Genome-wide association study of problematic opioid prescription use in 132,113 23andMe research participants of European ancestry

Sandra Sanchez-Roige[1-2*], Pierre Fontanillas[3], Mariela V Jennings[1], Sevim Bianchi[1], Yuye Huang[1], Alex Hatoum[4], Julia Sealock[2], Sarah L Elson[3], 23andMe Research Team[3], Abraham A Palmer[1,7*]

[1]Department of Psychiatry, University of California San Diego; [2]Department of Medicine, Division of Genetic Medicine, Vanderbilt University; [3]23andMe, Inc., Sunnyvale, CA, USA. A full list of members and affiliations appears at the end of the paper; [4] Department of Psychiatry Washington University School of Medicine, St. Louis, MO, USA; [7]Institute for Genomic Medicine, University of California San Diego

*** Co-corresponding Authors**: Sandra Sanchez-Roige, sanchezroige@ucsd.edu, Abraham A. Palmer, aap@ucsd.edu

ABSTRACT

Rates of opioid use disorder (**OUD**) constitute an urgent health crisis. Ample evidence indicates that risk for OUD is heritable. As a surrogate (or proxy) for OUD, we explored the genetic basis of using opioids 'not as prescribed'. We hypothesized that misuse of opiates might be a heritable risk factor for OUD. To test this hypothesis, we performed a genome-wide association study (**GWAS**) of problematic opioid use (**POU**; 'ever taking opioid prescriptions not as prescribed') in 132,113 23andMe research participants of European ancestry (N_{cases}=27,805). Our GWAS identified two genome-wide significant loci (rs3791033, an intronic variant of *KDM4A*; rs640561, an intergenic variant near *LRRIQ3*). POU showed a positive genetic correlation with opioid dependence and OUD, as measured in the largest available GWAS (rg=0.57-0.80). We also identified numerous additional genetic correlations with POU, including alcohol dependence (rg=0.74), smoking initiation (rg=0.63), pain relief medication intake (rg=0.49), major depressive disorder (rg=0.44), chronic pain (rg=0.42), insomnia (rg=0.39), and loneliness (rg=0.28). Although POU was positively genetically correlated with risk-taking (rg=0.38), conditioning POU on risk-taking did not substantially alter the magnitude or direction of these genetic correlations, suggesting that POU does not simply reflect a general tendency for risky behavior. We conclude that opioid misuse can be measured in population-based cohorts and provides a cost-effective complementary strategy for understanding the genetic basis of OUD.

INTRODUCTION

Opioid use disorders (**OUD**) represent a global epidemic [1]. Every day, 128 people in the United States die after overdosing on opiates. The pathway to opiate addiction has changed over the last few decades. In the 1960s, more than 80% of people who began using opioids initiated with heroin [2]; by 2013, nearly 80% of opioid users reported that their first regular opioid was a prescription pain reliever [3]. The misuse of and addiction to opiates—including prescription pain relievers, heroin, and synthetic opioids such as fentanyl— is thus a serious emergency that affects public health as well as social and economic welfare. More recently, the COVID-19 pandemic made it increasingly difficult for individuals with OUD to access treatment and impacted mental health, triggering both initial and continued use of opioids [4], which is likely to further increase rates of OUD [5].

Although OUD is known to be moderately heritable [6, 7], genomic discovery has been severely limited due to the complexity of obtaining large, well-characterized samples of cases and opiate-exposed controls [e.g. 6, 8–10]. The largest genome-wide association study (**GWAS**) to date (~10K cases) identified only one locus in *OPRM1*, which encodes the μ-opioid receptor [7]. Larger sample sizes are urgently needed to identify additional loci associated with OUD.

An alternative, complementary approach is to study the genetic liability for OUD across different stages, particularly the transition from use to misuse. By taking this approach, we can explore specific aspects of OUD liability, such as initial use, subjective drug response, transition to hazardous use, dependence, withdrawal and relapse. This strategy can be applied in population-based cohorts, allowing for a dramatic increase in sample sizes for a fraction of the cost.

In the present study, we pursued a novel strategy in which we measured problematic prescription opioid use (**POU**) in 23andMe research participants (n=132,113) of European ancestry. We asked a single question “Have you ever in your life used prescription painkillers (taken not as prescribed), e.g., Vicodin, Oxycontin?”, and conducted a GWAS defining cases as those who answered ‘yes’. We used a variety of bioinformatic analyses to explore the relationship between POU and OUD as well as substance use disorder related behaviors. We also explored POU’s genetic relationship with various psychiatric, behavioral and medical conditions.

MATERIALS AND METHODS

GWAS Cohort and phenotype

We utilized a cohort of 132,113 male and female research participants of European ancestry. All participants were drawn from the customer base of 23andMe, Inc., a direct-to-consumer genetics company. Participants provided informed consent and participated in the research online, under a protocol approved by the external AAHRPP-accredited IRB, Ethical & Independent Review Services (www.eandireview.com). During four months in 2015 and 14 months in 2018-2020, participants responded to a decision-making survey that, depending on branching logic, included up to 139 questions pertaining to aspects of impulsivity and substance use and abuse. A single item in this survey asked “Have you ever in your life used prescription painkillers (taken not as prescribed), e.g., Vicodin, Oxycontin?”. A total of 132,113 individuals responded, with 27,805 (10,164 males) answering ‘yes’ (cases) and the remaining 104,308 (36,246 males) answering ‘no’ (controls). Controls were not screened for prior opioid use so they represent a combination of individuals who have taken opioids but only as prescribed, and others who have never taken an opioid. Only individuals who were categorized as being of European ancestry based on empirical genotype data [11] were included in this study. Basic demographic information about this sample is presented in **Supplementary Table 1**.

