20)30339-4.
688	22.	Levey, D.F., Galimberti, M., Deak, J.D., Wendt, F.R., Bhattacharya, A., Koller, D.,
689		Harrington, K.M., Quaden, R., Johnson, E.C., Gupta, P., et al. (2023). Multi-ancestry
690		genome-wide association study of cannabis use disorder yields insight into disease
691		biology and public health implications. Nat Genet 55, 2094-2103. 10.1038/s41588-023-
692		01563-z.
693	23.	Xu, H., Toikumo, S., Crist, R.C., Glogowska, K., Jinwala, Z., Deak, J.D., Justice, A.C.,
694		Gelernter, J., Johnson, E.C., Kranzler, H.R., and Kember, R.L. (2023). Identifying
695		genetic loci and phenomic associations of substance use traits: A multi-trait analysis of
696		GWAS (MTAG) study. Addiction 118, 1942-1952. 10.1111/add.16229.
697	24.	Sanchez-Roige, S., Palmer, A.A., and Clarke, T.K. (2020). Recent Efforts to Dissect the
698		Genetic Basis of Alcohol Use and Abuse. Biol Psychiatry 87, 609-618.
699		10.1016/j.biopsych.2019.09.011.
700	25.	McLellan, A.T., Koob, G.F., and Volkow, N.D. (2022). Preaddiction-A Missing Concept
701		for Treating Substance Use Disorders. JAMA Psychiatry 79, 749-751.
702		10.1001/jamapsychiatry.2022.1652.
703	26.	Agrawal, A., Madden, P.A., Bucholz, K.K., Heath, A.C., and Lynskey, M.T. (2014). Initial
704		reactions to tobacco and cannabis smoking: a twin study. Addiction 109, 663-671.
705		10.1111/add.12449.
706	27.	Degenhardt, L., Coffey, C., Carlin, J.B., Swift, W., Moore, E., and Patton, G.C. (2010).
707		Outcomes of occasional cannabis use in adolescence: 10-year follow-up study in
708		Victoria, Australia. Br J Psychiatry 196, 290-295. 10.1192/bjp.bp.108.056952.
709	28.	Scherrer, J.F., Grant, J.D., Duncan, A.E., Sartor, C.E., Haber, J.R., Jacob, T., and
710		Bucholz, K.K. (2009). Subjective effects to cannabis are associated with use, abuse and

- 711 dependence after adjusting for genetic and environmental influences. Drug Alcohol
 712 Depend *105*, 76-82. 10.1016/j.drugalcdep.2009.06.014.
- 713 29. Swift, W., Coffey, C., Carlin, J.B., Degenhardt, L., and Patton, G.C. (2008). Adolescent
- cannabis users at 24 years: trajectories to regular weekly use and dependence in young
- 715 adulthood. Addiction *103*, 1361-1370. 10.1111/j.1360-0443.2008.02246.x.
- 716 30. Windle, M., and Wiesner, M. (2004). Trajectories of marijuana use from adolescence to
- young adulthood: predictors and outcomes. Dev Psychopathol *16*, 1007-1027.
- 718 10.1017/s0954579404040118.
- 719 31. Lyons, M.J., Toomey, R., Meyer, J.M., Green, A.I., Eisen, S.A., Goldberg, J., True, W.R.,

and Tsuang, M.T. (1997). How do genes influence marijuana use? The role of subjective
effects. Addiction *92*, 409-417.

- 32. Leung, J., Chan, G.C.K., Hides, L., and Hall, W.D. (2020). What is the prevalence and
 risk of cannabis use disorders among people who use cannabis? a systematic review
 and meta-analysis. Addict Behav *109*, 106479. 10.1016/j.addbeh.2020.106479.
- 725 33. Haberstick, B.C., Zeiger, J.S., Corley, R.P., Hopfer, C.J., Stallings, M.C., Rhee, S.H.,
- and Hewitt, J.K. (2011). Common and drug-specific genetic influences on subjective
- field reflects to alcohol, tobacco and marijuana use. Addiction 106, 215-224. 10.1111/j.1360-
- 728 0443.2010.03129.x.
- 729 34. Pasman, J.A., Verweij, K.J.H., Gerring, Z., Stringer, S., Sanchez-Roige, S., Treur, J.L.,
 730 Abdellaoui, A., Nivard, M.G., Baselmans, B.M.L., Ong, J.S., et al. (2018). GWAS of
- 731lifetime cannabis use reveals new risk loci, genetic overlap with psychiatric traits, and a732causal influence of schizophrenia. Nat Neurosci *21*, 1161-1170. 10.1038/s41593-018-
- 733 0206-1.
- 35. Stringer, S., Minica, C.C., Verweij, K.J., Mbarek, H., Bernard, M., Derringer, J., van Eijk,
 K.R., Isen, J.D., Loukola, A., Maciejewski, D.F., et al. (2016). Genome-wide association

- 736study of lifetime cannabis use based on a large meta-analytic sample of 32 330 subjects
- from the International Cannabis Consortium. Transl Psychiatry *6*, e769.
- 738 10.1038/tp.2016.36.
- 739 36. Sanchez-Roige, S., and Palmer, A.A. (2020). Emerging phenotyping strategies will
- advance our understanding of psychiatric genetics. Nat Neurosci 23, 475-480.
- 741 10.1038/s41593-020-0609-7.
- 742 37. Thorpe, H.H.A., Talhat, M.A., and Khokhar, J.Y. (2021). High genes: Genetic
- 743 underpinnings of cannabis use phenotypes. Prog Neuropsychopharmacol Biol
- 744 Psychiatry *106*, 110164. 10.1016/j.pnpbp.2020.110164.
- 745 38. McBain, R.K., Wong, E.C., Breslau, J., Shearer, A.L., Cefalu, M.S., Roth, E., Burnam,
- 746 M.A., and Collins, R.L. (2020). State medical marijuana laws, cannabis use and
- cannabis use disorder among adults with elevated psychological distress. Drug Alcohol
- 748 Depend *215*, 108191. 10.1016/j.drugalcdep.2020.108191.
- Hughes, J.R., Naud, S., Budney, A.J., Fingar, J.R., and Callas, P.W. (2016). Attempts to
 stop or reduce daily cannabis use: An intensive natural history study. Psychol Addict
- 751 Behav *30*, 389-397. 10.1037/adb0000155.
- 40. Mallard, T.T., and Sanchez-Roige, S. (2021). Dimensional Phenotypes in Psychiatric
- Genetics: Lessons from Genome-Wide Association Studies of Alcohol Use Phenotypes.
 Complex Psychiatry 7, 45-48. 10.1159/000518863.
- 755 41. Fogel, A.I., Akins, M.R., Krupp, A.J., Stagi, M., Stein, V., and Biederer, T. (2007).
- 756 SynCAMs organize synapses through heterophilic adhesion. J Neurosci 27, 12516-
- 757 12530. 10.1523/JNEUROSCI.2739-07.2007.
- 42. Yan, X., Wang, Z., Schmidt, V., Gauert, A., Willnow, T.E., Heinig, M., and Poy, M.N.
- 759 (2018). Cadm2 regulates body weight and energy homeostasis in mice. Mol Metab 8,

