Supplementary Table 1. Distribution of psychiatric diagnoses in the sample

Number of individuals among CUD cases and the control group having a diagnosis of one or more of five major psychiatric disorders studied in iPSYCH.

Disorder	CUD cases	% of cases	controls	% controls
Schizophrenia	569	23.84	1,789	3.65
Bipolar disorder	188	7.88	1,176	2.40
Autism spectrum disorder	153	6.41	8,698	17.76
Attention deficit hyperactivity disorder	926	38.79	9,701	19.80
Major depression disorder	1,134	47.51	16,393	33.47
Individuals without any of the above disorders	97	4.06	16,269	33.21

Note: numbers are after QC (genetic outliers and related individuals are excluded) numbers does not represent unique individuals, meaning that individuals could count more than once if diagnosed with more than one psychiatric disorder in combination with CUD.

Supplementary Table 2. Distribution of CUD cases across CUD diagnosis sub-categories

Sub-categories	Number of cases within diagnostic sub-categories	Number of cases having only the specific sub-diagnosis
F12.1	877	500
F12.2	1633	1139
F12.5	341	114
F12.7	25	6
F12.8	36	15
F12.9	115	51

Note: numbers are after QC (genetic outliers and related individuals are excluded)

Supplementary Table 3. Distribution of individuals with psychiatric disorders over CUD diagnostic sub-categories

Distribution of comorbid psychiatric disorders (schizophrenia (SZ), bipolar disorder (BD), autism spectrum disorder (ASD), attention-deficit/hyperactivity disorder (ADHD), major depressive disorder (MDD)) among CUD cases across CUD diagnostic sub-categories. The counts across subcategories for individuals without a psychiatric disorder is given as well (CO). For each category counts of individuals are listed and the percentage of the counts in relation to the total N of each comorbid psychiatric disorder.

	CUD+SZ	% CUD+SZ	CUD+BD	% CUD+BD	CUD+ASD	% CUD+ASD	CUD+ADHD	% CUD+ADHD	CUD+MDD	% CUD+MDD	CUD+CO	% CUD+CO	All cases	% Cases
Total N	**569**		**188**		**153**		**926**		**1134**		**97**		**2387**	
F12.1	229	40.246	85	45.213	53	34.641	302	32.613	440	38.801	31	31.959	877	36.741
F12.2	409	71.880	129	68.617	106	69.281	682	73.650	759	66.931	60	61.856	1633	68.412
F12.5	169	29.701	37	19.681	18	11.765	99	10.691	137	12.081	13	13.402	341	14.286
F12.7	7	1.230	2	1.064	1	0.654	11	1.188	11	0.970	1	1.031	25	1.047
F12.8	11	1.933	5	2.660	2	1.307	12	1.296	18	1.587	2	2.062	36	1.508
F12.9	28	4.921	10	5.319	9	5.882	39	4.212	63	5.556	2	2.062	115	4.818

Note the numbers are counts of individuals (after QC) in each diagnosis sub-category having CUD and a psychiatric diagnosis or no psychiatric diagnosis. Individuals might occur in more than one category if they have more than one CUD diagnosis and/or more than one psychiatric diagnosis.

Supplementary Table 4. CUD polygenic risk score analysis in deCODE samples

Association of CUD polygenic risk scores in deCODE samples with CUD (2,257 cases and 145,069 controls), schizophrenia (SCZ, 729 cases and 153,007 controls) and bipolar disorder (BD, 1,751 cases and 151,923 controls). PRS was estimated using the iPSYCH CUD GWAS as discovery (2,387 individuals with CUD and 48,985 controls) using different P-value thresholds in the discovery GWAS (Threshold). The effect is given in odds ratio per standard deviation PRS (Effect) and the variance explained by Nagelkerke's R^2 is given in percentages (Rsq(%)). Two-sided P-values (Pval) are from linear regression. Values in bold represents significance after Bonferroni correction (P=0.003).

