Leveraging genome-wide data to investigate differences between opioid use *vs.* opioid dependence in 41,176 individuals from the Psychiatric Genomics Consortium

Renato Polimanti[1], Raymond K. Walters[2], Emma C. Johnson[3], Jeanette N. McClintick[4], Amy E. Adkins[5], Daniel E. Adkins[6], Silviu-Alin Bacanu[7], Laura J. Bierut[3], Tim B. Bigdeli[8], Sandra Brown[9], Kathy Bucholz[3], William E. Copeland[10], E. Jane Costello[11], Louisa Degenhardt[12], Lindsay A Farrer[13], Tatiana M. Foroud[14], Louis Fox[3], Alison M. Goate[15], Richard Grucza[3], Laura M. Hack[16], Dana B. Hancock[17], Sarah M. Hartz[3], Andrew C. Heath[3], John K. Hewitt[18], Christian J. Hopfer[19], Eric O. Johnson[17], Kenneth S. Kendler[20], Henry R. Kranzler[21], Ken Krauter[22], Dongbing Lai[14], Pamela A. F. Madden[3], Nicholas G. Martin[23], Hermine H. Maes[20], Elliot C. Nelson[3], Roseann E. Peterson[24], Bernice Porjesz[8], Brien P. Riley[7], Nancy Saccone[25], Michael Stallings[18], Tamara Wall[9], Bradley T. Webb[7], Leah

CONFLICT OF INTEREST

Wetherill[14], **Psychiatric Genomics Consortium Substance Use Disorders Workgroup**[*], **Howard J. Edenberg**[4], **Arpana Agrawal**[3], **Joel Gelernter**[1]

[1]Department of Psychiatry, Yale University School of Medicine, West Haven CT, United States;

[2]Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston MA, United States;

[3]Department of Psychiatry, Washington University School of Medicine, Saint Louis MO, United States;

[4]Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis IN, United States;

[5]Department of Psychology, Virginia Commonwealth University, Richmond VA, United States;

[6]Department of Psychiatry, University of Utah, Salt Lake City UT, United States;

[7]Virginia Commonwealth University Alcohol Research Center, Virginia Institute for Psychiatric and Behavioral Genetics, Richmond, VA, United States;

[8]Department of Psychiatry and Behavioral Sciences, State University of New York Downstate Medical Center, Brooklyn NY, United States;

[9]Department of Psychiatry, University of California San Diego, San Diego CA, United States;

[10]Department of Psychiatry, University of Vermont Medical Center, Burlington VT, United States;

[11]Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham NC, United States;

[12]National Drug and Alcohol Research Centre, University of New South Wales, Sydney NSW, Australia;

[13]Department of Medicine (Biomedical Genetics), Boston University School of Medicine, Boston MA, United States;

[14]Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis IN, United States;

[15]Department of Neuroscience, Icahn School of Medicine at Mount Sinai, New York NY, United States;

[16]Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford CA, United States;

[17]Center for Omics Discovery and Epidemiology, RTI International, Research Triangle Park NC, United States;

[18]Institute for Behavioral Genetics, University of Colorado Boulder, Boulder CO, United States;

[19]Department of Psychiatry, University of Colorado Denver, Aurora CO, United States;

[20]Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond VA, United States;

[21]Center for Studies of Addiction, University of Pennsylvania Perelman School of Medicine, Philadelphia PA, United States;

[22]Department of Molecular, Cellular, and Developmental Biology, University of Colorado Boulder, Boulder CO, United States;

[23]QIMR Berghofer Medical Research Institute, Brisbane QLD, Australia;

[24]Department of Psychiatry, Virginia Commonwealth University, Richmond VA, United States;

[25]Department of Genetics, Washington University School of Medicine, Saint Louis MO, United States

Abstract

To provide insights into the biology of opioid dependence (OD) and opioid use (i.e., exposure, OE), we completed a genome-wide analysis comparing 4,503 OD cases, 4,173 opioid-exposed controls, and 32,500 opioid-unexposed controls, including participants of European and African descent (EUR and AFR, respectively). Among the variants identified, rs9291211 was associated with OE (exposed vs. unexposed controls; EUR z=−5.39, p=7.2×10^{-8}). This variant regulates the transcriptomic profiles of *SLC30A9* and *BEND4* in multiple brain tissues and was previously associated with depression, alcohol consumption, and neuroticism. A phenome-wide scan of rs9291211 in the UK Biobank (N>360,000) found association of this variant with propensity to use dietary supplements (p=1.68×10^{-8}). With respect to the same OE phenotype in the gene-based analysis, we identified *SDCCAG8* (EUR+AFR z=4.69, p=10^{-6}), which was previously associated with educational attainment, risk-taking behaviors, and schizophrenia. In addition, rs201123820 showed a genome-wide significant difference between OD cases and unexposed controls (AFR z=5.55, p=2.9×10^{-8}) and a significant association with musculoskeletal disorders in the UK Biobank (p=4.88×10^{-7}). A polygenic risk score (PRS) based on a GWAS of risk-tolerance (N=466,571) was positively associated with OD (OD vs. unexposed controls, p=8.1×10^{-5}; OD cases vs. exposed controls, p=0.054) and OE (exposed vs. unexposed controls, p=3.6×10^{-5}). A PRS based on a GWAS of neuroticism (N=390,278) was positively associated with OD (OD vs. unexposed controls, p=3.2×10^{-5}; OD vs. exposed controls, p=0.002) but not with OE (p=0.67). Our analyses highlight the difference between dependence and exposure and the importance of considering the definition of controls in studies of addiction.

Keywords

Opioids; Substance; Abuse; Genetics; Genome-wide association study

INTRODUCTION

The prevalence of opioid dependence (OD) is at epidemic levels and significantly affects public health and social and economic well-being. The use of opioid medications for analgesia is common, and opioids are considered a gold standard for pain control. However, they are also highly addictive, and are, along with heroin[1], the leading contributors to the ongoing epidemic of opioid misuse and the high rate of fatal overdoses[2–4].

Understanding the biology of human responses to opioids may lead to effective preventive strategies and treatments to reduce OD and its harmful consequences. Human genetic research has the potential to dissect the basis of inter-individual variability in the response to opioid exposure (i.e., whether an individual develops dependence on opioids). Genome-wide association studies (GWAS) of large cohorts have identified a number of risk loci and molecular pathways involved in the predisposition to numerous psychiatric disorders and behavioral traits[5, 6]. Previous OD GWAS included up to 10,000 participants and identified genome-wide significant (GWS) associations in *KCNG2*, *KCNC1*, *APBB2*, *CNIH3*, and *RGMA*[7–10]. However, there was no consistency across the previous OD GWAS with respect to the individual GWS loci, probably due to the limited statistical power and differences in case and control definitions in the context of polygenic architecture (thousands of causal loci with small effect).

Another key potential contributor to the lack of consistency in findings from prior GWAS is that different study designs were used. The most relevant design variation is related to the assessment of opioid exposure in controls. Two different control definitions have been considered: i) individuals exposed to opioids (OE) medically or illegally who did not develop OD; or ii) individuals without an OD diagnosis who were not assessed for opioid exposure. Although including individuals not exposed to opioids in the control group increases the overall sample size, it also potentially adds noise by including individuals who would have been likely to become OD if exposed, given the highly addictive nature of opioid drugs. Furthermore, exposure to opioids is a behavioral trait *per se*, and likely to be associated with its own specific genetic architecture, which may be different between licit and illicit exposure. Opioid use is rarer than the use of many other substances and it is often observed in individuals affected by severe mental and physical illnesses[11, 12]. Comparisons of OD cases with predominantly unexposed controls is likely to confound genetic risk for exposure to opioids with genetic factors specific to the transition to OD. Indeed, at least one prior smaller GWAS[7] found that comparisons of OD cases to controls with significant exposure and from similar neighborhoods resulted in a GWS finding while comparisons with general population controls did not identify any GWS variants.