760 180-188. 10.1016/j.molmet.2017.11.010.

761	43.	Tyler, R.E., Besheer, J., and Joffe, M.E. (2022). Advances in translating mGlu(2) and
762		mGlu(3) receptor selective allosteric modulators as breakthrough treatments for affective
763		disorders and alcohol use disorder. Pharmacol Biochem Behav 219, 173450.
764		10.1016/j.pbb.2022.173450.
765	44.	Karlsson Linner, R., Mallard, T.T., Barr, P.B., Sanchez-Roige, S., Madole, J.W., Driver,
766		M.N., Poore, H.E., de Vlaming, R., Grotzinger, A.D., Tielbeek, J.J., et al. (2021).
767		Multivariate analysis of 1.5 million people identifies genetic associations with traits
768		related to self-regulation and addiction. Nat Neurosci 24, 1367-1376. 10.1038/s41593-
769		021-00908-3.
770	45.	Agrawal, A., Budney, A.J., and Lynskey, M.T. (2012). The co-occurring use and misuse
771		of cannabis and tobacco: a review. Addiction 107, 1221-1233. 10.1111/j.1360-
772		0443.2012.03837.x.
773	46.	Galimberti, M., Levey, D.F., Deak, J.D., Zhou, H., Stein, M.B., and Gelernter, J. (2024).
774		Genetic influences and causal pathways shared between cannabis use disorder and
775		other substance use traits. Mol Psychiatry. 10.1038/s41380-024-02548-y.
776	47.	Poore, H.E., Hatoum, A., Mallard, T.T., Sanchez-Roige, S., Waldman, I.D., Palmer, A.A.,
777		Harden, K.P., Barr, P.B., and Dick, D.M. (2023). A multivariate approach to
778		understanding the genetic overlap between externalizing phenotypes and substance use
779		disorders. Addict Biol 28, e13319. 10.1111/adb.13319.
780	48.	Sanchez-Roige, S., Jennings, M.V., Thorpe, H.H.A., Mallari, J.E., van der Werf, L.C.,
781		Bianchi, S.B., Huang, Y., Lee, C., Mallard, T.T., Barnes, S.A., et al. (2023). CADM2 is
782		implicated in impulsive personality and numerous other traits by genome- and phenome-
783		wide association studies in humans and mice. Transl Psychiatry 13, 167.
784		10.1038/s41398-023-02453-y.

785	49	lang S.K. Saunders G. Liu M. and Me Research T. Jiang Y. Liu D. L. and Vrieze
700	40.	0 (0000) Operation and the provide the second second states between substances
786		S. (2022). Genetic correlation, pleiotropy, and causal associations between substance
787		use and psychiatric disorder. Psychol Med 52, 968-978. 10.1017/S003329172000272X.
788	50.	Chang, L.H., Ong, J.S., An, J., Verweij, K.J.H., Vink, J.M., Pasman, J., Liu, M.,
789		MacGregor, S., Cornelis, M.C., Martin, N.G., and Derks, E.M. (2020). Investigating the
790		genetic and causal relationship between initiation or use of alcohol, caffeine, cannabis
791		and nicotine. Drug Alcohol Depend 210, 107966. 10.1016/j.drugalcdep.2020.107966.
792	51.	Hatoum, A.S., Colbert, S.M.C., Johnson, E.C., Huggett, S.B., Deak, J.D., Pathak, G.,
793		Jennings, M.V., Paul, S.E., Karcher, N.R., Hansen, I., et al. (2023). Multivariate genome-
794		wide association meta-analysis of over 1 million subjects identifies loci underlying
795		multiple substance use disorders. Nat Ment Health 1, 210-223. 10.1038/s44220-023-
796		00034-у.
797	52.	Abdellaoui, A., Smit, D.J.A., van den Brink, W., Denys, D., and Verweij, K.J.H. (2021).
798		Genomic relationships across psychiatric disorders including substance use disorders.
799		Drug Alcohol Depend 220, 108535. 10.1016/j.drugalcdep.2021.108535.
800	53.	Waldman, I.D., Poore, H.E., Luningham, J.M., and Yang, J. (2020). Testing structural
801		models of psychopathology at the genomic level. World Psychiatry 19, 350-359.
802		10.1002/wps.20772.
803	54.	Mallard, T.T., Savage, J.E., Johnson, E.C., Huang, Y., Edwards, A.C., Hottenga, J.J.,
804		Grotzinger, A.D., Gustavson, D.E., Jennings, M.V., Anokhin, A., et al. (2022). Item-Level
805		Genome-Wide Association Study of the Alcohol Use Disorders Identification Test in
806		Three Population-Based Cohorts. Am J Psychiatry 179, 58-70.
807		10.1176/appi.ajp.2020.20091390.
808	55.	Kranzler, H.R., Zhou, H., Kember, R.L., Vickers Smith, R., Justice, A.C., Damrauer, S.,
809		Tsao, P.S., Klarin, D., Baras, A., Reid, J., et al. (2019). Genome-wide association study