Training GWAS	Threshold	Phenotype, deCODE	Pval	Effect*	Rsq(%)
CUD iPSYCH	**0.01**	**CUD**	**2.17×10^{-09}**	**1.173**	**0.253**
CUD iPSYCH	0.01	SCZ	7.01×10^{-02}	1.084	0.053
CUD iPSYCH	0.01	BD	1.18×10^{-01}	1.044	0.018
CUD iPSYCH	**0.03**	**CUD**	**3.03×10^{-09}**	**1.172**	**0.249**
CUD iPSYCH	0.03	SCZ	1.16×10^{-01}	1.073	0.040
CUD iPSYCH	0.03	BD	1.08×10^{-01}	1.046	0.019
CUD iPSYCH	**0.1**	**CUD**	**3.57×10^{-09}**	**1.171**	**0.246**
CUD iPSYCH	0.1	SCZ	1.36×10^{-01}	1.069	0.036
CUD iPSYCH	0.1	BD	1.04×10^{-01}	1.046	0.020
CUD iPSYCH	**0.3**	**CUD**	**3.71×10^{-09}**	**1.171**	**0.246**
CUD iPSYCH	0.3	SCZ	1.40×10^{-01}	1.068	0.035
CUD iPSYCH	0.3	BD	1.02×10^{-01}	1.046	0.020
CUD iPSYCH	**1**	**CUD**	**3.84×10^{-09}**	**1.171**	**0.245**
CUD iPSYCH	1	SCZ	1.43×10^{-01}	1.068	0.034
CUD iPSYCH	1	BD	1.02×10^{-01}	1.046	0.020

Supplementary Table 5. Association of previously identified cannabis risk variants

Association two-sided P-values (P-value) from logistic regression and odds ratio (OR) with respect to affect allele (A1) in the present study (2,387 individuals with CUD and 48,985 controls) of index variants passing the threshold for genome-wide significance ($P=5x10^{-8}$) in previous GWAS studies of cannabis related phenotypes.

Variant	Phenotype (previous GWAS)	Previous GWAS	Sample size (Previous GWAS)	A1 (current study)	OR (current study)	P-value (current study)
rs143244591	Cannabis criterion count	Sherva et al.[1]	3,394 cases 12,861 controls	NA	NA (MAF < 0.01 in Europeans)	NA
rs77378271	Cannabis criterion count	Sherva et al.[1]	3,394 cases 12,861 controls	A	0.96	0.46
rs146091982	Cannabis criterion count	Sherva et al.[1]	3,394 cases 12,861 controls	NA	NA (monomorphic in Europeans)	NA
rs1409568	Cannabis dependence vs cannabis exposed controls	Agrawal et al.[2]	2,080 cases 6435 controls	T	0.97	0.72
rs2875907	Lifetime cannabis use	Pasman et al.[3]	184,765 (22-43% users)	A	1.05	0.11
rs1448602	Lifetime cannabis use	Pasman et al.[3]	184,765 (22-43% users)	A	0.98	0.49
rs7651996	Lifetime cannabis use	Pasman et al.[3]	184,765 (22-43% users)	T	1.05	0.16
rs10085617	Lifetime cannabis use	Pasman et al.[3]	184,765 (22-43% users)	A	1.00	0.95
rs9773390	Lifetime cannabis use	Pasman et al.[3]	184,765 (22-43% users)	T	1.06	0.31
rs9919557	Lifetime cannabis use	Pasman et al.[3]	184,765 (22-43% users)	T	0.96	0.17
rs10499	Lifetime cannabis use	Pasman et al.[3]	184,765 (22-43% users)	A	1.02	0.51
rs17761723	Lifetime cannabis use	Pasman et al.[3]	184,765 (22-43% users)	T	1.01	0.80

NA = information not available

Supplementary Table 6. Tissues and number of genes tested using PrediXcan

Overview of tissues and number of genes included in the analysis of imputed genetically regulated expression using PrediXcan[4]. The analysis is based on 2,387 individuals with CUD and 48,985 controls, two sided P-value from logistic regression. Result in bold is significant after correction for multiple testing ($P=3.8x10^{-6}$).

Tissue	**Number of genes tested**	**Direction of expression in cases compared to controls (Beta)**	**P-value for association of *CHRNA2* imputed expression with cannabis use disorder**
Dorsolateral prefrontal cortex	10,929	-1.210	5.912×10^{-4} (SE = 0.35)
Anterior cingulate cortex	2,658		NA
Caudate basal ganglia	3,620		NA
Cerebellar Hemisphere	4,234	-0.121	5.304×10^{-3} (SE = 0.04)
Cerebellum	**4,836**	**-0.210**	**2.713×10^{-6} (SE = 0.05)**
Cortex	3,586		NA
Frontal cortex	3,212		NA
Hippocampus	2,537		NA
Hypothalamus	2,459		NA
Nucleus accumbens basal ganglia	3,133		NA
Putamen basal ganglia	2,810		NA

NA = results not available, due to no valid model for gene expression in the tissue.