Genome-wide association analysis

DNA extraction and genotyping were performed on saliva samples by CLIA-certified and CAP-accredited clinical laboratories of Laboratory Corporation of America. Quality control, imputation, and genome-wide analysis were performed by 23andMe (**Supplementary Table 2**; see [12, 13] for further information about genotyping and quality control).

As previously described [13, 14], 23andMe’s analysis pipeline performs logistic regression assuming an additive model for allelic effects (**Supplementary Material**). Covariates included age (inverse-normal transformed), sex, the top five principal components of genotype, and indicator variables for genotyping platforms. P-values were not corrected for genomic control.

Biological annotation, gene and transcriptome-based association analyses

We used a variety of bioinformatic methods to further characterize the loci identified by the GWAS. First, we used the default version of the FUMA web-based platform v1.3.6a [15] to identify independent SNPs (r2>0.10) and to study their functional consequences. We also used MAGMA v1.08 [15, 16] to perform competitive gene-set and pathway analyses. SNPs were

mapped to 18,546 protein-coding genes from Ensembl (build 85). Gene-sets were obtained from Msigdb v7.0 ("Curated gene sets", "GO terms"). We also used Hi-C coupled MAGMA (**H-MAGMA**; [17]) to assign non-coding (intergenic and intronic) SNPs to genes based on their chromatin interactions. Exonic and promoter SNPs were assigned to genes based on physical position. H-MAGMA uses four Hi-C datasets, which were derived from fetal brain, adult brain, iPSC-derived neurons and iPSC-derived astrocytes (https://github.com/thewonlab/H-MAGMA).

Lastly, we used MetaXcan v0.7.0 (an extension of S-PrediXcan v0.6.2 [18]) to identify specific eQTL-linked genes associated with POU. This approach uses genetic information to predict transcript abundance in 13 brain tissues, and tests whether the predicted transcripts correlate with POU. S-PrediXcan uses pre-computed tissue weights from the Genotype-Tissue Expression (**GTEx**) v8 project database (https://www.gtexportal.org/) as the reference transcriptome dataset. For S-PrediXcan and MetaXcan analyses we chose to use sparse (elastic net) prediction models, which are available at http://predictdb.hakyimlab.org/. We applied a conservative Bonferroni correction based on the total number of gene-tissue pairs tested (14,159 gene-tissue pairs tested; α=3.53E-06).

Gene-drug interaction analysis

We examined the 17 genes that were significantly associated with POU in the MAGMA gene-based analysis (10% FDR) for known interactions with prescription medications using the Drug Gene Interaction Database v3.0 (dgidb.genome.wustl.edu) [19]. We used the Anatomical Therapeutic Chemical (**ATC**) classification system to determine the second level classification of each medication we identified. ATC classifications for medications were retrieved from the Kyoto Encyclopedia of Genes and Genomics (**KEGG**; https://www.genome.jp/kegg/drug/) and the World Health Organization Collaborating Center for Drug Statistics Methodology (https://www.whocc.no/atc_ddd_index/). We used the R package circlize v0.4.1 [20] to visualize the interactions between each gene and the ATC classifications of drugs it interacts with.

Heritability

We used the LD Score regression (**LDSC** [21]) python package to estimate the heritability explained by SNPs (SNP-h_2). We used pre-computed LD scores ("eur_w_ld_chr/"), which are publicly available (https://data.broadinstitute.org/alkesgroup/LDSCORE/). LD scores were computed for every SNP using individuals from European ancestry from the 1000 Genomes Project. We restricted the analysis to well-imputed SNPs, filtered to HapMap3 SNPs, with MAF above 1%. We removed InDels, structural variants, strand-ambiguous SNPs and

SNPs with extremely large effect sizes (χ^2>80). Heritability was calculated on the liability scale by accounting for differences in population prevalence (4%) and sample prevalence (21%). The population prevalence was retrieved from the 2018 National Survey on Drug Use and Health [22].

Genetic correlation analyses

We used LDhub or local LDSC [21] to calculate genetic correlations (r_g) between POU and 935 other traits or diseases (852 from LDHub and 83 local) [21]. Local traits were selected based on previously known phenotypic associations between OUD and other substance use disorder phenotypes and related traits (e.g., cannabis use disorder, various measures of impulsivity) that were not available on LDhub. We used the standard Benjamini–Hochberg false discovery rate correction (FDR 5%) to correct for multiple testing. We also calculated a Bonferroni correction for 935 comparisons (α=5.35-05); however, this correction is overly conservative because many of the 935 traits are highly correlated with one another.

mtCOJO

We used mtCOJO [20] to individually condition the POU summary statistics on loci associated with other comorbid traits, including risk-taking behavior [23], smoking [24], cannabis use disorder [25], alcohol dependence [26], chronic pain [27], and OUD [28]. This analysis allowed us to examine whether the genetic associations with POU would be preserved when controlling for those covariate phenotypes. To test as many SNPs while preserving computational efficiency, we used a p-value threshold of 2E-05, 5E-08,1E-05, 1E-05, 5E-08, and 1E-05, respectively for risk-taking behavior, smoking, cannabis use disorder, alcohol dependence, chronic pain, and OUD. We then computed genetic correlations using the POU summary statistics adjusted for the covariates of interest.