810		of alcohol consumption and use disorder in 274,424 individuals from multiple
811		populations. Nat Commun 10, 1499. 10.1038/s41467-019-09480-8.
812	56.	Toikumo, S., Jennings, M.V., Pham, B.K., Lee, H., Mallard, T.T., Bianchi, S.B., Meredith,
813		J.J., Vilar-Ribo, L., Xu, H., Hatoum, A.S., et al. (2024). Multi-ancestry meta-analysis of
814		tobacco use disorder identifies 461 potential risk genes and reveals associations with
815		multiple health outcomes. Nature Human Behavior. 10.1038/s41562-024-01851-6.
816	57.	Zhou, H., Kember, R.L., Deak, J.D., Xu, H., Toikumo, S., Yuan, K., Lind, P.A.,
817		Farajzadeh, L., Wang, L., Hatoum, A.S., et al. (2023). Multi-ancestry study of the
818		genetics of problematic alcohol use in over 1 million individuals. Nat Med 29, 3184-3192.
819		10.1038/s41591-023-02653-5.
820	58.	Sanchez-Roige, S., Palmer, A.A., Fontanillas, P., Elson, S.L., 23andMe Research Team,
821		Substance Use Disorder Working Group of the Psychiatric Genomics Consortium,
822		Adams, M.J., Howard, D.M., Edenberg, H.J., Davies, G., et al. (2019). Genome-Wide
823		Association Study Meta-Analysis of the Alcohol Use Disorders Identification Test
824		(AUDIT) in Two Population-Based Cohorts. Am J Psychiatry 176, 107-118.
825		10.1176/appi.ajp.2018.18040369.
826	59.	Savage, J.E., Barr, P.B., Phung, T., Lee, Y.H., Zhang, Y., Ge, T., Smoller, J.W., Davis,
827		L.K., Meyers, J., Porjesz, B., et al. (2023). Genetic Heterogeneity Across Dimensions of
828		Alcohol Use Behaviors. medRxiv, 2023.2012.2026.23300537.
829		10.1101/2023.12.26.23300537.
830	60.	Saunders, G.R.B., Wang, X., Chen, F., Jang, S.K., Liu, M., Wang, C., Gao, S., Jiang, Y.,
831		Khunsriraksakul, C., Otto, J.M., et al. (2022). Genetic diversity fuels gene discovery for
832		tobacco and alcohol use. Nature 612, 720-724. 10.1038/s41586-022-05477-4.
833	61.	Trubetskoy, V., Pardinas, A.F., Qi, T., Panagiotaropoulou, G., Awasthi, S., Bigdeli, T.B.,
834		Bryois, J., Chen, C.Y., Dennison, C.A., Hall, L.S., et al. (2022). Mapping genomic loci

implicates genes and synaptic biology in schizophrenia. Nature *604*, 502-508.

836 10.1038/s41586-022-04434-5.

- 62. Lam, M., Chen, C.Y., Li, Z., Martin, A.R., Bryois, J., Ma, X., Gaspar, H., Ikeda, M.,
- 838 Benyamin, B., Brown, B.C., et al. (2019). Comparative genetic architectures of
- schizophrenia in East Asian and European populations. Nat Genet *51*, 1670-1678.
- 840 10.1038/s41588-019-0512-x.
- 841 63. Pardinas, A.F., Holmans, P., Pocklington, A.J., Escott-Price, V., Ripke, S., Carrera, N.,
- Legge, S.E., Bishop, S., Cameron, D., Hamshere, M.L., et al. (2018). Common
- schizophrenia alleles are enriched in mutation-intolerant genes and in regions under

strong background selection. Nat Genet *50*, 381-389. 10.1038/s41588-018-0059-2.

- 845 64. Li, Z., Chen, J., Yu, H., He, L., Xu, Y., Zhang, D., Yi, Q., Li, C., Li, X., Shen, J., et al.
- 846 (2017). Genome-wide association analysis identifies 30 new susceptibility loci for
 847 schizophrenia. Nat Genet *49*, 1576-1583. 10.1038/ng.3973.
- 848 65. Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014).
- Biological insights from 108 schizophrenia-associated genetic loci. Nature *511*, 421-427.
- 850 10.1038/nature13595.
- 851 66. Nagel, M., Watanabe, K., Stringer, S., Posthuma, D., and van der Sluis, S. (2018). Item852 level analyses reveal genetic heterogeneity in neuroticism. Nat Commun *9*, 905.
- 853 10.1038/s41467-018-03242-8.
- 854 67. Luciano, M., Hagenaars, S.P., Davies, G., Hill, W.D., Clarke, T.K., Shirali, M., Harris,

855 S.E., Marioni, R.E., Liewald, D.C., Fawns-Ritchie, C., et al. (2018). Association analysis

- in over 329,000 individuals identifies 116 independent variants influencing neuroticism.
- 857 Nat Genet *50*, 6-11. 10.1038/s41588-017-0013-8.
- 858 68. Lee, J.J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., Nguyen-Viet, T.A.,
- Bowers, P., Sidorenko, J., Karlsson Linner, R., et al. (2018). Gene discovery and