Supplementary Table 7. Association of cannabis risk locus with smoking behaviour

Association of the risk locus for cannabis use disorder in the large GWAS meta-analyses of smoking behaviour in 1.2 mill people performed by Liu et al[5]. Index SNPs from the meta-analysis located in the risk locus (LD region of the CUD risk variant rs56372821) are shown. Regression beta (Beta) for the effect of the risk allele is given, together with the association P-value (P) and standard error (SE) and LD with the CUD risk variant rs56372821 (R^2). Two side P-values from logistic regression (Smoking cessation and smoking initiation) and linear regression (cigarettes per day and age at smoking initiation).

Phenotype	Sample size	Variant	Ref allele	Risk allele	Pvalue	Beta	SE	N	R^2
Age at smoking initiation	341,427	rs11780471	G	A	$9.44x10^{-11}$	0.033	0.005	341,427	0.288
Cigarettes per day	335,394	rs73229090	C	A	$8.09x10^{-09}$	0.016	0.003	337,334	0.701
Smoking cessation	547,219	rs1565735	T	A	$1.54x10^{-12}$	-0.035	0.005	547,219	0.727
Smoking initiation	1,232,091	rs11783093	C	T	$2.07x10^{-41}$	-0.047	0.003	1,232,091	1.000

Supplementary Table 8. mtCOJO results for the CUD risk locus

Results from mtCOJO analysis[6], which adjust the association P-values from the CUD GWAS for the potential confounding of smoking by using smoking associated SNPs from GWAS of smoking initiation[5]. P-value thresholds for inclusion of variants from smoking GWAS (P-thres) are given together with the number of variants used for the estimation of the effect of smoking on the risk of CUD (N-snps). The original two-sided CUD GWAS P-values (P) and OR (OR) and corrected OR (OR_corrected) and two-sided P-values (P_corrected) are presented. The results for the CUD index variant (rs56372821) and the smoking initiation index variant (rs11783093) are shown.

SNP	P-thres	N-snps	A1	A2	OR	P	OR_corrected	P_corrected
rs56372821	$5x10^{-8}$	61	A	G	0.728	9.3059x10-12	0.759	4.443x10-09
rs11783093	$5x10^{-8}$	61	T	C	0.728	1.089x10-11	0.759	4.494x10-09

Supplementary Table 9. Distribution of individuals with respect to rs56372821 genotypes

In the table below individuals having one or two minor alleles of rs56372821 are grouped together within each unique phenotype (schizophrenia (SZ), bipolar disorder (BP), autism spectrum disorder (ASD), attention deficit hyper activity disorder (ADHD), major depressive disorder (MDD), controls (ICON)), and the number of individuals with a comorbid CUD diagnosis among each group in parentheses. Furthermore, the table has the maximum likelihood estimates of the proportion π_1 of CUD within carriers of one or two minor alleles, and the proportion π_2 of CUD within carriers of two major alleles, assuming the worst-case scenario of an expected odds ratio 0.954.

	SZ (CUD)	BP (CUD)	ASD (CUD)	ADHD (CUD)	MDD (CUD)	ICON (CUD)
At least one minor allele	624 (128)	356 (32)	2476 (25)	2466 (185)	4182 (153)	4792 (24)
major allele	1715 (440)	902 (128)	6066 (103)	6064 (533)	9858 (529)	11394 (71)
π_1	0.23653	0.12347	0.01450	0.08161	0.04706	0.00568
π_2	0.24514	0.12865	0.01518	0.08521	0.04922	0.00595

Supplementary Table 10. CUD polygenic risk score analysis of smoking in deCODE

Association (two-sided P-value from linear regression) of CUD polygenic risk scores in deCODE samples with CUD and tobacco smoking (EverSmoker). PRS was estimated using the iPSYCH CUD GWAS as discovery using a P-value threshold of P=0.01 in the discovery GWAS. The effect is given in odds ratio per standard deviation PRS (OR) and the variance explained by Nagelkerke's R^2 (in percentages) is reported on the observed scale ($R^2_{observed}$) and on the liability ($R^2_{liability}$) scale using the conversion recommended by Lee et al.[7]

PRS	Pheno	Cases	Controls	$R^2_{observed}$ (%)	$R^2_{liability}$(%)	OR	SE	P-value
CUD	CUD	2,257	145,069	0.246	0.337*	1.167	0.026	$2.613x10^{-09}$
CUD	EverSmoker	46,941	20,602	0.104	0.107*	1.066	0.010	$9.320x10^{-10}$

*In the conversion to R^2 on the liability scale a prevalence of 0.01 was used for CUD and 0.17 for smoking.