We leveraged genotypic and phenotypic information from 41,176 participants from 11 studies that are part of the Psychiatric Genomics Consortium Substance Use Disorder working group (PGC-SUD) to investigate genetic differences between OD cases (n=4,503), OE controls (n=4,173), and opioid-unexposed (OU) controls (n=32,500) using GWAS and polygenic risk score (PRS) analyses. In addition to identifying loci related to OD and OE, we also examined whether OD and OE could be differentiated with respect to their relationship with genetic liability to risk-taking behaviors and negative personality features (i.e., neuroticism), to provide further insights into the genetic architecture underlying opioid use and misuse.

MATERIALS AND METHODS

Study Design

This study leveraged the individual genotypic and phenotypic data available from the cohorts participating in the PGC-SUD workgroup. There is growing support for the idea that

the genetic architecture of substance *exposure* is different from that of substance *dependence*[13, 14]. Based on these previous findings, we hypothesized that OD and OE are biologically different and therefore focused the present opioid study on three association tests: i) difference between opioid dependent (DSM-IV) and opioid exposed; ii) difference between opioid dependent and opioid unexposed, and iii) difference between opioid exposed and opioid unexposed. The results of these analyses provided information regarding the genetic differences between OD and OE.

Cohorts and Phenotype Definitions

Of the 11 studies from the PGC-SUD workgroup, seven were case-control studies and four were family-based studies (Supplementary Table 1; Supplementary Methods). Lifetime OD diagnoses was based on DSM-IV OD criteria[15] and were derived either from clinician ratings or semi-structured interviews. The two control groups included OE controls (individuals without a lifetime OD diagnosis who were exposed to opioids at least once) and OU controls (individuals with no lifetime OD diagnosis who were not exposed to opioids). Lifetime opioid exposure included both licit, prescribed opioids and those used outside appropriate medical care. Some, but not all, studies distinguished between these forms of exposure. This study, which involved the analysis of de-identified data, was approved by the institutional review board (IRB) at Yale University School of Medicine and was conducted in accordance with all relevant ethical regulations. Each contributing study obtained informed consent from participants and ethics approvals for their study protocols from their respective review boards in accordance with applicable regulations.

Quality control and imputation

Individual genotype information was available for each subject. The Ricopili pipeline[16] (https://github.com/Nealelab/ricopili) was used for the QC and imputation of the case-control cohorts. Most family-based cohorts were analyzed with the Picopili pipeline[17] (https://github.com/Nealelab/picopili), which is designed to conduct genome-wide meta-analyses accounting for family structure. The genetic data from the Collaborative Studies on the Genetics of Alcoholism were imputed independently as previously described[18] because of the need in that study to merge data on members of large multiplex families who were genotyped across multiple genotyping arrays.

Details regarding the QC criteria were reported previously[17]. Briefly, after initial sample and variant QC, population outlier samples were excluded, and each retained individual was assigned to a specific ancestry on the basis of the principal components derived from genome-wide data. The 1000 Genomes Project Phase 3 reference panel[19] was used as a reference for the ancestry assignment. Based on genetic information, we identified 9,591 and 31,585 individuals of African and European descent, respectively. Other ancestry groups were not investigated due to the limited number of informative subjects. The final QC criteria included variant filters for call rate, heterozygosity, and departure from Hardy–Weinberg equilibrium expectations (HWE), performed within each ancestry group in each cohort stratified by genotyping array. We also used sample QC filters for cryptic relatedness and for departures from reported pedigree structures. Imputation was performed using SHAPEIT2[20] and IMPUTE2[21], and the 1000 Genomes Project Phase 3 reference panel,

which includes five continental groups[19]. High-quality imputed SNPs were retained for the association analysis, filtering for imputation INFO score > 0.8 and minor allele frequency (MAF) > 0.01 before analysis. After imputation, we tested for duplicated samples and cryptic relatedness among the cohorts analyzed. The association analysis was conducted considering variants present in at least 80% of the cohorts investigated (Supplementary Table 2).

Data analysis

The association analysis was conducted stratifying each cohort by ancestry (i.e., African and European ancestries) and genotyping array. For case–control studies, imputed dosages were entered in a logistic regression. For family-based studies, logistic mixed models were used to analyze hard-called best-guess genotypes. The association analyses were adjusted for sex and the within-ancestry top 10 principal components to account for possible confounding by population stratification. To investigate differences between OE and OD, three phenotype definitions were considered: i) OD cases vs. OE controls ($OD_{exposed}$; n=4,503 and 4,173, respectively); ii) OD cases vs. OU controls ($OD_{unexposed}$; n=4,238 and 17,700, respectively; iii) OE controls vs. OU controls ($OE_{controls}$; n=4,173 and 32,500). As explained in the Supplementary Methods, we removed some of the cohorts from the $OD_{unexposed}$ meta-analysis due to the deflation ($\lambda_{GC}<0.9$) caused by the low number of cases and the small case-control ratio. For each phenotype, meta-analyses of the results across the different cohorts were conducted in METAL with weights proportional to the square-root of the sample size for each study[22]. The effective sample size of each cohort was calculated based on the case-control ratio and the relatedness matrix. Ancestry-specific (African-ancestry and European-ancestry) and trans-ancestry meta-analyses were conducted. Heterogeneity was evaluated across all cohorts and between study designs.

To investigate the loci identified in the individual GWAS further, we performed a phenome-wide scan considering 4,082 traits assessed in up to 361,194 participants from the UK Biobank using previously generated GWAS association summary data[23]. Details regarding QC criteria and GWAS methods of this previous analysis are available at https://github.com/Nealelab/UK_Biobank_GWAS/tree/master/imputed-v2-gwas. Briefly, the association analyses for all phenotypes were conducted using regression models available in Hail (available at https://github.com/hail-is/hail) including the first 20 ancestry principal components, sex, age, age^2, sex×age, and sex×age^2 as covariates. We applied a false discovery rate (FDR) multiple testing correction (q<0.05) to account for the number of variants and phenotypes tested. Additionally, we investigated the associations of the loci we identified here with respect to 27 non-duplicated traits related to mental and behavioral disorders attributable to use of alcohol, cannabis, and tobacco (Supplementary Table 3). This information was derived from large-scale summary association data collected by the GWAS Atlas (available at https://atlas.ctglab.nl/)[24]. We also conducted a gene-based phenome-wide scan across 4,756 available datasets in the GWAS Atlas. A Bonferroni correction accounting for the number of traits tested was applied to this gene-based analysis ($p<1.05\times10^{-5}$).

Linkage disequilibrium (LD) score regression[25] was performed to estimate the heritability explained by common SNPs (h^2_g) in the European-ancestry meta-analysis of case-control

and family-based cohorts. The inclusion of related subjects may affect the LD score regression results due to the residual effect of family structure on the summary association data. To limit this potential confounder, the analyses was limited to variants assessed in more than 80% of the total sample and considering the effective sample size adjusted for both case-control ratio and family structure. The heritability analysis was not conducted on African-specific and trans-ancestry meta-analyses, because LD score regression is not suitable when analyzing GWAS summary data derived from admixed populations[25]. LD score regression analysis was performed considering HapMap3 SNPs[26] and LD scores computed from the 1000 Genomes Project reference for European populations. Conversion of h^2_g estimates from observed scale to liability scale was performed accounting for the difference between population prevalence ($OD_{exposed}=1\%$, $OD_{unexposed}=1\%$, and $OE_{controls}=5\%$) and sample prevalence ($OD_{exposed}=55\%$, $OD_{unexposed}=22\%$, and $OE_{controls}=12\%$).