860		polygenic prediction from a genome-wide association study of educational attainment in
861		1.1 million individuals. Nat Genet 50, 1112-1121. 10.1038/s41588-018-0147-3.
862	69.	Hysi, P.G., Choquet, H., Khawaja, A.P., Wojciechowski, R., Tedja, M.S., Yin, J., Simcoe,
863		M.J., Patasova, K., Mahroo, O.A., Thai, K.K., et al. (2020). Meta-analysis of 542,934
864		subjects of European ancestry identifies new genes and mechanisms predisposing to
865		refractive error and myopia. Nat Genet 52, 401-407. 10.1038/s41588-020-0599-0.
866	70.	Tedja, M.S., Wojciechowski, R., Hysi, P.G., Eriksson, N., Furlotte, N.A., Verhoeven,
867		V.J.M., Iglesias, A.I., Meester-Smoor, M.A., Tompson, S.W., Fan, Q., et al. (2018).
868		Genome-wide association meta-analysis highlights light-induced signaling as a driver for
869		refractive error. Nat Genet 50, 834-848. 10.1038/s41588-018-0127-7.
870	71.	Watanabe, K., Jansen, P.R., Savage, J.E., Nandakumar, P., Wang, X., 23andMe
871		Research Team, Hinds, D.A., Gelernter, J., Levey, D.F., Polimanti, R., et al. (2022).
872		Genome-wide meta-analysis of insomnia prioritizes genes associated with metabolic and
873		psychiatric pathways. Nat Genet 54, 1125-1132. 10.1038/s41588-022-01124-w.
874	72.	Di Menna, L., Joffe, M.E., Iacovelli, L., Orlando, R., Lindsley, C.W., Mairesse, J.,
875		Gressens, P., Cannella, M., Caraci, F., Copani, A., et al. (2018). Functional partnership
876		between mGlu3 and mGlu5 metabotropic glutamate receptors in the central nervous
877		system. Neuropharmacology 128, 301-313. 10.1016/j.neuropharm.2017.10.026.
878	73.	Ibrahim, K.S., Abd-Elrahman, K.S., El Mestikawy, S., and Ferguson, S.S.G. (2020).
879		Targeting Vesicular Glutamate Transporter Machinery: Implications on Metabotropic
880		Glutamate Receptor 5 Signaling and Behavior. Mol Pharmacol 98, 314-327.
881		10.1124/molpharm.120.000089.
882	74.	Jung, K.M., Mangieri, R., Stapleton, C., Kim, J., Fegley, D., Wallace, M., Mackie, K., and
883		Piomelli, D. (2005). Stimulation of endocannabinoid formation in brain slice cultures

- through activation of group I metabotropic glutamate receptors. Mol Pharmacol 68,
- 885 1196-1202. 10.1124/mol.105.013961.
- 886 75. The GTEx Consortium (2024). GTEx Portal. <u>https://gtexportal.org/</u>.
- 887 76. Schoeler, T., Speed, D., Porcu, E., Pirastu, N., Pingault, J.B., and Kutalik, Z. (2023).
- 888 Participation bias in the UK Biobank distorts genetic associations and downstream
- analyses. Nat Hum Behav 7, 1216-1227. 10.1038/s41562-023-01579-9.
- 890 77. Deak, J.D., Levey, D.F., Wendt, F.R., Zhou, H., Galimberti, M., Kranzler, H.R., Gaziano,
- J.M., Stein, M.B., Polimanti, R., Gelernter, J., and Million Veteran, P. (2022). Genome-
- 892 Wide Investigation of Maximum Habitual Alcohol Intake in US Veterans in Relation to
- Alcohol Consumption Traits and Alcohol Use Disorder. JAMA Netw Open 5, e2238880.
- 894 10.1001/jamanetworkopen.2022.38880.
- 895 78. Pasman, J.A., Demange, P.A., Guloksuz, S., Willemsen, A.H.M., Abdellaoui, A., Ten
- Have, M., Hottenga, J.J., Boomsma, D.I., de Geus, E., Bartels, M., et al. (2022). Genetic
- 897 Risk for Smoking: Disentangling Interplay Between Genes and Socioeconomic Status.

Behav Genet *52*, 92-107. 10.1007/s10519-021-10094-4.

- 899 79. Xu, K., Li, B., McGinnis, K.A., Vickers-Smith, R., Dao, C., Sun, N., Kember, R.L., Zhou,
- H., Becker, W.C., Gelernter, J., et al. (2020). Genome-wide association study of smoking
 trajectory and meta-analysis of smoking status in 842,000 individuals. Nat Commun *11*,
 5302. 10.1038/s41467-020-18489-3.
- 903 80. Zhou, H., Sealock, J.M., Sanchez-Roige, S., Clarke, T.K., Levey, D.F., Cheng, Z., Li, B.,
- 904 Polimanti, R., Kember, R.L., Smith, R.V., et al. (2020). Genome-wide meta-analysis of
- 905 problematic alcohol use in 435,563 individuals yields insights into biology and
- 906 relationships with other traits. Nat Neurosci 23, 809-818. 10.1038/s41593-020-0643-5.
- 907 81. Cai, N., Revez, J.A., Adams, M.J., Andlauer, T.F.M., Breen, G., Byrne, E.M., Clarke,
- 908 T.K., Forstner, A.J., Grabe, H.J., Hamilton, S.P., et al. (2020). Minimal phenotyping

- 909 yields genome-wide association signals of low specificity for major depression. Nat
 910 Genet *52*, 437-447. 10.1038/s41588-020-0594-5.
- 911 82. Evangelou, E., Gao, H., Chu, C., Ntritsos, G., Blakeley, P., Butts, A.R., Pazoki, R.,
- 912 Suzuki, H., Koskeridis, F., Yiorkas, A.M., et al. (2019). New alcohol-related genes
- 913 suggest shared genetic mechanisms with neuropsychiatric disorders. Nat Hum Behav 3,
- 914 950-961. 10.1038/s41562-019-0653-z.
- 915 83. Zhong, V.W., Kuang, A., Danning, R.D., Kraft, P., van Dam, R.M., Chasman, D.I., and
- 916 Cornelis, M.C. (2019). A genome-wide association study of bitter and sweet beverage
- 917 consumption. Hum Mol Genet *28*, 2449-2457. 10.1093/hmg/ddz061.
- 918 84. Sanchez-Roige, S., Fontanillas, P., Elson, S.L., Gray, J.C., de Wit, H., MacKillop, J., and
- 919 Palmer, A.A. (2019). Genome-Wide Association Studies of Impulsive Personality Traits
- 920 (BIS-11 and UPPS-P) and Drug Experimentation in up to 22,861 Adult Research
- 921 Participants Identify Loci in the CACNA11 and CADM2 genes. J Neurosci 39, 2562-2572.
- 922 10.1523/JNEUROSCI.2662-18.2019.
- 923 85. Karlsson Linner, R., Biroli, P., Kong, E., Meddens, S.F.W., Wedow, R., Fontana, M.A.,
- 924 Lebreton, M., Tino, S.P., Abdellaoui, A., Hammerschlag, A.R., et al. (2019). Genome-
- 925 wide association analyses of risk tolerance and risky behaviors in over 1 million
- 926 individuals identify hundreds of loci and shared genetic influences. Nat Genet 51, 245-
- 927 257. 10.1038/s41588-018-0309-3.
- 928 86. Liu, M., Jiang, Y., Wedow, R., Li, Y., Brazel, D.M., Chen, F., Datta, G., Davila-Velderrain,
- 929 J., McGuire, D., Tian, C., et al. (2019). Association studies of up to 1.2 million individuals
- 930 yield new insights into the genetic etiology of tobacco and alcohol use. Nat Genet 51,
- 931 237-244. 10.1038/s41588-018-0307-5.
- 932 87. Erzurumluoglu, A.M., Liu, M., Jackson, V.E., Barnes, D.R., Datta, G., Melbourne, C.A.,
- 933 Young, R., Batini, C., Surendran, P., Jiang, T., et al. (2020). Meta-analysis of up to

