SUPPLEMENTARY METHODS

Cohorts

Alcohol Dependence in African Americans (ADAA) data were collected between 2009 and 2013 and consisted of cases recruited from treatment centers in St. Louis Missouri and controls screened for the absence of substance use disorders recruited from households selected from neighborhoods in proximity to neighborhoods of residence of case participants. Cases met criteria for DSM-IV opioid dependence. Controls were opioid-exposed but did not meet criteria for opioid dependence (DSM-IV).

Australian Alcohol and Nicotine Studies (OZAL) study recruited participants from twins and their relatives who had participated in questionnaire- and interview-based studies on alcohol and nicotine use and alcohol-related events or symptoms[1]. They were living in Australia and of predominantly European ancestry Opioid dependence was defined using DSM-IV criteria.

Center on Antisocial Drug Dependence (CADD) cohort includes unrelated participants aggregated from several studies described elsewhere[2-5]. This cohort was over-selected for adolescent behavioral disinhibition, with half of the participants ascertained specifically from high-risk populations (i.e. recruited through substance abuse treatment, special schools, or involvement with the criminal justice system; see supplement of 19 for additional criteria for clinical probands). Lifetime

opioid dependence was assessed with the CIDI-SAM and defined as meeting opioid dependence at any wave for this longitudinal study.

Collaborative Study on the Genetics of Alcoholism (COGA) is a multi-site study of alcohol dependent probands and their family members. Alcohol dependent probands were recruited from inpatient and outpatient facilities. Community probands and their family members were also recruited from a variety of sources. Further details were described previously [6]. All participants were assessed using the Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA). Cases met criteria for a lifetime history of DSMIV opioid dependence. The exposed controls were defined as those who were exposed at least once in their lifetime to opioids but did not meet criteria for opioid dependence.

Comorbidity and Trauma Study (CATS) consisted of opioid dependent individuals aged 18 and older recruited from opioid substitution therapy clinics in the greater Sydney area and genetically unrelated individuals with little or no lifetime opioid misuse from neighborhoods in geographic proximity to these clinics. All subjects were of European-Australian descent. Further details were described previously[7]. All participants were assessed using the SSGA. Opioid dependence was defined using DSM-IV criteria. The exposed controls were defined as those who were exposed at least once in their lifetime to opioids but did not meet criteria for opioid dependence at the time of their last interview.

Gene-Environment-Development Initiative – the Great Smoky Mountains Study (GSMS) is prospective study funded by the National Institute on Drug Abuse (NIDA). The Participants were assessed via structured interviewing using the Young

Adult Psychiatric Assessment and its early life extension (i.e., YAPA and CAPA), yielding diagnoses and symptom scales for a wide range of substance use disorders (SUDs). Opioid dependence was defined using DSM-IV criteria. Exposed controls were defined as those who were exposed to opioids at least once but did not meet criteria for opioid dependence.

Gene-Environment-Development Initiative – Virginia Commonwealth University (VCUI) study combined existing phenotypic and environmental data from the Virginia Twin Study of Adolescent Behavioral Development (VTSABD) study, a population-based multi-wave, cohort-sequential twin study of adolescent psychopathology and its risk factors, and two followup studies, the Young Adult Follow Up (YAFU) and the Transitions to Substance Abuse (TSA) study. Further details of the VCUI cohort were reported previously[8, 9]. Participants were assessed via structured interviewing using the Child Adult Psychiatric Assessment (CAPA), a Structured Clinical Interview for DSM-IV (SCID)-based assessment of psycho-pathology in young adult twins for YAFU and the Life Experiences Interview (LEI) for TSA, yielding diagnoses and symptom scales for a wide range of substance use disorders (SUDs). Opioid dependence was defined using DSM-IV criteria.

Study of Addiction: Genetics and Environment (SAGE), Collaborative Genetic Study of Nicotine Dependence (COGEND) and Family Study of Cocaine Dependence (FSCD) are selected from three large, complementary studies focused of the genetics of substance use disorders[10-12]. We analyze these subsets separately and remove overlap between cohorts. SAGE participants were assessed using the SSAGA. FSCD and COGEND participants were assessed using

polydiagnostic instruments closely based on the SSAGA. Cases reported a lifetime history of DSM-IV opioid dependence. Genetically unrelated control subjects reported to be exposed to opioids at least once in their lifetime but had no significant opioid-dependence symptoms.