Unsupervised Learning to Determine POU clustering

Previous studies have shown that consumption and misuse/dependence phenotypes have a distinct genetic architecture. To explore whether POU clustered more with consumption or misuse/dependence phenotypes we used a data-driven unsupervised machine learning method known as agglomerative hierarchical clustering analysis (**HCA**) [29, 30]. HCA forms clusters iteratively by creating groups and successively joining or splitting those groups based on a prespecified algorithm [29]. Agglomerative nesting (**AGNES**) is a bottom-up process focused on individual traits to structure. Agglomerative clustering was chosen as this allowed us to compare different algorithms to maximize for the dissimilarity on each branch, with Ward's

minimum variance method performing best. All models were fit in R using the *cluster* package [29].

The product of HCA is a dendrogram, formed with multiple brackets called “branches”. Phenotypes on the same branch are more similar to each other based on their pairwise genetic associations with each other and with all other phenotypes on that branch. Branches can form subbranches of more specific clustering.

RESULTS

Genome-wide association analysis, biological annotation, gene and transcriptome-based association analyses

We examined 11,311,983 SNPs in all 132,113 study participants. The inflation factor of the GWAS was λ_{GC}=1.097 with an LDSC intercept of 1.004 (SE=0.008), suggesting that the majority of the inflation was due to polygenicity. The SNP heritability of POU on the liability scale was SNP-h_2=0.04 +/- 0.01.

We identified two genome-wide significant loci: rs3791033 near the genes *KDM4A* and *PTPRF* (*p*=3.80E-08), and rs640561 near *LRRIQ3* (*p*=3.80E-08; **Supplementary Table 3** and **Supplementary Figures 1-2** for locus zoom plots). Both loci are on chromosome 1, but rs3791033 and rs640561 are independent (R^2<0.001).

A phenome-wide scan in the UK Biobank (UKB; N>360,000) revealed that rs3791033 has been previously implicated in smoking phenotypes (e.g., ever smoked regularly, *p*=1.01E-12), other psychiatric conditions, such as ADHD (*p*=2.76E-10), and educational outcomes, such as educational attainment (*p*=1.01E-10). *KDM4A*, which is the nearest gene, has been previously implicated with similar traits across several independent GWAS studies (*e.g.*, [23, 31, 32]; **Supplementary Table 4**). The other gene in this region, *PTPRF*, has been previously identified in studies of smoking and other behaviors [14, 23, 33]; **Supplementary Table 4**).

For the second genome-wide significant SNP (rs640561), a phenome-wide scan in UKB revealed nominal associations (*p*>7.00E-03) with neurological, metabolic and other psychiatric traits, such as ADHD (**Supplementary Table 4**). *LRRIQ3*, which is the nearest gene, has previously been associated with lifetime smoking [23].

We did not observe GWAS-significant associations (p>0.05) with any of the 12 SNPs that have been previously associated with OUD and opioid dependence [6, 8–10] (**Supplementary Table 5**).

In addition to the GWAS, we performed several gene-based analyses. *KDM4A* was implicated by both MAGMA (FDR 5%) and H-MAGMA (**Supplementary Tables 6, 7**). *PTPRF* was implicated by H-MAGMA (**Supplementary Table 7**). None of these analyses identified *LRRIQ3*. H-MAGMA analysis also identified *ARTN* (previously showing a nominal association with tea consumption in UKB; **Supplementary Table 7**). S-PrediXcan did not identify any significant genes (**Supplementary Table 8**).

Lastly, gene-set analysis in MAGMA identified one significant gene set, which is involved in the activation of Phospholipase D (**Supplementary Table 9**). Intriguingly, activation of Phospholipase D2 modulates agonist□induced μ□opioid receptor desensitization and resensitization [34].

Gene-Drug Analysis

For this analysis we relaxed the FDR threshold from 5 to 10%, which produced a set of 17 genes; 3 of these 17 genes (*KDM4A, PTPRF, CACNA2D2*) had a total of 464 interactions with drugs belonging to ATC drug classes (see **Figure 4** and **Supplementary Table 10**). The most abundant second-level ATC classifications identified were antibacterials for systemic use, psycholeptics, and gynecological antiinfectives and antiseptics. Interactions with *KDM4A* included both dopamine agonists and agents, as well as adrenergic agents, drugs used in alcohol dependence such as disulfiram, opioid anaesthetics, such as sufentanil citrate, and antidepressants.

Genetic correlation analyses, and clustering solution

We performed a series of genetic correlations and identified consistent, moderate to strong associations with well-known OUD comorbid factors. We considered 935 traits, which represented 17 categories (substance use, psychiatric, impulsivity, personality, lifestyle, education, cognitive, health, longevity, metabolic, eyes, teeth, sleep, pain, reproductive, medication and anthropometric). We identified significant genetic associations between POU and 253 of these 935 traits (FDR 5%; **Supplementary Table 11**).

Notably, POU showed positive genetic correlations with OUD as measured by the MVP (r_g=0.58, p=1.13E-07), and opioid dependence versus unexposed controls (European) as measured by the PGC (r_g=0.80, p=3.24E-04; opioid dependence and opioid exposed controls showed no significant heritability and therefore could not be tested), and with opioid medication use (r_g=0.46, p=2.25E-12) in participants from UKB.

Beyond opioid-related traits, we also identified strong genetic correlations with alcohol, nicotine, and cannabis-related traits, from initiation (nicotine, r_g=0.63, p=7.14E-38; cannabis, r_g=0.35, p=1.06E-06), to high levels of consumption (drinks per week, r_g=0.36; p=2.11E-15), to misuse (AUDIT-P: r_g=0.42, p=5.60E-08) and dependence (alcohol, r_g=0.74, p=7.21E-08; the Fagerström Test for Nicotine Dependence, r_g=0.60, p=2.98E-10; cannabis use disorder, r_g=0.63, p=2.36E-13), and other smoking behaviors, such as maternal smoking around birth (r_g=0.64,

p=3.44E-26). Interestingly, hierarchical HCA with AGNES found that POU clustered closely with substance use disorders, as opposed to consumption phenotypes (**Supplementary Figure 3**).