- 934 622,409 individuals identifies 40 novel smoking behaviour associated genetic loci. Mol
 935 Psychiatry 25, 2392-2409. 10.1038/s41380-018-0313-0.
- 936 88. Kichaev, G., Bhatia, G., Loh, P.R., Gazal, S., Burch, K., Freund, M.K., Schoech, A.,
- 937 Pasaniuc, B., and Price, A.L. (2019). Leveraging Polygenic Functional Enrichment to
- 938 Improve GWAS Power. Am J Hum Genet *104*, 65-75. 10.1016/j.ajhg.2018.11.008.
- 939 89. Clifton, E.A.D., Perry, J.R.B., Imamura, F., Lotta, L.A., Brage, S., Forouhi, N.G., Griffin,
- 940 S.J., Wareham, N.J., Ong, K.K., and Day, F.R. (2018). Genome-wide association study
- 941 for risk taking propensity indicates shared pathways with body mass index. Commun
- 942 Biol 1, 36. 10.1038/s42003-018-0042-6.
- 943 90. Clarke, T.K., Adams, M.J., Davies, G., Howard, D.M., Hall, L.S., Padmanabhan, S.,
- 944 Murray, A.D., Smith, B.H., Campbell, A., Hayward, C., et al. (2017). Genome-wide
- 945 association study of alcohol consumption and genetic overlap with other health-related
- 946 traits in UK Biobank (N=112 117). Mol Psychiatry 22, 1376-1384. 10.1038/mp.2017.153.
- 947 91. Brazel, D.M., Jiang, Y., Hughey, J.M., Turcot, V., Zhan, X., Gong, J., Batini, C.,
- 948 Weissenkampen, J.D., Liu, M., CHD Exome+ Consortium Consortium for Genetics of
- 949 Smoking Behaviour, et al. (2019). Exome Chip Meta-analysis Fine Maps Causal
- 950 Variants and Elucidates the Genetic Architecture of Rare Coding Variants in Smoking
- 951 and Alcohol Use. Biol Psychiatry *85*, 946-955. 10.1016/j.biopsych.2018.11.024.
- 952 92. Baselmans, B., Hammerschlag, A.R., Noordijk, S., Ip, H., van der Zee, M., de Geus, E.,
- 953 Abdellaoui, A., Treur, J.L., and van 't Ent, D. (2022). The Genetic and Neural Substrates
- 954 of Externalizing Behavior. Biol Psychiatry Glob Open Sci 2, 389-399.
- 955 10.1016/j.bpsgos.2021.09.007.
- 956 93. Strawbridge, R.J., Ward, J., Cullen, B., Tunbridge, E.M., Hartz, S., Bierut, L., Horton, A.,
- 957 Bailey, M.E.S., Graham, N., Ferguson, A., et al. (2018). Genome-wide analysis of self-

- 958 reported risk-taking behaviour and cross-disorder genetic correlations in the UK Biobank
 959 cohort. Transl Psychiatry *8*, 39. 10.1038/s41398-017-0079-1.
- 960 94. Mills, M.C., Tropf, F.C., Brazel, D.M., van Zuydam, N., Vaez, A., eQTLGen Consortium,
- 961 Bios Consortium, Human Reproductive Behaviour Consortium, Pers, T.H., Snieder, H.,
- 962 et al. (2021). Identification of 371 genetic variants for age at first sex and birth linked to
- 963 externalising behaviour. Nat Hum Behav *5*, 1717-1730. 10.1038/s41562-021-01135-3.
- 964 95. Johnson, E.C., Sanchez-Roige, S., Acion, L., Adams, M.J., Bucholz, K.K., Chan, G.,

965 Chao, M.J., Chorlian, D.B., Dick, D.M., Edenberg, H.J., et al. (2021). Polygenic

- 966 contributions to alcohol use and alcohol use disorders across population-based and
- 967 clinically ascertained samples. Psychol Med *51*, 1147-1156.
- 968 10.1017/S0033291719004045.
- 969 96. Sun, Y., Chang, S., Wang, F., Sun, H., Ni, Z., Yue, W., Zhou, H., Gelernter, J., Malison,
- 970 R.T., Kalayasiri, R., et al. (2019). Genome-wide association study of alcohol
- 971 dependence in male Han Chinese and cross-ethnic polygenic risk score comparison.

972 Transl Psychiatry *9*, 249. 10.1038/s41398-019-0586-3.

- 973 97. Lai, D., Wetherill, L., Bertelsen, S., Carey, C.E., Kamarajan, C., Kapoor, M., Meyers,
- 974J.L., Anokhin, A.P., Bennett, D.A., Bucholz, K.K., et al. (2019). Genome-wide association975studies of alcohol dependence, DSM-IV criterion count and individual criteria. Genes
- 976 Brain Behav *18*, e12579. 10.1111/gbb.12579.

977 98. Chang, L.H., Couvy-Duchesne, B., Liu, M., Medland, S.E., Verhulst, B., Benotsch, E.G.,

- 978 Hickie, I.B., Martin, N.G., Gillespie, N.A., and GSCAN Consortium (2019). Association
- 979 between polygenic risk for tobacco or alcohol consumption and liability to licit and illicit
- 980 substance use in young Australian adults. Drug Alcohol Depend *197*, 271-279.
- 981 10.1016/j.drugalcdep.2019.01.015.