Gene-based association, enrichment analysis for molecular pathways, Gene Ontologies (GO) annotations and tissue-specific transcriptomic profiles were conducted using the MAGMA tool[27] implemented in the FUMA platform[28]. Information regarding molecular pathways and GO annotations was derived from MsigDB v6.2[29]. Tissue-specific transcriptomic profiles were derived from GTEx V7[30] and BrainSpan[31]. A Bonferroni multiple testing correction was used to control for the number of tests conducted in each enrichment analysis. GTEx data were also used to verify whether the GWS loci identified affect the transcriptomic regulation of the surrounding genes. To evaluate the effect across multiple tissues, we considered multi-tissue expression quantitative trait locus (eQTL) data. These were calculated using Meta-Tissue [32]. This meta-analytic approach calculates a posterior probability (m value) that an effect exists in each of the tissues tested assuming that the eQTL effect is consistent across the affected tissues. M values>0.9 indicate that the tissue was predicted to show the eQTL association.

A PRS analysis was conducted to test the genetic overlap with behavioral traits that could differentiate between OD and OE status using the PRSice software[33]. Risk-taking and neuroticism were selected as we expected they would capture genetic susceptibility to early versus later stages of opioid use and misuse. For polygenic profile scoring, we used summary statistics generated from large-scale GWAS of risk tolerance (N= 466,571)[34] and neuroticism (N=390,278)[35]. We considered multiple association P value thresholds (PT < 5×10^{-8}, 10^{-7}, 10^{-6}, 10^{-5}, 10^{-4}, 0.001, 0.05, 0.1, 0.3, 0.5, 1) for SNP inclusion to identify the best-fit for each target phenotype tested. The PRS were calculated after using P-value-informed clumping with a LD cut-off of $R^2 = 0.3$ within a 500 kb window and excluding the major histocompatibility complex region of the genome because of its complex LD structure. The PRS were calculated considering unrelated subjects of European descent available in both case-control and family-based cohorts ($OD_{exposed}$ $N_{effective}$=3,038; $OD_{unexposed}$ $N_{effective}$=4,728; and $OE_{controls}$ $N_{effective}$=5,376). The PRS were fitted in regression models with adjustments for sex and the top 10 within-ancestry principal components. We applied FDR multiple testing correction (q<0.05) to correct for the number of thresholds tested.

RESULTS

SNP-Heritability Estimates Comparing OD and OE Traits

The GWAS meta-analyses of $OD_{exposed}$, $OD_{unexposed}$, and $OE_{controls}$ phenotypes included up to 4,503 OD cases (African-ancestry=1,231; European-ancestry=3,272), 4,173 OE controls (African-ancestry=1,297; European-ancestry=2,876), and 32,500 OU controls (African-ancestry=7,063; European-ancestry=25,437). Significant SNP-heritability was observed for $OD_{unexposed}$ (liability-scale h^2_g=0.28, SE=0.1; population prevalence=0.01, sample prevalence=0.22), but not for $OD_{exposed}$ (liability-scale h^2_g=−0.08, SE=0.08; population prevalence=0.01, sample prevalence=0.55) and $OE_{controls}$ (liability-scale h^2_g=0.05, SE=0.1; population prevalence=0.05, sample prevalence=0.12). Moderate genome-wide inflation was observed in the European-specific meta-analyses, but genomic control using lambda or the LD score regression intercept did not affect the significance of any variants in the GWAS meta-analyses (Supplementary Table 4; Supplementary Methods).

Opioid Dependence vs. Exposed and Unexposed Controls

In the $OD_{exposed}$ analysis, which is the comparison most relevant to dependence liability given exposure but also the one that most constricted the sample size, no association survived the genome-wide significance threshold ($p=5\times10^{-8}$). Additionally, there were no significant enrichments for GO annotations, molecular pathways, nor tissue-specific regulation. The $OD_{unexposed}$ comparison identified a GWS association in the African-ancestry meta-analysis, rs201123820 on chromosome 18 (z=5.55, $p=2.9\times10^{-8}$; Figure 1A; Table 1; see Supplementary Table 5 for ancestry-specific results for each genome-wide significant variant). With respect to this locus, no heterogeneity was observed among the cohorts included in the meta-analysis (heterogeneity: I^2=0, p=0.473; Supplementary Table 6). This variant did not show significant genetic associations with traits related to other addictive substances (Supplementary Table 3). The gene-based association analysis identified a GWS gene in the same genomic region, *C18orf32* ($p=1.8\times10^{-6}$; Table 1; Supplementary Figure 1A). Additionally, in the African-ancestry meta-analysis, we also observed an enrichment for adipose tissue (beta=0.04, $p=4.21\times10^{-4}$; Figure 2A) and GO:0034498 – *early endosome to Golgi transport* (beta=1.01, $p=5.1\times10^{-8}$). In the trans-ancestry meta-analysis, we observed significant enrichment for specific adult stages of brain development (37 y; beta=0.06, $p=6.22\times10^{-4}$; 15 y; beta=0.06, p=0.001; 36yrs: beta=0.06, p=0.002; Figure 2B) and GO:0007143~*female meiotic division* (beta=0.73, $p=1.08\times10^{-7}$). In the European-ancestry $OD_{unexposed}$ GWAS meta-analysis, no result survived multiple testing correction.

To extend the phenotypic breadth of our findings, we conducted a phenome-wide scan (4,082 traits tested; Supplementary Table 7) identified in 361,194 participants of European descent from the UK Biobank. Rs201123820 was identified in the African-ancestry $OD_{unexposed}$ GWAS meta-analysis. Although this variant was not significant in the European-ancestry meta-analysis (Supplementary Table 5), Rs201123820 has reasonably similar MAF in African and European populations (1000 Genomes Project: AFR MAF=0.024; EUR MAF=0.042). In UK Biobank cohort, we observed several associations for rs201123820 (FDR q<0.05; Figure 3, center panel; Supplementary Table 7) including

postprocedural musculoskeletal disorders (UK Biobank Field ID: 41202 "*Diagnoses – main ICD10 [M96]*", p=4.88×10⁻⁷), other disorders of the musculoskeletal system and connective tissue (UK Biobank Field ID: 41270 "*Diagnoses - ICD10 [M13_MUSCULOSKELEOTH]*", p=1.25×10⁻⁵), postpartum care and examination (UK Biobank Field ID: 41202 "*Diagnoses - main ICD10 [Z39]*", p=3.74×10⁻⁵), and auto-refraction measurements for eye prescription (UK Biobank Field ID: 5159 "*3mm asymmetry index (right)*", p=6.54×10⁻⁵).