Supplementary Table 11. Polygenic risk score analysis

Results from PRS analysis in CUD for phenotypes related to cognition, personality, psychiatric disorders, reproduction and smoking behavior. Association P-value of PRS with CUD (P-value). The P-value threshold in the discovery GWAS (P-value threshold), the Z-score from logistic regression (Z-score) and Nagelkerke's R^2 (R^2) as estimate of the percentage of the variance in the phenotype explained by PRS. Nagelkerke's R^2 on the liability ($R^2_{liability}$) scale estimated using the conversion recommended by Lee et al.[7] Two-sided P-values from logistic regression. Values in bold represents significance after Bonferroni correction.

Phenotype/GWAS study	Trait type	Sample size	Z score	P-value	R^2	$R^2_{liability}$	Threshold
Attainment of college/university degree (UK Biobank)[8]	Cognition/education	111,114	-2.81	4.94×10^{-03}	4.6×10^{-04}	4.5×10^{-04}	0.5
Memory (UK Biobank)[8]	Cognition/education	112,067	1.23	2.18×10^{-01}	8.9×10^{-05}	8.7×10^{-05}	1×10^{-03}
Reaction time (UK Biobank)[8]	Cognition/education	111,483	1.96	5.00×10^{-02}	2.2×10^{-04}	2.1×10^{-04}	0.1
verbal-numerical reasoning (UK Biobank)[8]	Cognition/education	36,035	1.43	1.53×10^{-01}	1.2×10^{-04}	1.2×10^{-04}	0.2
Human Intelligence[9]	Cognition/education	78,308	-3.51	**4.33×10^{-04}**	7.2×10^{-04}	7.0×10^{-04}	0.5
Number of educational years (SSGAC)[10]	Cognition/education	328,917	-7.78	**7.44×10^{-15}**	3.5×10^{-03}	3.4×10^{-03}	5×10^{-02}
College completion (SSGAC)[11]	Cognition/education	126,559	-5.10	**3.33×10^{-07}**	1.5×10^{-03}	1.5×10^{-03}	1×10^{-02}
Openness (GPC-1)[12]	Personality	17,375	1.47	1.43×10^{-01}	1.3×10^{-04}	1.3×10^{-04}	1×10^{-04}
Neuroticism (GPC-2)[13]	Personality	63,661	-1.95	5.02×10^{-02}	2.2×10^{-04}	2.1×10^{-04}	1×10^{-03}
Extraversion (GPC-2)[14]	Personality	63,030	-0.52	6.03×10^{-01}	1.6×10^{-05}	1.6×10^{-05}	1×10^{-04}
Agreeableness (GPC-1)[12]	Personality	17,375	-1.73	8.23×10^{-02}	1.8×10^{-04}	1.8×10^{-04}	1×10^{-02}
Conscientiousness (GPC-1)[12]	Personality	17,375	-2.08	3.74×10^{-02}	2.5×10^{-04}	2.4×10^{-04}	1×10^{-04}
ADHD (iPSYCH+PGC)[15]	Psychiatric*	53,293	5.10	**3.45×10^{-07}**	2.3×10^{-03}	2.3×10^{-03}	0.5
Schizophrenia (PGC)[16]	Psychiatric*	150,064	5.47	**4.45×10^{-08}**	2.2×10^{-03}	2.3×10^{-03}	0.5
Major depressive disorder (PGC)[17]	Psychiatric*	18,759	-0.55	5.81×10^{-01}	3.1×10^{-05}	3.0×10^{-05}	1×10^{-03}
Depressive symptoms (SSGAC)[18]	Psychiatric*	161,460	4.34	**6.58×10^{-06}**	2.1×10^{-03}	2.5×10^{-03}	0.1
Age of first birth (SSGAC)[19]	Reproduction	251,151	-7.41	**1.26×10^{-13}**	3.2×10^{-03}	3.1×10^{-03}	0.2
Number of children ever born (SSGAC)[19]	Reproduction	343,072	2.81	5.01×10^{-03}	4.6×10^{-04}	4.5×10^{-04}	1.0
Smoking age of onset (TAG)[20]	Smoking behaviour	24,114	1.17	2.43×10^{-01}	7.9×10^{-05}	7.7×10^{-05}	1×10^{-03}
life-time smoking (TAG)[20]	Smoking behaviour	74,035	4.01	**6.08×10^{-05}**	9.4×10^{-04}	9.2×10^{-04}	0.1
Current vs former smoker (TAG)[20]	Smoking behaviour	41,969	-1.19	2.33×10^{-01}	8.3×10^{-05}	8.1×10^{-05}	1
Cigarettes per day (TAG)[20]	Smoking behaviour	68,028	1.75	8.01×10^{-02}	1.8×10^{-04}	1.8×10^{-04}	1×10^{-03}