- 982 99. Allegrini, A.G., Verweij, K.J.H., Abdellaoui, A., Treur, J.L., Hottenga, J.J., Willemsen, G.,
- 983 Boomsma, D.I., International Cannabis Consortium, and Vink, J.M. (2019). Genetic
- 984 Vulnerability for Smoking and Cannabis Use: Associations With E-Cigarette and Water
- 985 Pipe Use. Nicotine Tob Res *21*, 723-730. 10.1093/ntr/nty150.
- 986 100. Hodgson, K., Coleman, J.R.I., Hagenaars, S.P., Purves, K.L., Glanville, K., Choi, S.W.,
- 987 O'Reilly, P., Breen, G., Major Depressive Disorder Working Group of the Psychiatric
- 988 Genomics Consortium, and Lewis, C.M. (2020). Cannabis use, depression and self-
- harm: phenotypic and genetic relationships. Addiction *115*, 482-492.
- 990 10.1111/add.14845.
- 991 101. Smeland, O.B., and Andreassen, O.A. (2021). Polygenic risk scores in psychiatry -
- Large potential but still limited clinical utility. Eur Neuropsychopharmacol *51*, 68-70.
- 993 10.1016/j.euroneuro.2021.05.007.
- 102. Kember, R.L., Hartwell, E.E., Xu, H., Rotenberg, J., Almasy, L., Zhou, H., Gelernter, J.,
- and Kranzler, H.R. (2023). Phenome-wide Association Analysis of Substance Use
- 996 Disorders in a Deeply Phenotyped Sample. Biol Psychiatry *93*, 536-545.
- 997 10.1016/j.biopsych.2022.08.010.
- 998 103. Cheng, W., Parker, N., Karadag, N., Koch, E., Hindley, G., Icick, R., Shadrin, A.,
- 999 O'Connell, K.S., Bjella, T., Bahrami, S., et al. (2023). The relationship between cannabis
- 1000 use, schizophrenia, and bipolar disorder: a genetically informed study. Lancet Psychiatry
- 1001 *10*, 441-451. 10.1016/S2215-0366(23)00143-8.
- 1002 104. Johnson, E.C., Hatoum, A.S., Deak, J.D., Polimanti, R., Murray, R.M., Edenberg, H.J.,
- 1003 Gelernter, J., Di Forti, M., and Agrawal, A. (2021). The relationship between cannabis
- and schizophrenia: a genetically informed perspective. Addiction *116*, 3227-3234.
- 1005 10.1111/add.15534.

1006	105.	Reginsson, G.W., Ingason, A., Euesden, J., Bjornsdottir, G., Olafsson, S., Sigurdsson,
1007		E., Oskarsson, H., Tyrfingsson, T., Runarsdottir, V., Hansdottir, I., et al. (2018).
1008		Polygenic risk scores for schizophrenia and bipolar disorder associate with addiction.
1009		Addict Biol 23, 485-492. 10.1111/adb.12496.
1010	106.	Johnson, E.C., Austin-Zimmerman, I., Thorpe, H.H.A., Levey, D.F., Baranger, D.A.A.,
1011		Colbert, S.M.C., Thorgeirsson, T., Khokhar, J.Y., Davis, L.K., Di Forti, M., et al. (2024).
1012		Dissecting the relationships between tobacco smoking, cannabis use disorder, and
1013		schizophrenia using genomic approaches.
1014	107.	Hamilton, I., and Monaghan, M. (2019). Cannabis and Psychosis: Are We any Closer to
1015		Understanding the Relationship? Curr Psychiatry Rep 21, 48. 10.1007/s11920-019-
1016		1044-x.
1017	108.	Di Forti, M., Quattrone, D., Freeman, T.P., Tripoli, G., Gayer-Anderson, C., Quigley, H.,
1018		Rodriguez, V., Jongsma, H.E., Ferraro, L., La Cascia, C., et al. (2019). The contribution
1019		of cannabis use to variation in the incidence of psychotic disorder across Europe (EU-
1020		GEI): a multicentre case-control study. Lancet Psychiatry 6, 427-436. 10.1016/S2215-
1021		0366(19)30048-3.
1022	109.	Oluwoye, O., Monroe-DeVita, M., Burduli, E., Chwastiak, L., McPherson, S., McClellan,
1023		J.M., and McDonell, M.G. (2019). Impact of tobacco, alcohol and cannabis use on
1024		treatment outcomes among patients experiencing first episode psychosis: Data from the
1025		national RAISE-ETP study. Early Interv Psychiatry 13, 142-146. 10.1111/eip.12542.
1026	110.	Schoeler, T., Monk, A., Sami, M.B., Klamerus, E., Foglia, E., Brown, R., Camuri, G.,
1027		Altamura, A.C., Murray, R., and Bhattacharyya, S. (2016). Continued versus
1028		discontinued cannabis use in patients with psychosis: a systematic review and meta-
1029		analysis. Lancet Psychiatry 3, 215-225. 10.1016/S2215-0366(15)00363-6.