Exposed vs. Unexposed Controls

In GWAS meta-analysis of $OE_{controls}$ in the European-ancestry cohort, we observed a gene-based association for the *BEND4* locus that was GWS in the gene-based test (p=9.9×10⁻⁶; Table 1; Supplementary Figure 1B). In the *BEND4* gene region, we identified a genetic association that nearly reached GWS: rs9291211 on chromosome 4 (z=−5.38, p=7.2×10⁻⁸; Figure 1B; Table 1). With respect to this locus, no heterogeneity was observed among the cohorts (heterogeneity: I^2=0, p=0.879; Supplementary Table 6). This variant (or LD proxies in the same ancestry group) was identified in previous GWAS of behavioral traits: alcohol consumption (rs4501255, LD proxy r^2=0.94, p=5×10⁻¹⁰)[36]; neuroticism (rs9291211, p=2×10⁻⁸)[35]; and helping behavior (rs2880666, LD-proxy r^2=0.77, p=5×10⁻⁷)[37]. Additionally, rs9291211 is an eQTL for *SLC30A9* and *BEND4* in multiple tissues (GTEx multi-tissue eQTL p = 1.2×10⁻²⁶ and 2.88×10⁻⁹, respectively). The rs9291211×*SLC30A9* eQTL (i.e., rs9291211 regulating the *SLC30A9* expression) showed a posterior probability>90% in seven brain tissues (amygdala m=0.99; anterior cingulate cortex m=1; caudate m=1; cortex m=0.99; hypothalamus m=1; nucleus accumbens m=1, putamen m=0.99; Supplementary Figure 2A). The rs9291211×*BEND4* eQTL showed posterior probabilities>90% in two brain tissues (caudate m=0.9; cortex m=1; Supplementary Figure 2B). We found significant association of rs9291211 with additional traits from the GWAS Atlas[24] that were related to addictive substances including "*alcohol usually taken with meals*" (p=3.02×10⁻⁸), "*frequency of consuming six or more units of alcohol*" (p=2.78×10⁻⁴), "*alcohol intake frequency*" (p=5.91×10⁻⁴), and "*cigarettes per day*" (p=0.004) (Supplementary Table 3). The European-ancestry $OE_{controls}$ meta-analysis also showed enrichment for several brain development stages: post-conception weeks 9 (beta=0.04, p=1.28×10⁻⁴), 8 (beta=0.032, p=0.001), and 12 (beta=0.04, p=0.002) (Figure 2C). No result in the association and the enrichment analyses based on the African-ancestry $OE_{controls}$ GWAS meta-analysis survived multiple testing correction.

Although rs9291211 showed only a "suggestive" GWS association with the $OE_{controls}$ phenotype, it was the strongest signal responsible for the significant *BEND4* gene-based association. Accordingly, we tested its phenotypic spectrum in UK Biobank. This variant was associated with 22 phenotypes (FDR q<0.05; Figure 3, upper panel; Supplementary Table 7), which included dietary habits (e.g., UK Biobank Field ID: 6179 "*Mineral and other dietary supplements [None of the above]*", p=1.68×10⁻⁸), anthropometric traits (e.g., UK Biobank Field ID: 1687 "*Comparative body size at age 10*", p=7.58×10⁻⁸), behavioral traits (e.g., UK Biobank Field ID: 20127 "*Neuroticism score*", p=3.12×10⁻⁶), physical outcomes (e.g., UK Biobank Field ID: 6152 "*Hay fever, allergic rhinitis or eczema*", p=3.49×10⁻⁵), reproductive function (UK Biobank Field ID: 3581 *Age at menopause [last menstrual*

period], p=8.40×10⁻⁵), and cognitive tests (e.g., UK Biobank Field ID: 404 "*Reaction time [Duration to first press of snap-button in each round]*", p= 8.82×10⁻⁵).

Based on the gene-based and eQTL analyses, we also investigated the phenotypic spectrum of *SLC30A9* and *BEND4*. In line with the shared effect of rs9291211, we observed associations surviving multiple-testing correction (p<1.05×10⁻⁵) in both gene-based phenome-wide scans (Supplementary Table 8). Among the 14 common associations, we observed: *"alcohol usually taken with meals"* (*BEND4* p=6.5×10⁻¹¹; *SLC30A9* p=3.59×10⁻⁶), "*neuroticism sum score*" (*BEND4* p=1.92×10⁻⁸; *SLC30A9* p=3.63×10⁻⁷), and "*depressive symptoms*" (*BEND4* p=6.24×10⁻⁷; *SLC30A9* p=3.58×10⁻⁶).

In the trans-ancestry GWAS meta-analysis of $OE_{controls}$, we observed an additional single-variant GWS association, rs12461856 on chromosome 19 (z=−5.61, p=2.1×10⁻⁸; Figure 1C; Table 1). With respect to this locus, no heterogeneity was observed among the cohorts included in the meta-analysis (heterogeneity: I^2=0, p=0.554; Supplementary Table 6). This variant did not show significant genetic associations with traits related to other addictive substances (Supplementary Table 3). In the UK Biobank, no novel phenotypic associations with rs12461856 survived multiple testing correction (FDR q<0.05; Figure 3, bottom panel; Supplementary Table 7). A gene-based GWS association was identified for *SDCCAG8* on chromosome 1 (p=1.4×10⁻⁶, Table 1; Supplementary Figure 1C). In the gene-based phenome-wide scan, we observed 77 traits associated with *SDCCAG8* that survived multiple testing correction (p<1.05×10⁻⁵; Supplementary Table 8). Among them, we observed strong associations for schizophrenia (p=2.5×10⁻¹²) and risk taking (p=1.1×10⁻⁹). With respect to the molecular pathways, significant enrichments were also observed for GO:0017069~*small RNA binding* (beta=0.65, p=5.4×10⁻⁷) and a curated gene set related to genes downregulated 6 h after induction of *HoxA5* expression in a breast cancer cell line (Standard name: CHEN_HOXA5_TARGETS_6HR_DN; beta=1.37, p=8.6×10⁻⁶).

Polygenic Risk Score Analysis

We also used PRS to compare the three opioid-related phenotypes – dependence (with exposed controls – $OD_{exposed}$ – and with unexposed controls – $OD_{unexposed}$) and exposure in non-dependent individuals ($OE_{controls}$). Similarly to other addictive substances, illicit opioid users would be expected to have greater propensity to risk-taking behaviors, impulsivity, and stress responsivity than unexposed subjects[38]. Accordingly, we derived a PRS from the large-scale GWAS (N=466,571) conducted by the Social Science Genetic Association Consortium (SSGAC) on risk tolerance, which was defined as the tendency, preparedness, or willingness to take risks in general[34]. The PRS analysis was conducted on European-ancestry subjects only due to the well-known lack of large-scale GWAS in other ancestry groups[39]. The risk-tolerance PRS was positively associated with OD when contrasted with unexposed controls ($OD_{unexposed}$: $N_{effective}$=4,728, PT=1, z=3.94, p=8.1×10⁻⁵, FDR q=0.003), whereas OD contrasted with exposed controls displayed only a trend (p<0.1; $OD_{exposed}$: $N_{effective}$=3,038, PT=1, z=1.93, p=0.054, FDR q=0.13). OE ($OE_{control}$: $N_{effective}$=5,376, PT=0.05, z=3.57, p=3.6×10⁻⁵, FDR q=0.003) was also significant for the risk-tolerance PRS.

We also tested PRS derived from a large-scale GWAS of neuroticism (N=390,278)[35]. This behavioral trait represents a tendency to negative affect and was previously observed to be genetically correlated with several psychiatric disorders, including SUDs[17, 40], major depression[41], and posttraumatic stress disorder[42]. Consistent with our expectation of genetic liability to negative affect being related to dependence but not exposure alone, the neuroticism PRS was associated with dependence compared with unexposed (OD$_{unexposed}$: N$_{effective}$=4,728, PT=0.001, z=4.16, p=3.2×10^{-5}, FDR q=5.76×10^{-4}) and dependence compared with exposed controls (OD$_{exposed}$: N$_{effective}$=3,038, PT=0.001, z=3.1, p=0.002, FDR q=0.016), but not with exposed vs unexposed controls (OE$_{controls}$: N$_{effective}$=5,376, PT=0.5, z=−0.42, p=0.671, FDR q=0.919).