We also identified a number of associations with UKB pain phenotypes, including positive genetic correlations with several pain ICD diagnoses (*e.g.*, pain in throat, pelvic pain, r_g range=[0.37-0.39], p<8.22E-05), and prevalent pain conditions such as back and knee pain, and headaches (r_g range=[0.21-0.42], p<8.50E-03). As expected, we found positive genetic associations with chronic pain (r_g=0.42, p=4.04E-17), and higher pain relief medication intake (*e.g.*, Ibuprofen, r_g=0.49; p=1.61E-10).

We also identified positive genetic associations between POU and other behavioral and psychiatric traits, such as mood swings (r_g=0.40, p=1.04E-13), risk-taking behavior (r_g=0.38, p=2.09E-08), major depressive disorder (r_g=0.44, p=1.59E-10), insomnia (r_g=0.39, p=6.88E-11), loneliness (r_g=0.28, p=4.81E-06) and irritability (r_g=0.27, p=1.74E-05).

Lastly, POU was negatively correlated with educational attainment (*e.g.*, obtaining a college or university degree, r_g=-0.47; p=6.13E-25) and intelligence (r_g=-0.41; p=7.22E-11).

These genetic correlations generally remained consistent after conditioning POU on risk-taking behavior, nicotine and cannabis use disorders (**Supplementary Tables 12, 13** and **14** respectively). These associations were also broadly consistent when we conditioned on pain (**Supplementary Tables 16**), and OUD (**Supplementary Tables 17**). Intriguingly, when we corrected for alcohol dependence some of the associations increased, particularly with OUD and opioid dependence, cannabis use disorder, and depressive symptoms, whereas some associations dramatically decreased, such as nicotine dependence as measured via FTND (**Supplementary Table 15**).

DISCUSSION

In this study, we performed a GWAS of problematic opioid use ('ever taking prescription painkillers not as prescribed') in 132,113 23andMe research participants of European ancestry. This represented a novel approach to studying OUD in a population-based cohort. Our results show that this single question captured a genetic signal that is correlated with signals from well-characterized cohorts that have been clinically diagnosed with OUD. Notably, the genetic correlations with OUD persisted even after correcting for risk-taking behavior and other putatively similar dimensional phenotypes. While previous GWAS of OUD have used clinically-ascertained cohorts, our results suggest that POU provides a cost-effective alternative to diagnosed OUD that is viable in non-clinically ascertained populations, making it possible to rapidly obtain large sample sizes that can aid in OUD gene discovery.

Obtaining a large enough sample size to effectively identify risk loci has been a common obstacle for OUD GWAS [35]. Our approach represents a new and qualitatively different way of characterizing individuals at high risk for OUD and. Our approach is stimulated by the idea of fractionating OUD and looking at POU as an early stage of misuse [36]. In particular, we were motivated to characterize the common mechanism of taking opioids not as prescribed, which can lead to abuse. Sometimes referred to as minimal phenotyping, where a complex trait is reduced to a single yes or no question [36], the polygenic signal of POU is nevertheless informative for aspects of OUD risk that is not intended to be a 'noisy' measurement of the true underlying disease.

The most important finding in this study is the identification of the polygenic architecture of POU that is highly comparable to the findings from GWAS of clinically-ascertained OUD cohorts. We observed a positive association of POU with OUD and opioid dependence measured by two of the largest available GWAS, MVP and PGC, and with a GWAS of opioid medication use in UKB participants. As might have been expected, we also observed strong positive genetic correlations with alcohol dependence and tobacco smoking, as well as with various psychiatric traits associated with OUD, including mood swings, risk-taking, anxiety, depression, and insomnia. These sets of genetic correlations mirror those from previous GWAS of clinically ascertained OUD samples [6–8]. However, we also showed that the overlap is not complete, and whether POU could be an early manifestation of risk for subsequent OUD is not directly explored by our study.

We initially speculated that some of the genetic signal associated with POU could be confounded by genetic factors unrelated to opioid misuse, such as risk-taking behaviors,

comorbid substance use (such as tobacco or cannabis use disorders, and even OUD diagnosis) or pain. Previous reports have indicated that genetic studies using an unscreened sample for relevant comorbid conditions [6], including all individuals being exposed to opioids at least once, can introduce biases in the genetic analyses. However, when we adjusted for these known risk factors, the signal remained consistent, suggesting that we are capturing a signal that is specific to opioids and not merely confounded by signals associated with the secondary phenotypes.

This study has also revealed novel biological insights for opioid misuse. While we were unable to replicate the previously reported GWAS signals associated with OUD, we discovered two novel loci associated with POU. Both lead SNPs or nearby genes have shown previous associations with other psychiatric traits, particularly smoking phenotypes, which are known to be highly comorbid with OUD and were also strongly correlated with POU. In particular, we revealed novel associations with the gene *KDM4A*, which has known interactions with medications that are used to treat other psychiatric conditions, such as depression, or target dopaminergic or serotonergic systems. Furthermore, gene-set analyses revealed that the polygenic signal for POU is implicated in pathways that modulate μ-opioid receptor sensitization. Taken together, these findings suggest a potential avenue for identifying new therapeutic targets for problematic opioid use.