- 1030 111. Durvasula, A., and Price, A.L. (2023). Distinct explanations underlie gene-environment
 1031 interactions in the UK Biobank. medRxiv. 10.1101/2023.09.22.23295969.
- 1032 112. Abdellaoui, A., Dolan, C.V., Verweij, K.J.H., and Nivard, M.G. (2022). Gene-environment
- 1033 correlations across geographic regions affect genome-wide association studies. Nat
- 1034 Genet *54*, 1345-1354. 10.1038/s41588-022-01158-0.
- 1035 113. Thorpe, H.H.A., Fontanillas, P., Pham, B.K., Meredith, J.J., Jennings, M.V., Courchesne-
- 1036 Krak, N.S., Vilar-Ribó, L., Bianchi, S.B., Mutz, J., 23andMe Research Team, et al.
- 1037 (2024). Genome-wide association studies of coffee intake in UK/US participants of
- 1038 European ancestry uncover cohort-specific genetic associations.
- 1039 Neuropsychopharmacology. 10.1038/s41386-024-01870-x.
- 1040 114. Kayir, H., Ruffolo, J., McCunn, P., and Khokhar, J.Y. (2023). The Relationship Between
- 1041 Cannabis, Cognition, and Schizophrenia: It's Complicated. Curr Top Behav Neurosci 63,
- 1042 437-461. 10.1007/7854_2022_396.
- 1043 115. Farrelly, K.N., Wardell, J.D., Marsden, E., Scarfe, M.L., Najdzionek, P., Turna, J., and
- 1044 MacKillop, J. (2023). The Impact of Recreational Cannabis Legalization on Cannabis
- 1045 Use and Associated Outcomes: A Systematic Review. Subst Abuse 17,
- 1046 11782218231172054. 10.1177/11782218231172054.
- 1047 116. Fergusson, D.M., Boden, J.M., and Horwood, L.J. (2015). Psychosocial sequelae of
- 1048 cannabis use and implications for policy: findings from the Christchurch Health and
- 1049 Development Study. Soc Psychiatry Psychiatr Epidemiol *50*, 1317-1326.
- 1050 10.1007/s00127-015-1070-x.
- 1051 117. van der Pol, P., Liebregts, N., de Graaf, R., Korf, D.J., van den Brink, W., and van Laar,
- 1052 M. (2013). Predicting the transition from frequent cannabis use to cannabis dependence:
- a three-year prospective study. Drug Alcohol Depend *133*, 352-359.
- 1054 10.1016/j.drugalcdep.2013.06.009.

1055	118.	Walden, N., and Earleywine, M. (2008). How high: quantity as a predictor of cannabis-
1056		related problems. Harm Reduct J 5, 20. 10.1186/1477-7517-5-20.

- 1057 119. Zeisser, C., Thompson, K., Stockwell, T., Duff, C., Chow, C., Vallance, K., Ivsins, A.,
- 1058 Michelow, W., Marsh, D., and Lucas, P. (2012). A 'standard joint'? The role of quantity in
- 1059 predicting cannabis-related problems. Addiction Research & Theory 20, 82-92.
- 1060 10.3109/16066359.2011.569101.
- 1061 120. Substance Abuse and Mental Health Services Administration (2022). Key substance use
- and mental health indicators in the United States: Results from the 2022 National Survey
- 1063 on Drug Use and Health. Center for Behavioral Health Statistics and Quality, Substance
- 1064 Abuse and Mental Health Services Administration.
- 1065 https://www.samhsa.gov/data/sites/default/files/reports/rpt42731/2022-nsduh-nnr.pdf.
- 1066 121. Althubaiti, A. (2016). Information bias in health research: definition, pitfalls, and
- adjustment methods. J Multidiscip Healthc 9, 211-217. 10.2147/JMDH.S104807.
- 1068 122. Jeffers, A.M., Glantz, S., Byers, A., and Keyhani, S. (2021). Sociodemographic
- 1069 Characteristics Associated With and Prevalence and Frequency of Cannabis Use
- 1070 Among Adults in the US. JAMA Netw Open *4*, e2136571.
- 1071 10.1001/jamanetworkopen.2021.36571.
- 1072 123. Karriker-Jaffe, K.J. (2013). Neighborhood socioeconomic status and substance use by
- 1073 U.S. adults. Drug Alcohol Depend *133*, 212-221. 10.1016/j.drugalcdep.2013.04.033.
- 1074 124. Atkinson, E.G., Bianchi, S.B., Ye, G.Y., Martinez-Magana, J.J., Tietz, G.E., Montalvo-
- 1075 Ortiz, J.L., Giusti-Rodriguez, P., Palmer, A.A., and Sanchez-Roige, S. (2022). Cross-
- 1076 ancestry genomic research: time to close the gap. Neuropsychopharmacology 47, 1737-
- 1077 1738. 10.1038/s41386-022-01365-7.

- 1078 125. Martin, A.R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B.M., and Daly, M.J. (2019).
- 1079 Clinical use of current polygenic risk scores may exacerbate health disparities. Nat 1080 Genet *51*, 584-591. 10.1038/s41588-019-0379-x.
- 1081 126. Bryc, K., Durand, E.Y., Macpherson, J.M., Reich, D., and Mountain, J.L. (2015). The
- 1082 genetic ancestry of African Americans, Latinos, and European Americans across the
- 1083 United States. Am J Hum Genet *96*, 37-53. 10.1016/j.ajhg.2014.11.010.
- 1084 127. National Academies of Sciences, Engineering, and Medicine Health and Medicine
- 1085 Division, Division of Behavioral and Social Sciences and Education, Board on Health
- 1086 Sciences Policy, Committee on Population, Committee on the Use of Race, and
- 1087 Ethnicity, and Ancestry as Population Descriptors in Genomics Research (2023). In
- 1088Using Population Descriptors in Genetics and Genomics Research: A New Framework
- 1089 for an Evolving Field. 10.17226/26902.
- 1090 128. Durand, E.Y., Do, C.B., Mountain, J.L., and Macpherson, J.M. (2014). Ancestry
- 1091 Composition: A Novel, Efficient Pipeline for Ancestry Deconvolution. bioRxiv. [Preprint].
 1092 10.1101/010512.
- 1093 129. Yin, B., Wang, X., Huang, T., and Jia, J. (2022). Shared Genetics and Causality
- 1094 Between Decaffeinated Coffee Consumption and Neuropsychiatric Diseases: A Large-
- Scale Genome-Wide Cross-Trait Analysis and Mendelian Randomization Analysis. Front
 Psychiatry *13*, 910432. 10.3389/fpsyt.2022.910432.
- 1097 130. Agrawal, A., Chou, Y.L., Carey, C.E., Baranger, D.A.A., Zhang, B., Sherva, R., Wetherill,
- 1098 L., Kapoor, M., Wang, J.C., Bertelsen, S., et al. (2018). Genome-wide association study
- 1099 identifies a novel locus for cannabis dependence. Mol Psychiatry 23, 1293-1302.
- 1100 10.1038/mp.2017.200.
- 1101 131. Sherva, R., Wang, Q., Kranzler, H., Zhao, H., Koesterer, R., Herman, A., Farrer, L.A.,
- and Gelernter, J. (2016). Genome-wide Association Study of Cannabis Dependence