Figure 4 shows the association of risk-taking and neuroticism PRS (Panels A and B, respectively) across the cohorts included in the meta-analyses. We did not observe significant heterogeneity (Supplementary Table 9), which indicates that the meta-analytic results were driven by the sizes of the samples investigated and not by the different recruitment strategies.

DISCUSSION

We investigated the genetic architecture of opioid-related traits in informative cohorts. Our comparison of opioid dependence (OD) compared with unexposed (OD$_{unexposed}$) and exposed (OD$_{exposed}$) controls, as well as of opioid exposure among controls (OE$_{controls}$) provides new insights into opioid addiction. We identified GWS loci and genes for OD$_{unexposed}$ and OE$_{controls}$, and found these variants to be associated with health relevant traits in the UK Biobank. Most critically, our PRS analyses highlighted distinctions between exposed and unexposed controls, as well as the progression from exposure to dependence. Use and dependence are behaviors with different relationships to other genetically influenced traits, as has been shown for alcohol use vs. alcohol dependence[13, 14]. To our knowledge, no previous study investigated the specific genetic differences between OD and OE; our current findings provide the first insights based on genome-wide data into the molecular mechanisms by which OE and OD differ. The lack of sufficiently large numbers of OD cases and OE controls is a fundamental barrier to facilitating our understanding of the biological underpinnings of this serious public health epidemic, as it limited the power of what we would regard as the most informative comparison.

With respect to the single-variant associations observed, the strongest bioinformatics support from other studies was observed for rs9291211, identified in the European-ancestry GWAS meta-analysis of the OE$_{controls}$ phenotype. Although it reached only a suggestive GWS threshold (p=7.2×10^{-8}), this variant was the leading signal in the significant *BEND4* gene-based association and also showed strong regulatory effects on the brain-specific transcriptomic profile of *BEND4* and *SLC30A9*. The function of *BEND4* gene is unclear, but previous GWAS identified several variants at this locus (including rs9291211) that were associated with depression[43], alcohol consumption[36], autism spectrum disorder[43], neuroticism[35], height[44], and helping behavior[37]. *SLC30A9* encodes a zinc transporter involved in intracellular zinc homeostasis, which also plays a role in transcriptional activation of Wnt-responsive genes[45]. In previous GWAS, variants located in *SLC30A9* gene

(rs9291211 is located in *BEND4* but also affects *SLC30A9* gene expression) were associated with neuroticism[44] and depression[46]. The phenome-wide scan of rs9291211 in the UK Biobank showed an effect of this SNP on a wide range of complex traits, some with an easy-to-conceptualize relationship to OD such as alcohol consumption, neuroticism, depression, and anxious feelings. Considering traits related to mental and behavioral disorders attributable to use of alcohol, cannabis, and tobacco, we observed that rs9291211 is associated with alcohol consumption and tobacco smoking and not with dependence-relevant phenotypes. This support that this variant may have pleiotropic effects across consumption of multiple substancs. In our phenome-wide analysis, the strongest results were observed with respect to dietary habits: rs9291211*A was positively associated with reduced OE risk in the PGC-SUD cohorts and with increased propensity to use dietary supplements, such as vitamin and mineral supplements in the UK Biobank. A recent GWAS identified several loci associated with dietary habits and indicated a causal relationship between educational attainment and healthy eating[47]. With respect to rs9291211, we also observed a nominally significant association with traits related to educational attainment (e.g., UK Biobank Field ID: 6138 *Qualification [College or University degree]*, p=0.033). Accordingly, we hypothesize that rs9291211 could be involved in the individual variability to consume chemicals ranging from dietary supplements to opioids, independent from educational attainment.

In the trans-ancestry GWAS meta-analysis of the $OE_{controls}$ phenotype comparison, we identified GWS loci in the single-variant and the gene-based analyses. No external validations were observed for rs12461856 and further studies will be needed to confirm this finding. Conversely, *SDCCAG8* identified in the gene-based analysis (but not related to any individual GWS variants) was shown in studies available in the GWAS catalog[48] to have 49 single-variant associations with educational attainment[49], blood-related parameters[50], risk-taking behaviors[34], anthropometric traits[44], kidney function[51], and schizophrenia[52]. The previous associations with behavioral traits support *SDCCAG8* as potentially associated with behaviors that, in turn, associated with increased risk of OE.

With respect to the $OD_{unexposed}$ phenotype comparison, we identified rs201123820 in the African-ancestry meta-analysis. This is a non-coding deletion located 2 kb upstream of *LOC101928144*, an uncharacterized long intergenic non-protein coding RNA. The gene-based association analysis identified a GWS locus in the same region, *C18orf32*, a gene involved in the activation of the NF-kappaB and MAPK signaling pathways, which play a key role in immune and inflammatory responses[53]. The phenome-wide analysis in the UK Biobank, despite its being in a predominantly European cohort, showed a significant association of rs201123820 with physical conditions, particularly musculoskeletal disorders. This is particularly interesting as opioids are commonly prescribed for pain management in musculoskeletal disorders and early use is associated with prolonged work disability[54], which may be related to the consequences of opioid abuse and/or the severity of the underlying disorder that required treatment. While these associations merit replication, this result highlights how human genetic research can also be relevant to improve pain management protocols.

The GWS risk loci identified by previous OD GWAS (e.g., *KCNG2*, *CNIH3*, and *RGMA*)[7–10] were not concordant with the present investigation (Supplementary Table 10). Such discrepancies are not unexpected, given that these analyses were underpowered, and the reported findings are likely to be affected by phenotypic heterogeneity and the random variation allowing for discovery of alternate subsets of risk loci in small datasets[55, 56].

The available genome-wide data also permitted us to compare the opioid-related phenotypes with respect to shared genetic risk of relevant behavioral traits. Although there is variability in the effective sample size and therefore statistical power of the phenotypes tested ($OD_{unexposed}$ $N_{effective}$=4,728; $OE_{controls}$: $N_{effective}$=5,376; $OD_{exposed}$ $N_{effective}$=3,038), the PRS results showed an interesting pattern. The risk-tolerance PRS was positively associated with all three phenotypes with the strength of association mostly related to the effective sample size of the target sample. The association between $OE_{controls}$ and genetic liability to risk-taking highlights the importance of accounting for the genetic factors related to the individual differences in exposure when examining those contribution to dependence. Such a finding would support the hypothesis that the inclusion of exposed controls can "fine-tune" our ability to separate loci related to generalized risk-taking from those specific to repeated use that lead to opioid dependence. The neuroticism PRS showed positive associations with $OD_{unexposed}$ and $OD_{exposed}$ phenotype comparisons. Although it was non-significant, we observed a negative association of neuroticism PRS with the $OE_{controls}$ phenotype. This suggestive negative relationship parallels the rs9291211 result where the allele A was associated with reduced OE in the PGC GWAS meta-analysis and increased neuroticism score in the UK Biobank. Genetic liability to neuroticism may thus overlap with genetic liability to OD but not to OE.