This study is not without limitations. The screening and composition of the control group is crucial for studies of substance use and abuse [6, 27, 35]. Our controls indicated they had never used opiates not as prescribed, which could include individuals who had simply never used an opiate, along with individuals who had anywhere from minimal to extensive experience with opiates, but had never deviated from their prescribed use. POU might be even more useful if we had additional data that allowed us to exclude individuals that have never used opioids. Using an unscreened control group can lead to considerable phenotypic heterogeneity across samples [7, 37]. A second important limitation is our inability to evaluate whether the individuals included in our analyses suffered from mild versus severe pain. Although high genetic predisposition for chronic pain may itself be a risk factor for OUD, this concern was partially addressed by our mtCOJO analysis in which we conditioned on pain. That analysis revealed that the association between POU and OUD was persevered even after correcting for pain, suggesting that POU was not correlated with OUD via its ability to capture genetic predisposition to pain.

In summary, we have shown that the genetic signature for opioid misuse that we were able to capture via self-report in an unselected population is similar to genetic risk for OUD. Our

work sets the stage for future analyses incorporating a multivariate framework (*e.g.*, genomic Structural Equation Modelling), and larger sample sizes. Our approach provides new, cost-efficient tools for genetic research related to OUD and provides insights into an intermediate behavior that may set the stage for later transition to OUD.

Data availability. We have provided summary statistics for the top 10,000 SNPs (Supplementary Data Set). Full GWAS summary statistics will be made available through 23andMe to qualified researchers under an agreement with 23andMe that protects the privacy of the 23andMe participants. Interested investigators should email dataset-request@23andme.com and reference this paper for more information.

Acknowledgements. We would like to thank the research participants and employees of 23andMe (Michelle Agee, Babak Alipanahi, Adam Auton, Robert K. Bell, Katarzyna Bryc, Sarah L. Elson, Pierre Fontanillas, Nicholas A. Furlotte, David A. Hinds, Karen E. Huber, Aaron Kleinman, Nadia K. Litterman, Jennifer C. McCreight, Matthew H. McIntyre, Joanna L. Mountain, Elizabeth S. Noblin, Carrie A.M. Northover, Steven J. Pitts, J. Fah Sathirapongsasuti, Olga V. Sazonova, Janie F. Shelton, Suyash Shringarpure, Chao Tian, Joyce Y. Tung, Vladimir Vacic, and Catherine H. Wilson) for making this work possible.

We would also like to thank The Externalizing Consortium, for sharing the GWAS summary statistics of externalizing. The Externalizing Consortium: Principal Investigators: Danielle M. Dick, Philipp Koellinger, K. Paige Harden, Abraham A. Palmer. Lead Analysts: Richard Karlsson Linnér, Travis T. Mallard, Peter B. Barr, Sandra Sanchez-Roige. Significant Contributors: Irwin D. Waldman. The Externalizing Consortium has been supported by the National Institute on Alcohol Abuse and Alcoholism (R01AA015416 -administrative supplement), and the National Institute on Drug Abuse (R01DA050721). Additional funding for investigator effort has been provided by K02AA018755, U10AA008401, P50AA022537, as well as a European Research Council Consolidator Grant (647648 EdGe to Koellinger). The content is solely the responsibility of the authors and does not necessarily represent the official views of the above funding bodies. The Externalizing Consortium would like to thank the following groups for making the research possible: 23andMe, Add Health, Vanderbilt University Medical Center's BioVU, Collaborative Study on the Genetics of Alcoholism (COGA), the Psychiatric Genomics Consortium's Substance Use Disorders working group, UK10K Consortium, UK Biobank, and Philadelphia Neurodevelopmental Cohort.

We are grateful to Renato Polimanti for his comments on an earlier draft of this manuscript.

Funding. MVJ, YH, SSR and AAP were supported by funds from the California Tobacco-Related Disease Research Program (TRDRP; Grant Number 28IR-0070 and T29KT0526).

PF and SLE are employees of 23andMe, Inc., and hold stock or stock options in 23andMe. The other authors report no conflict of interest.

References

1. Kolodny A, Courtwright DT, Hwang CS, Kreiner P, Eadie JL, Clark TW, et al. The Prescription Opioid and Heroin Crisis: A Public Health Approach to an Epidemic of Addiction. Annu Rev Public Health. 2015;36:559–574.
2. Cicero TJ, Ellis MS, Surratt HL, Kurtz SP. The Changing Face of Heroin Use in the United States: A Retrospective Analysis of the Past 50 Years. JAMA Psychiatry. 2014;71:821.
3. Muhuri PK. Associations of Nonmedical Pain Reliever Use and Initian of Heroin Use in the United States. CBHSQ Data Rev. 2013. https://www.samhsa.gov/data/sites/default/files/DR006/DR006/nonmedical-pain-reliever-use-2013.htm.
4. Khatri UG, Perrone J. Opioid Use Disorder and COVID-19: Crashing of the Crises. J Addict Med. 2020;14:e6–e7.
5. Volkow ND, Blanco C. The changing opioid crisis: development, challenges and opportunities. Mol Psychiatry. 2021;26:218–233.
6. Polimanti R, Walters RK, Johnson EC, McClintick JN, Adkins AE, Adkins DE, et al. Leveraging genome-wide data to investigate differences between opioid use vs. opioid dependence in 41,176 individuals from the Psychiatric Genomics Consortium. Mol Psychiatry. 2020;25:1673–1687.
7. Zhou H, Rentsch CT, Cheng Z, Kember RL, Nunez YZ, Sherva RM, et al. Association of *OPRM1* Functional Coding Variant With Opioid Use Disorder: A Genome-Wide Association Study. JAMA Psychiatry. 2020;77:1072.
8. Hancock DB, Levy JL, Gaddis NC, Glasheen C, Saccone NL, Page GP, et al. Cis-Expression Quantitative Trait Loci Mapping Reveals Replicable Associations with Heroin Addiction in OPRM1. Biol Psychiatry. 2015;78:474–484.
9. Nelson EC, Agrawal A, Heath AC, Bogdan R, Sherva R, Zhang B, et al. Evidence of CNIH3 involvement in opioid dependence. Mol Psychiatry. 2016;21:608–614.
10. 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;84:762–770.
11. Durand EY, Do CB, Mountain JL, Macpherson JM. Ancestry Composition: A Novel, Efficient Pipeline for Ancestry Deconvolution:16.
12. Eriksson N, Macpherson JM, Tung JY, Hon LS, Naughton B, Saxonov S, et al. Web-Based, Participant-Driven Studies Yield Novel Genetic Associations for Common Traits. PLoS Genet. 2010;6:e1000993.
13. Hyde CL, Nagle MW, Tian C, Chen X, Paciga SA, Wendland JR, et al. Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. Nat Genet. 2016;48:1031–1036.
14. Lam M, Trampush JW, Yu J, Knowles E, Davies G, Liewald DC, et al. Large-Scale Cognitive GWAS Meta-Analysis Reveals Tissue-Specific Neural Expression and Potential Nootropic Drug Targets. Cell Rep. 2017;21:2597–2613.
15. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun. 2017;8:1826.