- Severity, Novel Risk Variants, and Shared Genetic Risks. JAMA Psychiatry *73*, 472-480.
 10.1001/jamapsychiatry.2016.0036.
- 1105 132. Agrawal, A., Lynskey, M.T., Hinrichs, A., Grucza, R., Saccone, S.F., Krueger, R.,
- 1106 Neuman, R., Howells, W., Fisher, S., Fox, L., et al. (2011). A genome-wide association
- 1107 study of DSM-IV cannabis dependence. Addict Biol 16, 514-518. 10.1111/j.1369-
- 1108 1600.2010.00255.x.
- 1109 133. Minica, C.C., Dolan, C.V., Hottenga, J.J., Pool, R., Genome of the Netherlands
- 1110 Consortium, Fedko, I.O., Mbarek, H., Huppertz, C., Bartels, M., Boomsma, D.I., and
- 1111 Vink, J.M. (2015). Heritability, SNP- and Gene-Based Analyses of Cannabis Use
- 1112 Initiation and Age at Onset. Behav Genet *45*, 503-513. 10.1007/s10519-015-9723-9.
- 1113 134. Zhao, Y., Han, X., and Zheng, Z.L. (2023). Analysis of the brain transcriptome for
- substance-associated genes: An update on large-scale genome-wide association
 studies. Addict Biol 28, e13332. 10.1111/adb.13332.
- 1116 135. Greco, L.A., Reay, W.R., Dayas, C.V., and Cairns, M.J. (2023). Exploring opportunities
- 1117 for drug repurposing and precision medicine in cannabis use disorder using genetics.
- 1118 Addict Biol 28, e13313. 10.1111/adb.13313.
- 1119 136. Carreras-Gallo, N., Dwaraka, V.B., Caceres, A., Smith, R., Mendez, T.L., Went, H., and
- 1120 Gonzalez, J.R. (2023). Impact of tobacco, alcohol, and marijuana on genome-wide DNA
- 1121 methylation and its relationship with hypertension. Epigenetics *18*, 2214392.
- 1122 10.1080/15592294.2023.2214392.
- 1123 137. Greco, L.A., Reay, W.R., Dayas, C.V., and Cairns, M.J. (2022). Pairwise genetic meta-
- analyses between schizophrenia and substance dependence phenotypes reveals novel
- association signals with pharmacological significance. Transl Psychiatry *12*, 403.
- 1126 10.1038/s41398-022-02186-4.

1127	138.	Minica, C.C., Verweij, K.J.H., van der Most, P.J., Mbarek, H., Bernard, M., van Eijk, K.R.,
1128		Lind, P.A., Liu, M.Z., Maciejewski, D.F., Palviainen, T., et al. (2018). Genome-wide
1129		association meta-analysis of age at first cannabis use. Addiction 113, 2073-2086.
1130		10.1111/add.14368.
1131	139.	Sey, N.Y.A., Hu, B., Mah, W., Fauni, H., McAfee, J.C., Rajarajan, P., Brennand, K.J.,
1132		Akbarian, S., and Won, H. (2020). A computational tool (H-MAGMA) for improved
1133		prediction of brain-disorder risk genes by incorporating brain chromatin interaction
1134		profiles. Nat Neurosci 23, 583-593. 10.1038/s41593-020-0603-0.
1135	140.	Barbeira, A.N., Dickinson, S.P., Bonazzola, R., Zheng, J., Wheeler, H.E., Torres, J.M.,
1136		Torstenson, E.S., Shah, K.P., Garcia, T., Edwards, T.L., et al. (2018). Exploring the
1137		phenotypic consequences of tissue specific gene expression variation inferred from
1138		GWAS summary statistics. Nat Commun 9, 1825. 10.1038/s41467-018-03621-1.
1139	141.	Bulik-Sullivan, B.K., Loh, P.R., Finucane, H.K., Ripke, S., Yang, J., Schizophrenia
1140		Working Group of the Psychiatric Genomics Consortium, Patterson, N., Daly, M.J.,
1141		Price, A.L., and Neale, B.M. (2015). LD Score regression distinguishes confounding from
1142		polygenicity in genome-wide association studies. Nat Genet 47, 291-295.
1143		10.1038/ng.3211.
1144	142.	All of Us Research Program Genomics Investigators (2024). Genomic data in the All of
1145		Us Research Program. Nature 627, 340-346. 10.1038/s41586-023-06957-x.
1146	143.	All of Us Research Program Investigators, Denny, J.C., Rutter, J.L., Goldstein, D.B.,
1147		Philippakis, A., Smoller, J.W., Jenkins, G., and Dishman, E. (2019). The "All of Us"
1148		Research Program. N Engl J Med 381, 668-676. 10.1056/NEJMsr1809937.
1149	144.	Ge, T., Chen, C.Y., Ni, Y., Feng, Y.A., and Smoller, J.W. (2019). Polygenic prediction via
1150		Bayesian regression and continuous shrinkage priors. Nat Commun 10, 1776.

1151 10.1038/s41467-019-09718-5.

1152	145.	Lee, S.H., Goddard, M.E., Wray, N.R., and Visscher, P.M. (2012). A better coefficient of
1153		determination for genetic profile analysis. Genet Epidemiol 36, 214-224.
1154		10.1002/gepi.21614.
1155	146.	Roden, D.M., Pulley, J.M., Basford, M.A., Bernard, G.R., Clayton, E.W., Balser, J.R.,
1156		and Masys, D.R. (2008). Development of a large-scale de-identified DNA biobank to
1157		enable personalized medicine. Clin Pharmacol Ther 84, 362-369. 10.1038/clpt.2008.89.
1158	147.	Dennis, J., Sealock, J., Levinson, R.T., Farber-Eger, E., Franco, J., Fong, S., Straub, P.,
1159		Hucks, D., Song, W.L., Linton, M.F., et al. (2021). Genetic risk for major depressive
1160		disorder and loneliness in sex-specific associations with coronary artery disease. Mol
1161		Psychiatry 26, 4254-4264. 10.1038/s41380-019-0614-y.