Yale-Penn (YP) study includes participants recruited in the eastern US, predominantly in Connecticut and Pennsylvania. They were administered the Semi-Structured Assessment for Drug Dependence and Alcoholism (SSADDA) to derive DSM-IV diagnoses of lifetime substance use disorders and other major psychiatric traits). The study received IRB approval from all participating institutions and written informed consent was obtained from all study participants. Additional information is available in the previous GWAS publications[13-17]. Opioid dependence was defined using DSM-IV criteria.

Inflation and Deflation in the cohorts investigated

Genomic inflation factor (λ_{GC}) and LD score regression intercept were calculated to investigate the presence of the deviations of the genome-wide association results from the expected distribution. Inflation (λ_{GC}>1.04) could be due to the effect of unaccounted population stratification and polygenicity[18]. To distinguish polygenicity from unaccounted population stratification, we calculated the LD score regression intercept in the datasets tested[19]. Additionally, we observed in some of the cohorts investigated a strong deflation (λ_{GC}<0.9), which is due to the low number of cases and the small case-control ratio present in some of the cohorts. Accordingly, we excluded NICO, OZAL, SAGE, and YP2 (AFR) from the $OD_{unexposed}$ meta-analysis due to the deflation observed.

Supplementary Table 1: Sample size of the opioid-informative cohorts investigated.

Cohort		Ancestry	Cases	Exposed Controls	Unexposed Controls
case-control studies: logistic regression					
COGEND Study of Addiction: Genetics and Environment (SAGE)		EUR	22	73	966
Collaborative Study on the Genetics of Nicotine Dependence (NICO)		EUR	15	117	794
Center on Antisocial Drug Dependence (CADD)		EUR	44	386	666
Gene–Environment Development Initiative: Great Smoky Mountains Study (GSMS)		EUR	21	151	615
Family Study of Cocaine Dependence (FSCD)		AFR	57	66	477
		EUR	81	98	361
Comorbidity and Trauma Study (CATS)		EUR	1025	95	291
Alcohol Dependence in African Americans (ADAA)		AFR	241	177	1393
Family-based studies: logistic mixed model					
Yale-Penn	Phase1 (YP1)	AFR	626	580	1840
		EUR	1030	295	426
	Phase2 (YP2)	AFR	204	214	952
		EUR	642	210	685
Gene–Environment Development Initiative: Virginia Commonwealth University (VCUI)		EUR	48	10	2482
Australian Alcohol and Nicotine Studies (OZAL)		EUR	24	173	12088
Collaborative Study on the Genetics of Alcoholism (COGA)		AFR	103	260	2401
		EUR	320	1268	6063
TOTAL			**4,503**	**4,173**	**32,500**

Supplementary Table 2: Number of variants tested in the ancestry-specific and trans-ancestry meta-analysis across the phenotypes investigated. The association analysis was conducted considering variants present in at least 80% of the cohorts investigated.

Meta-analysis	$OD_{exposed}$	$OD_{unexposed}$	$OE_{controls}$
African-ancestry	8,956,178	8,896,928	9,065,342
European-ancestry	4,211,587	4,887,055	5,986,959
Trans-ancestry	4,000,397	4,010,688	5,122,696

Supplementary Table 3: Associations of the variants identified in the opioid GWAS with respect to phenotypes related to mental and behavioral disorders due to use of alcohol, cannabinoid, and tobacco. PMID: PubMed Identifier; NS: Not significant (p>0.05).