Although several putative single-variant, gene-based, and PRS associations were identified based on the different OD and OE phenotypes, the sample size of the current investigation is still small, given the polygenic architecture of psychiatric disorders[57]. Novel studies specifically targeting SUDs and assessing opioid-related behaviors will be necessary to recruit cohorts informative for OD and OE GWAS. Another important limitation is the phenotypic heterogeneity within the opioid exposure sample, which included individual exposed to opioids via licit use (i.e., medical prescriptions) and illicit use. There may be important differences between these two subgroups (e.g., risk-taking may be more strongly associated with illicit exposure). However, several of the cohorts investigated lacked this information, and, due to the limited sample size, we were not able to make this comparison. In addition, this may have resulted in heterogeneity in the OU controls (i.e., those who reported not using opioids illicitly were unassessed for medical exposure). Future large opioid-informative datasets will be needed to determine whether illicit and licit opioid exposure have distinct effects on the molecular basis of opioid dependence. Finally, while the phenome-wide investigation in the UK Biobank provides encouraging support for the plausibility of our findings, it may reflect complex pleiotropic effects of these variants on multiple traits, including an unmeasured third variable. Replication of these association signals with opioid dependence and exposure phenotypes will be required.

In conclusion, we provide a comprehensive genome-wide investigation of opioid-related traits, highlighting different molecular mechanisms that could underlie exposure and

dependence. These findings draw attention to challenges associated with the use of unexposed controls in genetic association studies for OD and potentially for other SUDs (where exposure is not widespread, as is the case for alcohol, or more recently marijuana). This information should be used to guide the next generation of human genetic studies of opioid-related behaviors.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

REFERENCES

1. Cicero TJ, Ellis MS. Nonmedical Prescription-Opioid Use and Heroin Use. N Engl J Med 2016; 374(13): 1295–1296. [PubMed: 27028931]

2. Volkow ND, Frieden TR, Hyde PS, Cha SS. Medication-assisted therapies--tackling the opioid-overdose epidemic. N Engl J Med 2014; 370(22): 2063–2066. [PubMed: 24758595]

3. Centers for Disease Control Prevention. CDC grand rounds: prescription drug overdoses-a U.S. epidemic. MMWR Morb Mortal Wkly Rep 2012; 61(1): 10–13. [PubMed: 22237030]

4. Rudd RA, Aleshire N, Zibbell JE, Gladden RM. Increases in Drug and Opioid Overdose Deaths--United States, 2000–2014. MMWR Morb Mortal Wkly Rep 2016; 64(50–51): 1378–1382. [PubMed: 26720857]

5. Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA et al. 10 Years of GWAS Discovery: Biology, Function, and Translation. Am J Hum Genet 2017; 101(1): 5–22. [PubMed: 28686856]

6. Sullivan PF, Agrawal A, Bulik CM, Andreassen OA, Borglum AD, Breen G et al. Psychiatric Genomics: An Update and an Agenda. Am J Psychiatry 2018; 175(1): 15–27. [PubMed: 28969442]

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. [PubMed: 26239289]

8. 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. [PubMed: 24143882]

9. 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. [PubMed: 29478698]

10. 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. [PubMed: 27936112]

11. Kelly TM, Daley DC, Douaihy AB. Treatment of substance abusing patients with comorbid psychiatric disorders. Addict Behav 2012; 37(1): 11–24. [PubMed: 21981788]

12. Clark RE, Baxter JD, Aweh G, O'Connell E, Fisher WH, Barton BA. Risk Factors for Relapse and Higher Costs Among Medicaid Members with Opioid Dependence or Abuse: Opioid Agonists, Comorbidities, and Treatment History. J Subst Abuse Treat 2015; 57: 75–80. [PubMed: 25997674]

13. Sanchez-Roige S, Palmer AA, Fontanillas P, Elson SL, andMe Research Team tSUDWGotPGC, Adams MJ et al. Genome-Wide Association Study Meta-Analysis of the Alcohol Use Disorders Identification Test (AUDIT) in Two Population-Based Cohorts. Am J Psychiatry 2019; 176(2): 107–118. [PubMed: 30336701]

14. Kranzler HR, Zhou H, Kember RL, Vickers Smith R, Justice AC, Damrauer S et al. Genome-wide association study of alcohol consumption and use disorder in 274,424 individuals from multiple populations. Nat Commun 2019; 10(1): 1499. [PubMed: 30940813]

15. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision. APA Publishing: Washington DC, USA, 2000.

16. Lam M, Awasthi S, Watson HJ, Goldstein J, Panagiotaropoulou G, Trubetskoy V et al. RICOPILI: Rapid Imputation for COnsortias PIpeLIne. Bioinformatics 2019.

17. Walters RK, Polimanti R, Johnson EC, McClintick JN, Adams MJ, Adkins AE et al. Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. Nat Neurosci 2018; 21(12): 1656–1669. [PubMed: 30482948]

18. 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. [PubMed: 31090166]

19. Consortium. GP, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM et al. A global reference for human genetic variation. Nature 2015; 526(7571): 68–74. [PubMed: 26432245]

20. O'Connell J, Gurdasani D, Delaneau O, Pirastu N, Ulivi S, Cocca M et al. A general approach for haplotype phasing across the full spectrum of relatedness. PLoS Genet 2014; 10(4): e1004234. [PubMed: 24743097]

21. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet 2009; 5(6): e1000529. [PubMed: 19543373]

22. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010; 26(17): 2190–2191. [PubMed: 20616382]

23. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 2015; 12(3): e1001779. [PubMed: 25826379]

24. 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. [PubMed: 31427789]

25. 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. [PubMed: 25642630]

26. International HapMap Consortium, Altshuler DM, Gibbs RA, Peltonen L, Altshuler DM, Gibbs RA et al. Integrating common and rare genetic variation in diverse human populations. Nature 2010; 467(7311): 52–58. [PubMed: 20811451]

27. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput Biol 2015; 11(4): e1004219. [PubMed: 25885710]

28. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun 2017; 8(1): 1826. [PubMed: 29184056]

29. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdottir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. Bioinformatics 2011; 27(12): 1739–1740. [PubMed: 21546393]

30. Consortium GTEx. Genetic effects on gene expression across human tissues. Nature 2017; 550(7675): 204–213. [PubMed: 29022597]

31. Ramasamy A, Trabzuni D, Guelfi S, Varghese V, Smith C, Walker R et al. Genetic variability in the regulation of gene expression in ten regions of the human brain. Nat Neurosci 2014; 17(10): 1418–1428. [PubMed: 25174004]

32. Sul JH, Han B, Ye C, Choi T, Eskin E. Effectively identifying eQTLs from multiple tissues by combining mixed model and meta-analytic approaches. PLoS Genet 2013; 9(6): e1003491. [PubMed: 23785294]

33. Euesden J, Lewis CM, O'Reilly PF. PRSice: Polygenic Risk Score software. Bioinformatics 2015; 31(9): 1466–1468. [PubMed: 25550326]

34. Karlsson Linner R, Biroli P, Kong E, Meddens SFW, Wedow R, Fontana MA et al. Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. Nat Genet 2019; 51(2): 245–257. [PubMed: 30643258]

35. Nagel M, Jansen PR, Stringer S, Watanabe K, de Leeuw CA, Bryois J et al. Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. Nat Genet 2018; 50(7): 920–927. [PubMed: 29942085]

36. Liu M, Jiang Y, Wedow R, Li Y, Brazel DM, Chen F et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. Nat Genet 2019; 51(2): 237–244. [PubMed: 30643251]

37. Primes G, Fieder M. Real-life helping behaviours in North America: A genome-wide association approach. PLoS One 2018; 13(1): e0190950. [PubMed: 29324852]