16. Leeuw CA de, Mooij JM, Heskes T, Posthuma D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. PLOS Comput Biol. 2015;11:e1004219.

17. Sey NYA, Hu B, Mah W, Fauni H, McAfee JC, Rajarajan P, et al. A computational tool (H-MAGMA) for improved prediction of brain-disorder risk genes by incorporating brain chromatin interaction profiles. Nat Neurosci. 2020;23:583–593.

18. Barbeira AN, Dickinson SP, Bonazzola R, Zheng J, Wheeler HE, Torres JM, et al. Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. Nat Commun. 2018;9:1825.

19. Griffith M, Griffith OL, Coffman AC, Weible JV, McMichael JF, Spies NC, et al. DGIdb: mining the druggable genome. Nat Methods. 2013;10:1209–1210.

20. Gu Z, Gu L, Eils R, Schlesner M, Brors B. circlize implements and enhances circular visualization in R. Bioinformatics. 2014;30:2811–2812.

21. Bulik-Sullivan BK, Loh P-R, Finucane HK, Ripke S, Yang J, Patterson N, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet. 2015;47:291–295.

22. Lipari RN. Key Substance Use and Mental Health Indicators in the United States: Results from the 2018 National Survey on Drug Use and Health. 2018:82.

23. 23and Me Research Team, eQTLgen Consortium, International Cannabis Consortium, Social Science Genetic Association Consortium, Karlsson Linnér R, Biroli P, et al. Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. Nat Genet. 2019;51:245–257.

24. 23andMe Research Team, HUNT All-In Psychiatry, Liu M, Jiang Y, Wedow R, Li Y, et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. Nat Genet. 2019;51:237–244.

25. Johnson EC, Demontis D, Thorgeirsson TE, Walters RK, Polimanti R, Hatoum AS, et al. A large-scale genome-wide association study meta-analysis of cannabis use disorder. Lancet Psychiatry. 2020;7:1032–1045.

26. 23andMe Research Team, Walters RK, Polimanti R, Johnson EC, McClintick JN, Adams MJ, et al. Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. Nat Neurosci. 2018;21:1656–1669.

27. Johnston KJA, Adams MJ, Nicholl BI, Ward J, Strawbridge RJ, Ferguson A, et al. Genome-wide association study of multisite chronic pain in UK Biobank. PLOS Genet. 2019;15:e1008164.

28. Zhou H, Rentsch CT, Cheng Z, Kember RL, Nunez YZ, Tate JP, et al. GWAS including 82,707 subjects identifies functional coding variant in *OPRM1* gene associated with opioid use disorder. Genetic and Genomic Medicine; 2019.

29. Maechler,M., Rousseeuw, P., Struyf, A., Hubert, M., Hornik, K. Cluster: Cluster Analysis Basics and Extentions. R Software; 2013.

30. Bunge JA, Judson DH. Data Mining. Encycl. Soc. Meas., Elsevier; 2005. p. 617–624.

31. Artigas MS, Sánchez-Mora C, Rovira P, Richarte V, García-Martínez I, Pagerols M, et al. Attention-Deficit/Hyperactivity Disorder and lifetime cannabis use: genetic overlap and causality. 2021:20.

32. Pirastu N, Pirastu M, Polasek O, Posthuma D, Power C, Samani NJ, et al. Gene discovery and polygenic prediction from a 1.1-million-. 2019:29.

33. Davies G, Marioni R, Liewald D, Hill W, Hagenaars S, Harris S, et al. Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N = 112 151). Mol Psychiatry. 2016:10.

34. Koch T, Brandenburg L-O, Liang Y, Schulz S, Beyer A, Schröder H, et al. Phospholipase D2 modulates agonist-induced μ-opioid receptor desensitization and resensitization: PLD2 modulates μ-opioid receptor desensitization. J Neurochem. 2003;88:680–688.

35. Crist RC, Reiner BC, Berrettini WH. A review of opioid addiction genetics. Curr Opin Psychol. 2019;27:31–35.

36. Sanchez-Roige S, Palmer AA. Emerging phenotyping strategies will advance our understanding of psychiatric genetics. Nat Neurosci. 2020;23:475–480.

37. Song W, Kossowsky J, Torous J, Chen C-Y, Huang H, Mukamal KJ, et al. Genome-wide association analysis of opioid use disorder: A novel approach using clinical data. Drug Alcohol Depend. 2020;217:108276.