PMID	Phenotype	rs201123820	rs9291211	rs12461856
31427789	Alcohol usually taken with meals	NS	3.02E-08	NS
31427789	Frequency of consuming six or more units of alcohol	NS	2.78E-04	NS
31427789	Alcohol intake frequency	NS	5.91E-04	NS
31427789	Cigarettes per day	NS	3.82E-03	NS
20418890	Age of smoking initiation	NS	NS	NS
31427789	Age stopped smoking	NS	NS	NS
30482948	Alcohol dependence	NS	NS	NS
BioRxiv: 261081	Alcohol intake	NS	NS	NS
31427789	Alcohol intake versus 10 years previously	NS	NS	NS
31427789	Beer/cider intake	NS	NS	NS
31427789	Current tobacco smoking	NS	NS	NS
31427789	Ever had known person concerned about, or recommended reduction of alcohol consumption	NS	NS	NS
30643251	Ever smoked regulary	NS	NS	NS
31427789	Ever taken cannabis	NS	NS	NS
30150663	Ever used cannabis	NS	NS	NS
BioRxiv: 261081	Ever vs never smokers	NS	NS	NS
20418890	Former vs current smokers	NS	NS	NS
31427789	Frequency of feeling guilt or remorse after drinking alcohol in last year	NS	NS	NS
31427789	Frequency of memory loss due to drinking alcohol in last year	NS	NS	NS
31427789	Light smokers	NS	NS	NS
31427789	Past tobacco smoking	NS	NS	NS
31427789	Reason for reducing amount of alcohol drunk: Health precaution	NS	NS	NS

31427789	Reason for reducing amount of alcohol drunk: Other reason	NS	NS	NS
31427789	Red wine intake	NS	NS	NS
30643251	Smoking cessation	NS	NS	NS
31427789	Tobacco smoking	NS	NS	NS
31427789	Why stopped smoking: Health precaution	NS	NS	NS

Supplementary Table 4: Genomic inflation factors (λ_{GC}) across the meta-analyses conducted with respect to different phenotypic definitions and ancestry groups. LD score regression intercept was calculated when $\lambda_{GC} > 1.04$ to distinguish polygenicity from population stratification.

Meta-analysis	**OD vs. OE**	**OD vs. OU**	**OE vs. OU**
African-ancestry	0.994	1.018	0.999
European-ancestry	1.033	1.055 (LDSC intercept=1.033)	1.044 (LDSC intercept=1.020)
Trans-ancestry	1.028	1.061 (LDSC intercept=1.029)	1.031

Supplementary Table 5: Ancestry-specific results for each genome-wide significant variant identified.

Phenotype	Meta-analysis	rsID	Effect Allele	Other Allele	Effect Allele Frequency	Z score	P value
$OD_{unexposed}$	AFR	rs201123820	T	TAAACAAAAACA	0.0193	5.547	2.90E-08
	EUR				0.0402	-0.903	0.367
	TRANS				0.0318	2.472	0.013
$OE_{controls}$	TRANS	rs12461856	A	G	0.8402	-5.606	2.07E-08
	EUR				0.8436	-4.833	1.35E-06
	AFR				0.833	-2.866	0.004
	EUR	rs9291211	A	G	0.7817	-5.387	7.16E-08
	AFR				0.331	0.829	0.407
	TRANS				0.6356	-3.956	7.61E-05

Supplementary Table 6: directions and heterogeneity estimates of the three GWS variants among the cohorts meta-analyzed.

rsID	Effect Allele	Other Allele	Z score	P value	Direction	Heterogeneity			
						I^2	$\square^2$	Df	P value
rs201123820	T	TAAACAAAAACA	5.547	2.90E-08	+++++	0	3.534	4	0.472
rs9291211	A	G	-5.387	7.16E-08	-------?--?	0	3.74	8	0.880
rs12461856	A	G	-5.606	2.07E-08	------??-?----?-	0	9.744	11	0.554

Supplementary Table 7: full summary association data of the phenome-wide scan conducted in the UK biobank [xlsx file attached].

Supplementary Table 8: Results surviving multiple testing correction in the gene-based phenome-wide scan conducted with respect to *SLC30A9, BEND4,* and *SDCCAG8* using data from the GWAS atlas[20] [xlsx file attached]. We did not observe association surviving multiple testing correction for *C18orf32* gene.

Supplementary Table 9: Association and heterogeneity of the polygenic risk scores tested among the cohorts meta-analyzed.