38. Kreek MJ, Nielsen DA, Butelman ER, LaForge KS. Genetic influences on impulsivity, risk taking, stress responsivity and vulnerability to drug abuse and addiction. Nat Neurosci 2005; 8(11): 1450–1457. [PubMed: 16251987]

39. Sirugo G, Williams SM, Tishkoff SA. The Missing Diversity in Human Genetic Studies. Cell 2019; 177(1): 26–31. [PubMed: 30901543]

40. Pasman JA, Verweij KJH, Gerring Z, Stringer S, Sanchez-Roige S, Treur JL et al. GWAS of lifetime cannabis use reveals new risk loci, genetic overlap with psychiatric traits, and a causal influence of schizophrenia. Nat Neurosci 2018; 21(9): 1161–1170. [PubMed: 30150663]

41. Howard DM, Adams MJ, Clarke TK, Hafferty JD, Gibson J, Shirali M et al. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. Nat Neurosci 2019; 22(3): 343–352. [PubMed: 30718901]

42. Gelernter J, Sun N, Polimanti R, Pietrzak RH, Levey D, Bryois J et al. Genome-wide association study of post-traumatic stress disorder reexperiencing symptoms in >165,000 US veterans. Nature Neuro 2019.

43. Grove J, Ripke S, Als TD, Mattheisen M, Walters RK, Won H et al. Identification of common genetic risk variants for autism spectrum disorder. Nat Genet 2019; 51(3): 431–444. [PubMed: 30804558]

44. Kichaev G, Bhatia G, Loh PR, Gazal S, Burch K, Freund MK et al. Leveraging Polygenic Functional Enrichment to Improve GWAS Power. Am J Hum Genet 2019; 104(1): 65–75. [PubMed: 30595370]

45. Perez Y, Shorer Z, Liani-Leibson K, Chabosseau P, Kadir R, Volodarsky M et al. SLC30A9 mutation affecting intracellular zinc homeostasis causes a novel cerebro-renal syndrome. Brain 2017; 140(4): 928–939. [PubMed: 28334855]

46. Wray NR, Ripke S, Mattheisen M, Trzaskowski M, Byrne EM, Abdellaoui A et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. Nat Genet 2018; 50(5): 668–681. [PubMed: 29700475]

47. Cole JB, Florez JC, Hirschhorn JN. Comprehensive genomic analysis of dietary habits in UK Biobank identifies hundreds of genetic loci and establishes causal relationships between educational attainment and healthy eating. bioRxiv 2019: 662239.

48. MacArthur J, Bowler E, Cerezo M, Gil L, Hall P, Hastings E et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleic Acids Res 2017; 45(D1): D896–D901. [PubMed: 27899670]

49. Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nat Genet 2018; 50(8): 1112–1121. [PubMed: 30038396]

50. Hoffmann TJ, Ehret GB, Nandakumar P, Ranatunga D, Schaefer C, Kwok PY et al. Genome-wide association analyses using electronic health records identify new loci influencing blood pressure variation. Nat Genet 2017; 49(1): 54–64. [PubMed: 27841878]

51. Yasukochi Y, Sakuma J, Takeuchi I, Kato K, Oguri M, Fujimaki T et al. Identification of CDC42BPG as a novel susceptibility locus for hyperuricemia in a Japanese population. Mol Genet Genomics 2018; 293(2): 371–379. [PubMed: 29124443]

52. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. Nature 2014; 511(7510): 421–427. [PubMed: 25056061]

53. Matsuda A, Suzuki Y, Honda G, Muramatsu S, Matsuzaki O, Nagano Y et al. Large-scale identification and characterization of human genes that activate NF-kappaB and MAPK signaling pathways. Oncogene 2003; 22(21): 3307–3318. [PubMed: 12761501]

54. Busse JW, Craigie S, Sadeghirad B, Couban R, Hong P, Oparin Y et al. Management of acute musculoskeletal pain (excluding low back pain): protocol for a systematic review and network meta-analysis of randomised trials. BMJ Open 2019; 9(4): e024441.

55. Jiang W, Yu W. Power estimation and sample size determination for replication studies of genome-wide association studies. BMC Genomics 2016; 17 Suppl 1: 3. [PubMed: 26818952]

56. Zou J, Zhou J, Faller S, Brown R, Eskin E. Accurate modeling of replication rates in genome-wide association studies by accounting for winner's curse and study-specific heterogeneity. bioRxiv 2019: 856898.