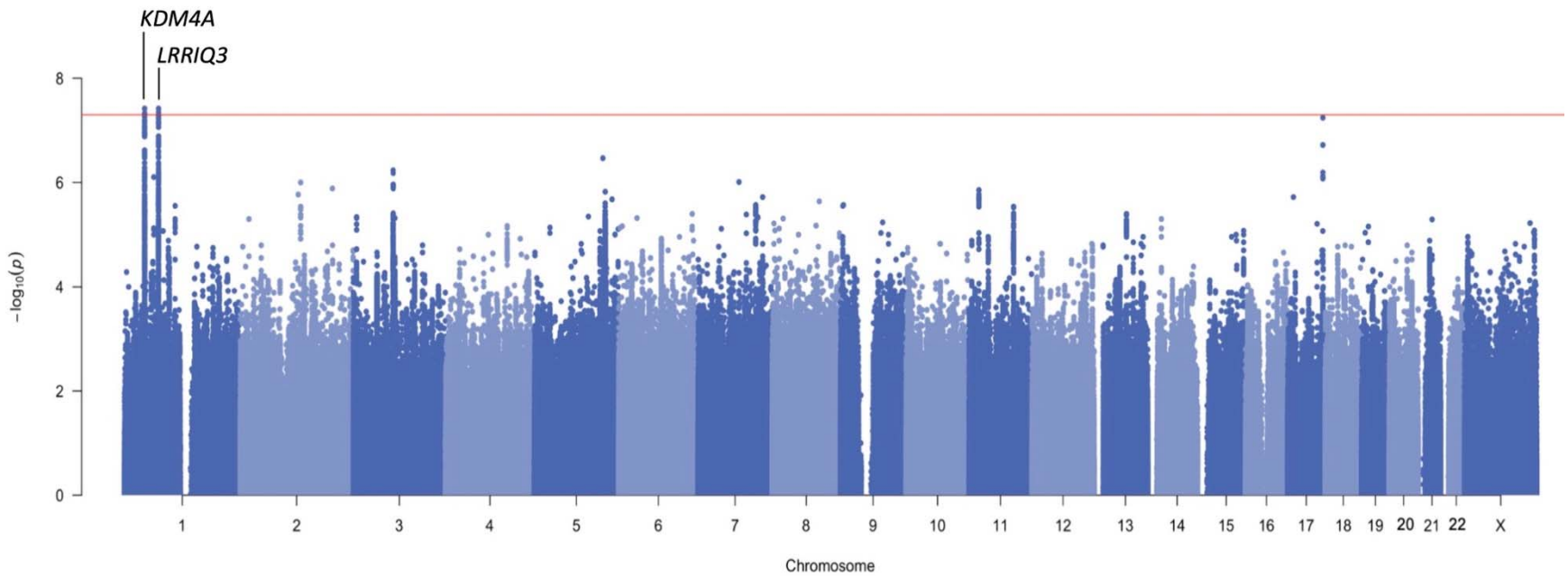

Figure 1. Genome-wide association analysis for problematic opioid use. The nearest genes are labeled for the 2 lead SNPs (rs3791033, rs640561). The x-axis shows chromosomal position and the y-axis shows significance on –log10 scale. The horizontal red line denotes genome-wide significance (*p*=5.00E-08).

Figure 2. LDSC FDR-significant genetic correlations with POU. Traits with positive genetic correlation (r_g) values are plotted above the line; traits with negative r_g values are plotted below the line. All traits surpass 5% FDR correction for multiple testing.