PRS	Test	Z score	P value	FDR Q value	I^2	Heterogeneity P value
Risk-Taking	OD vs OU (PT=1)	3.94	8.1E-05	0.003	0	0.681
	OE vs OU (PT=0.05)	3.57	3.6E-04	0.003	38.9	0.090
	OD vs OE (PT=1)	1.93	0.054	0.133	0	0.487
Neuroticism	OD vs OU (PT=0.001)	4.16	3.22E-05	0.001	0	0.960
	OE vs OU (PT=0.5)	-0.43	0.6706	0.919	0	0.812
	OD vs OE (PT=0.001)	3.10	0.002	0.016	0	0.665

Supplementary Table 10: Results in the present study related to variants identified in prior opioid dependence GWAS. We report the results of the variants that survived the quality control criteria applied to the different meta-analysis (e.g., information present in at least 80% of the cohorts meta-analyzed; imputation info score > 0.8; minor allele frequency > 0.01).

Phenotype	Ancestry	rsID (Gene, Reference)	Effect Allele	Other Allele	Effect Allele Frequency	Z score	P value
$OE_{controls}$	TRANS	rs10494334 (Intergenic,[21])	A	G	0.1035	0.439	0.6605
$OE_{controls}$	EUR				0.0677	0.336	0.7367
$OD_{exposed}$	AFR				0.1851	-0.23	0.818
$OE_{controls}$	TRANS	rs12442183 (*RGMA*,[22])	T	C	0.3793	-1.709	0.08753
$OE_{controls}$	EUR				0.408	-1.819	0.06886
$OD_{exposed}$	AFR				0.3202	1.043	0.2968
$OE_{controls}$	TRANS	rs1436175 (*CNIH3*,[7])	A	G	0.4638	0.049	0.9609
$OE_{controls}$	EUR				0.4092	0.085	0.9325
$OD_{exposed}$	AFR				0.5901	0.269	0.7881
$OD_{exposed}$	TRANS				0.4861	-1.827	0.0677
$OD_{exposed}$	EUR				0.4264	-2.496	0.01256
$OD_{unexposed}$	TRANS				0.4771	-0.786	0.4319
$OD_{unexposed}$	EUR				0.4054	-0.869	0.3847
$OE_{controls}$	TRANS	rs62103177 (*KCNG2*,[13])	A	G	0.1188	0.988	0.323
$OE_{controls}$	EUR				0.1436	0.5	0.6173
$OD_{exposed}$	AFR				0.0725	-1.625	0.1043
$OD_{exposed}$	TRANS				0.1159	-2.09	0.03664
$OD_{exposed}$	EUR				0.1456	-1.368	0.1712
$OD_{unexposed}$	TRANS				0.1165	-1.666	0.09574
$OD_{unexposed}$	EUR				0.1458	-1.557	0.1195

Supplementary Figure 1: Manhattan plots from the gene-based GWAS meta-analysis of $OD_{unexposed}$ phenotype in African-ancestry individuals (**A**); $OE_{controls}$ phenotypes in European-ancestry individuals (**B**) and in the trans-ancestry meta-analysis (**C**) [tiff file attached].

Supplementary Figure 2: Multi-tissue eQTL results of rs9291211 with respect to *SLC30A9* (**A**) and *BEND4* (**B**) transcriptomic profiles [tiff file attached].

References

1. Heath AC, Whitfield JB, Martin NG, Pergadia ML, Goate AM, Lind PA *et al.* A quantitative-trait genome-wide association study of alcoholism risk in the community: findings and implications. *Biol Psychiatry* 2011; **70**(6)**:** 513-518.

2. Hartman CA, Gelhorn H, Crowley TJ, Sakai JT, Stallings M, Young SE *et al.* Item response theory analysis of DSM-IV cannabis abuse and dependence criteria in adolescents. *J Am Acad Child Adolesc Psychiatry* 2008; **47**(2)**:** 165-173.

3. Rhea SA, Gross AA, Haberstick BC, Corley RP. Colorado Twin Registry. *Twin Res Hum Genet* 2006; **9**(6)**:** 941-949.

4. Stallings MC, Corley RP, Dennehey B, Hewitt JK, Krauter KS, Lessem JM *et al.* A genome-wide search for quantitative trait Loci that influence antisocial drug dependence in adolescence. *Arch Gen Psychiatry* 2005; **62**(9)**:** 1042-1051.

5. Derringer J, Corley RP, Haberstick BC, Young SE, Demmitt BA, Howrigan DP *et al.* Genome-Wide Association Study of Behavioral Disinhibition in a Selected Adolescent Sample. *Behav Genet* 2015; **45**(4)**:** 375-381.