* In PRS analyses of psychiatric disorders, individuals having a diagnosis of the respective psychiatric disorder being analysed (ADHD, schizophrenia, major depressive disorder and depressive symptoms) were excluded in the CUD target sample.

References

1 Sherva, R. *et al.* Genome-wide Association Study of Cannabis Dependence Severity, Novel Risk Variants, and Shared Genetic Risks. *JAMA Psychiatry* **73**, 472-480, doi:10.1001/jamapsychiatry.2016.0036 (2016).

2 Agrawal, A. *et al.* Genome-wide association study identifies a novel locus for cannabis dependence. *Molecular psychiatry*, doi:10.1038/mp.2017.200 (2017).

3 Pasman, J. A. *et al.* GWAS of lifetime cannabis use reveals new risk loci, genetic overlap with psychiatric traits, and a causal influence of schizophrenia. *Nature neuroscience* **21**, 1161-1170, doi:10.1038/s41593-018-0206-1 (2018).

4 Gamazon, E. R. *et al.* A gene-based association method for mapping traits using reference transcriptome data. *Nature genetics* **47**, 1091-1098, doi:10.1038/ng.3367 (2015).

5 Liu, M. *et al.* Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nature genetics*, doi:10.1038/s41588-018-0307-5 (2019).

6 Zhu, Z. *et al.* Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nature communications* **9**, 224, doi:10.1038/s41467-017-02317-2 (2018).

7 Lee, S. H., Goddard, M. E., Wray, N. R. & Visscher, P. M. A better coefficient of determination for genetic profile analysis. *Genet Epidemiol* **36**, 214-224, doi:10.1002/gepi.21614 (2012).

8 Davies, G. *et al.* Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N=112 151). *Molecular psychiatry* **21**, 758-767, doi:10.1038/mp.2016.45 (2016).

9 Sniekers, S. *et al.* Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. *Nature genetics* **49**, 1107-1112, doi:10.1038/ng.3869 (2017).

10 Okbay, A. *et al.* Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* **533**, 539-542, doi:10.1038/nature17671 (2016).

11 Rietveld, C. A. *et al.* GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science* **340**, 1467-1471, doi:10.1126/science.1235488 (2013).

12 de Moor, M. H. *et al.* Meta-analysis of genome-wide association studies for personality. *Molecular psychiatry* **17**, 337-349, doi:10.1038/mp.2010.128 (2012).

13 Genetics of Personality, C. *et al.* Meta-analysis of Genome-wide Association Studies for Neuroticism, and the Polygenic Association With Major Depressive Disorder. *JAMA Psychiatry* **72**, 642-650, doi:10.1001/jamapsychiatry.2015.0554 (2015).

14 van den Berg, S. M. *et al.* Meta-analysis of Genome-Wide Association Studies for Extraversion: Findings from the Genetics of Personality Consortium. *Behav Genet* **46**, 170-182, doi:10.1007/s10519-015-9735-5 (2016).

15 Demontis, D. *et al.* Discovery Of The First Genome-Wide Significant Risk Loci For ADHD. *bioRxiv*, doi:10.1101/145581 (2017).

16 Schizophrenia Working Group of the Psychiatric Genomics, C. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421-427, doi:10.1038/nature13595 (2014).

17 Major Depressive Disorder Working Group of the Psychiatric, G. C. *et al.* A mega-analysis of genome-wide association studies for major depressive disorder. *Molecular psychiatry* **18**, 497-511, doi:10.1038/mp.2012.21 (2013).

18 Okbay, A. *et al.* Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nature genetics* **48**, 624-633, doi:10.1038/ng.3552 (2016).

19 Barban, N. *et al.* Genome-wide analysis identifies 12 loci influencing human reproductive behavior. *Nature genetics* **48**, 1462-1472, doi:10.1038/ng.3698 (2016).
20 Tobacco & Genetics, C. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nature genetics* **42**, 441-447, doi:10.1038/ng.571 (2010).