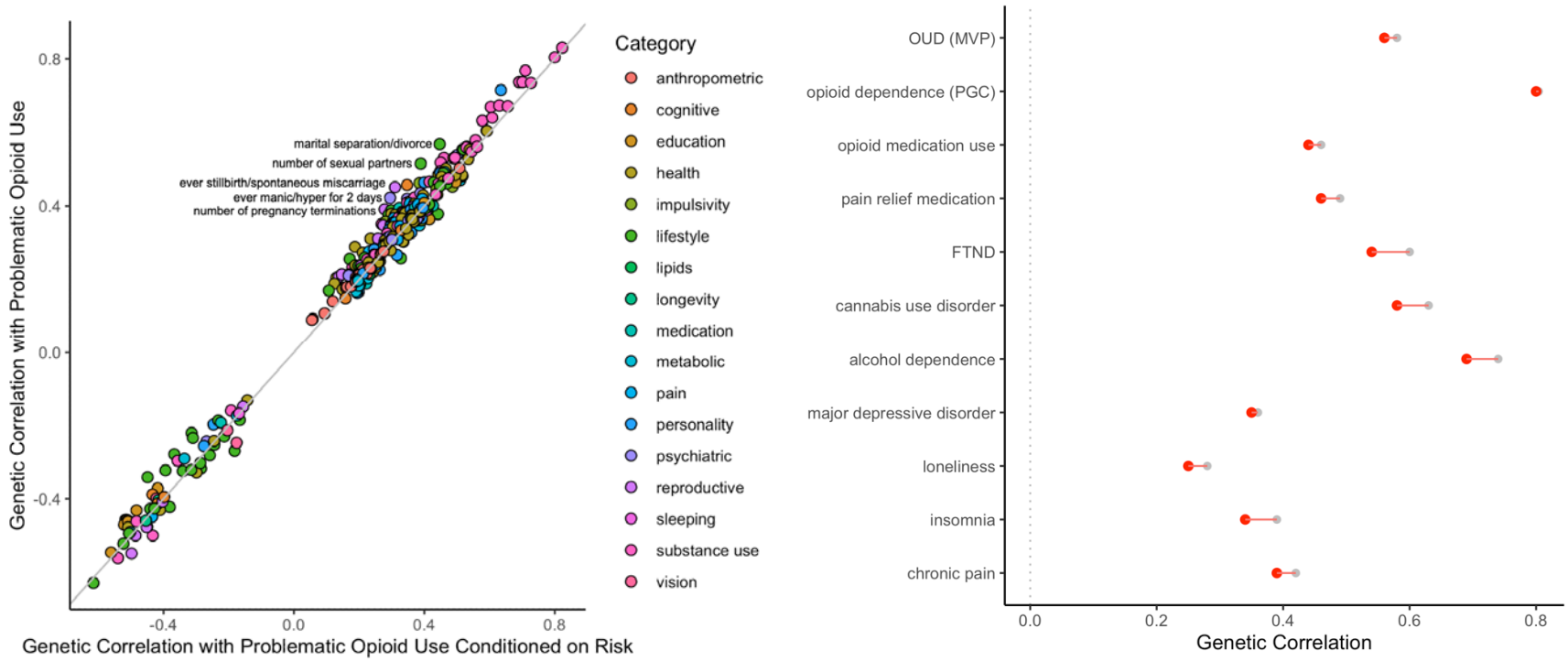

Figure 3. Genetic correlations (r_g) with POU before and after conditioning on risk-taking behavior. All FDR-significant results are plotted on the left panel, selected relevant traits are shown on the right panel (original r_g in grey; corrected r_g in red). The top 5 traits with biggest change in r_g value are labelled (left).

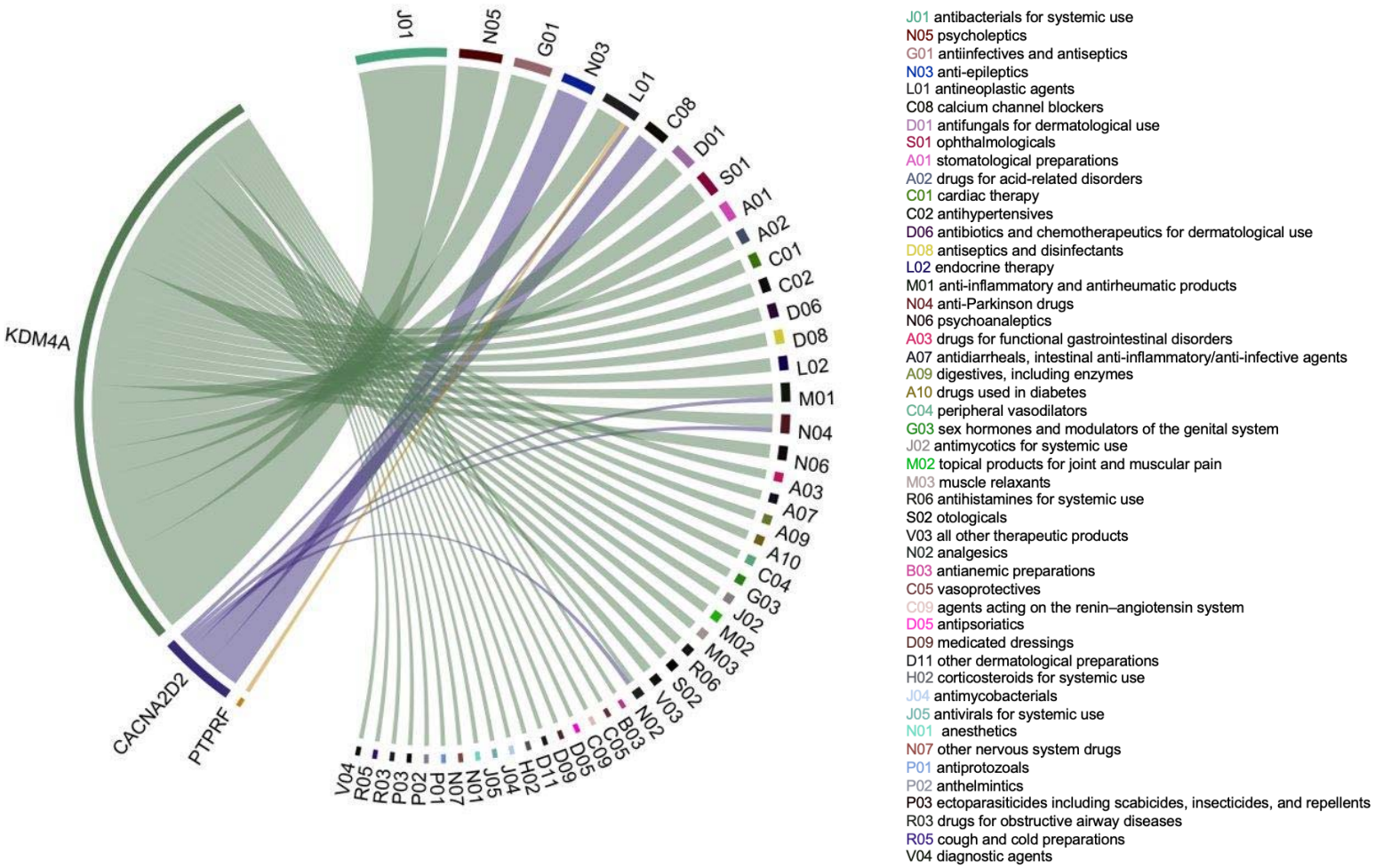

Figure 4. Chord diagram of genes significantly associated with POU at 10% FDR and the Anatomical Therapeutic Chemical classifications of drugs. The width of each line is determined by the number of drugs known to interact with each gene.