6. Lai D, Wetherill L, Bertelsen S, Carey CE, Kamarajan C, Kapoor M *et al.* Genome-wide association studies of alcohol dependence, DSM-IV criterion count and individual criteria. *Genes Brain Behav* 2019; **18**(6)**:** e12579.

7. Nelson EC, Agrawal A, Heath AC, Bogdan R, Sherva R, Zhang B *et al.* Evidence of CNIH3 involvement in opioid dependence. *Mol Psychiatry* 2016; **21**(5)**:** 608-614.

8. Costello EJ, Eaves L, Sullivan P, Kennedy M, Conway K, Adkins DE *et al.* Genes, environments, and developmental research: methods for a multi-site study of early substance abuse. *Twin Res Hum Genet* 2013; **16**(2)**:** 505-515.

9. Maes HH, Silberg JL, Neale MC, Eaves LJ. Genetic and cultural transmission of antisocial behavior: an extended twin parent model. *Twin Res Hum Genet* 2007; **10**(1)**:** 136-150.

10. Edenberg HJ. The collaborative study on the genetics of alcoholism: an update. *Alcohol Res Health* 2002; **26**(3)**:** 214-218.

11. Bierut LJ, Strickland JR, Thompson JR, Afful SE, Cottler LB. Drug use and dependence in cocaine dependent subjects, community-based individuals, and their siblings. *Drug Alcohol Depend* 2008; **95**(1-2)**:** 14-22.

12. Bierut LJ, Madden PA, Breslau N, Johnson EO, Hatsukami D, Pomerleau OF *et al.* Novel genes identified in a high-density genome wide association study for nicotine dependence. *Hum Mol Genet* 2007; **16**(1)**:** 24-35.

13. Gelernter J, Kranzler HR, Sherva R, Koesterer R, Almasy L, Zhao H *et al.* Genome-wide association study of opioid dependence: multiple associations mapped to calcium and potassium pathways. *Biol Psychiatry* 2014; **76**(1)**:** 66-74.

14. Gelernter J, Sherva R, Koesterer R, Almasy L, Zhao H, Kranzler HR *et al.* Genome-wide association study of cocaine dependence and related traits: FAM53B identified as a risk gene. *Mol Psychiatry* 2014; **19**(6)**:** 717-723.

15. Gelernter J, Kranzler HR, Sherva R, Almasy L, Koesterer R, Smith AH *et al.* Genome-wide association study of alcohol dependence:significant findings in African- and European-Americans including novel risk loci. *Mol Psychiatry* 2014; **19**(1)**:** 41-49.

16. Gelernter J, Kranzler HR, Sherva R, Almasy L, Herman AI, Koesterer R *et al.* Genome-wide association study of nicotine dependence in American populations: identification of novel risk loci in both African-Americans and European-Americans. *Biol Psychiatry* 2015; **77**(5)**:** 493-503.

17. Sherva R, Wang Q, Kranzler H, Zhao H, Koesterer R, Herman A *et al.* Genome-wide Association Study of Cannabis Dependence Severity, Novel Risk Variants, and Shared Genetic Risks. *JAMA Psychiatry* 2016; **73**(5)**:** 472-480.

18. Yang J, Weedon MN, Purcell S, Lettre G, Estrada K, Willer CJ *et al.* Genomic inflation factors under polygenic inheritance. *Eur J Hum Genet* 2011; **19**(7)**:** 807-812.

19. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics C *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* 2015; **47**(3)**:** 291-295.

20. Watanabe K, Stringer S, Frei O, Umicevic Mirkov M, de Leeuw C, Polderman TJC *et al.* A global overview of pleiotropy and genetic architecture in complex traits. *Nat Genet* 2019; **51**(9)**:** 1339-1348.

21. Kalsi G, Euesden J, Coleman JR, Ducci F, Aliev F, Newhouse SJ *et al.* Genome-Wide Association of Heroin Dependence in Han Chinese. *PLoS One* 2016; **11**(12)**:** e0167388.

22. Cheng Z, Zhou H, Sherva R, Farrer LA, Kranzler HR, Gelernter J. Genome-wide Association Study Identifies a Regulatory Variant of RGMA Associated With Opioid Dependence in European Americans. *Biol Psychiatry* 2018; **84**(10)**:** 762-770.