57. Sullivan PF, Geschwind DH. Defining the Genetic, Genomic, Cellular, and Diagnostic Architectures of Psychiatric Disorders. Cell 2019; 177(1): 162–183. [PubMed: 30901538]
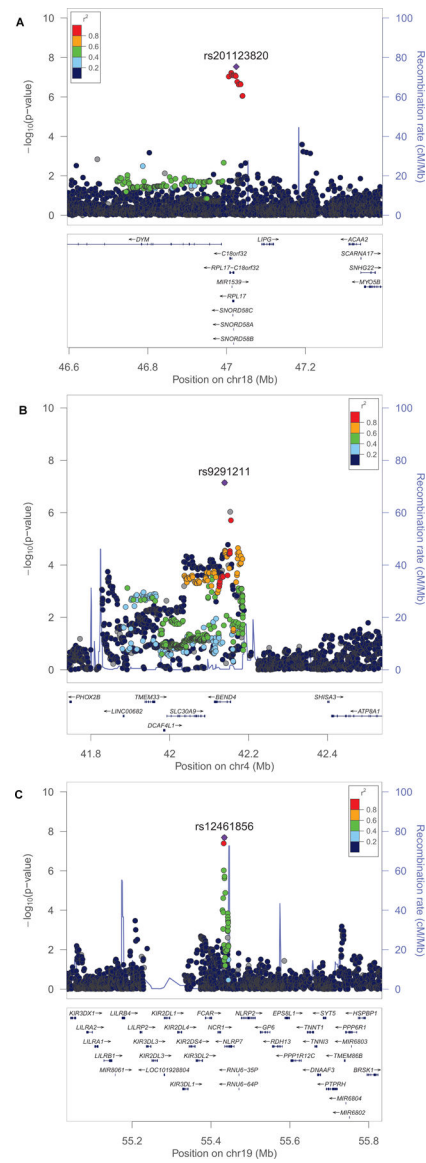
Figure 1:
Regional Manhattan plots of the genetic association identified: rs201123820 in the African-ancestry OD$_{unexposed}$ GWAS meta-analysis (**A**); rs92911211 in the European-ancestry OE$_{controls}$ GWAS meta-analysis (**B**); rs12461856 in the trans-ancestry OE$_{controls}$ GWAS meta-analysis (**C**).
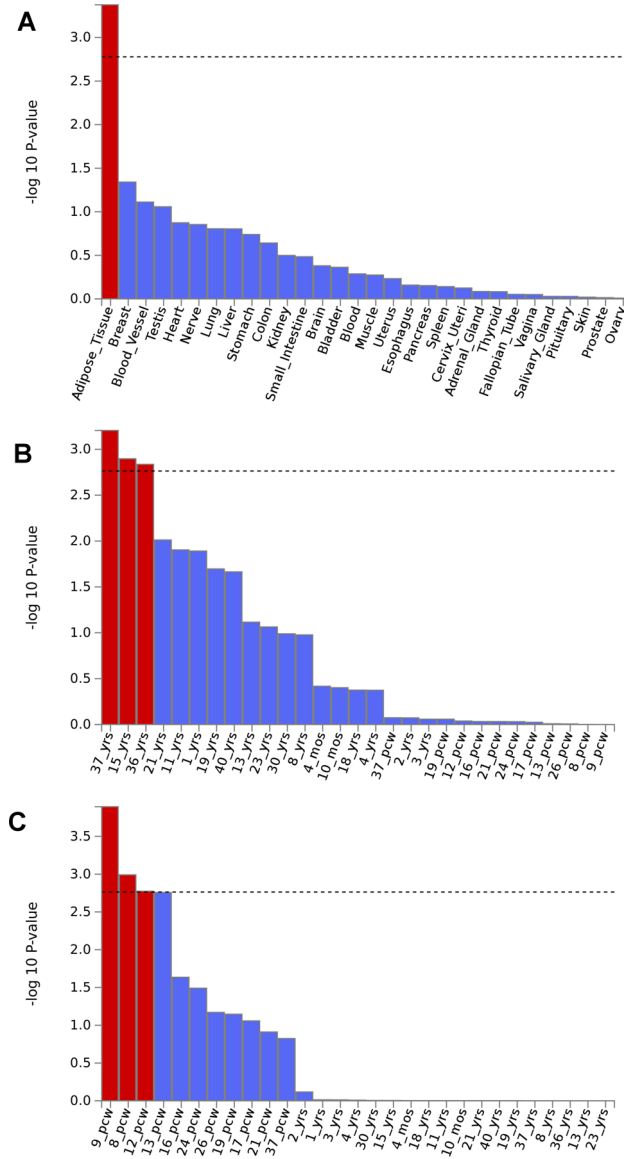
Figure 2:
Significant tissue enrichments identified in the African-ancestry OD$_{unexposed}$ meta-analysis
(**A**); the trans-ancestry OD$_{unexposed}$ GWAS meta-analysis (**B**), and the European-ancestry
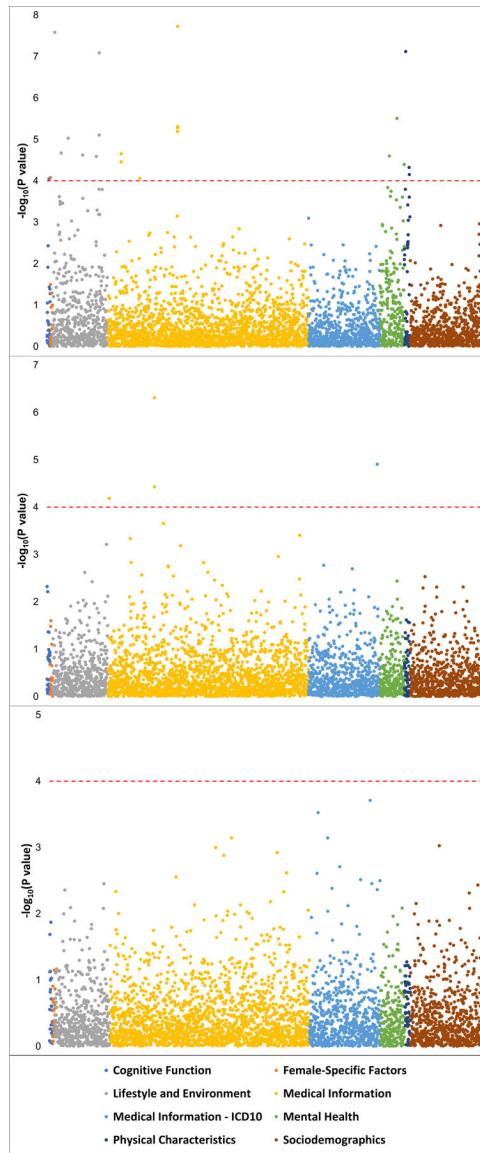OE$_{controls}$ GWAS meta-analysis (**C**).

Figure 3:
Manhattan plot of the phenome-wide scan conducted in the UK Biobank with respect to rs12461856, rs201123820, and rs9291211 (bottom, center, and upper panels, respectively). As indicated in the legend, the phenotypic categories are color coded. Red dashed line indicates the significance threshold accounting for the number of variants and phenotypes tested (FDR q < 0.05).
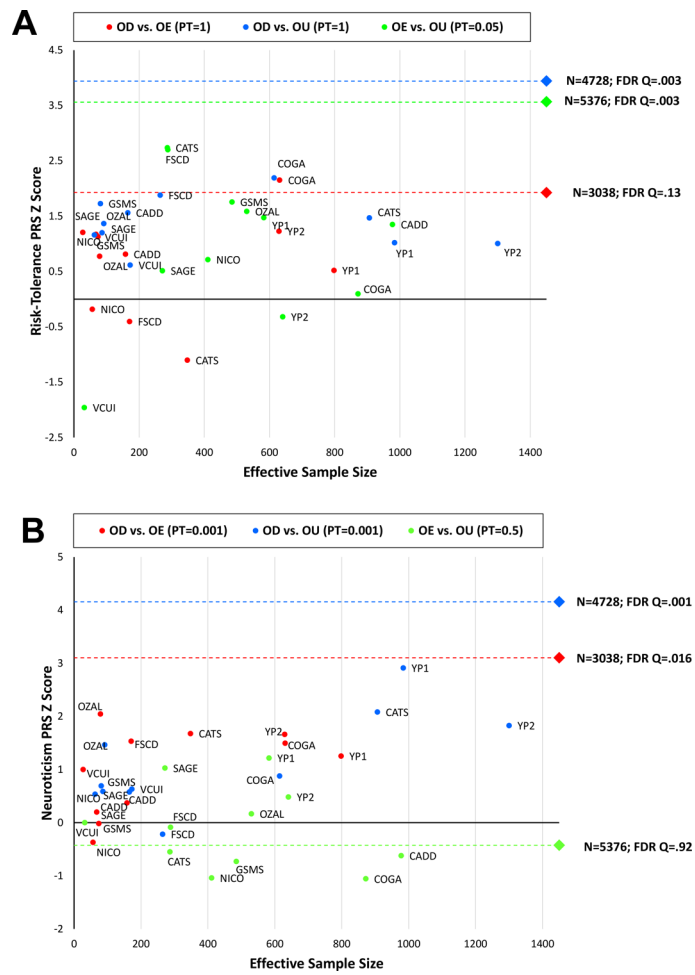
Figure 4:
Relationship between PRS z scores (**A**: risk tolerance; **B**: neuroticism) and effective sample size across the opioid-related phenotypes tested. Each circle represents an individual cohort; the diamond represents the results from the meta-analysis with respect to the phenotypes tested.

Table 1:

Top loci identified considering different OD and OE phenotypic definitions.

				Single-variant Associations					
Phenotype	Meta-analysis	rsid	Chromosome	Location (bp)	Effect Allele	Other Allele	Effect Allele Frequency	Z score	P value
OD$_{unexposed}$	AFR	rs201123820	18	47025347	T	TAAAACAAAAACA	0.019	5.547	2.90E-08
OE$_{controls}$	EUR	rs9291211	4	42139132	A	G	0.782	−5.387	7.16E-08
	Trans-ancestry	rs12461856	19	55433852	A	G	0.8402	−5.606	2.07E-08

				Gene-based Associations				
Phenotype	Meta-analysis	Gene	Chromosome	Location-Start (bp)	Location-End (bp)	SNP (n)	Z score	P value
OD$_{unexposed}$	AFR	C18orf32	18	47008028	47013622	18	4.633	2E-06
OE$_{controls}$	EUR	BEND4	4	42112955	42154895	81	4.756	1E-06
	Trans-ancestry	SDCCAG8	1	243419320	243663394	407	4.686	1E-06