



Published in final edited form as:

Addict Biol. 2020 March ; 25(2): e12811. doi:10.1111/adb.12811.

Genome-wide Scan Identifies Opioid Overdose Risk Locus Close to *MCOLN1*

Zhongshan Cheng^{a,*}, Bao-Zhu Yang^{a,*}, Hang Zhou^a, Yaira Nunez^a, Henry R. Kranzler^c, Joel Gelernter^{a,b,*}

a. Division of Human Genetics, Department of Psychiatry, Yale University School of Medicine and VA CT Healthcare Center, New Haven, CT

b. Departments of Genetics and Neuroscience, Yale University School of Medicine, New Haven, CT

c. Center for Studies of Addiction, Department of Psychiatry, University of Pennsylvania Perelman School of Medicine and Crescenz Veterans Affairs Medical Center, Philadelphia, PA

Abstract

The US is experiencing the worst opioid overdose (OpOD) crisis in its history. We carried out a genome-wide association study on OpOD severity among 3,477 opioid-exposed individuals, 1,019 of whom experienced opioid overdoses, including 2,032 European Americans (EAs) (653 overdose cases) and 1,445 African Americans (AAs) (366 overdose cases). Participants were scored 1–4 based on their reported overdose status and the number of times that medical treatment was required. Genome-wide association study (GWAS) of EAs and AAs separately resulted in two genome-wide significant (GWS) signals in AAs but none in EAs. The first signal was represented by three closely-mapped SNPs (rs115208233, rs116181528, and rs114077267) located near *MCOLN1* (Mucolipin-1) and *PNPLA6* (Patatin like phospholipase domain containing 6), and the other signal was represented by rs369098800 near *DDX18*; There were no additional GWS signals in the trans-population meta-analysis, so that post-GWAS analysis focused on these loci. In network analysis, *MCOLN1* was co-expressed with *PNPLA6*, but only *MCOLN1*-associated genes were enriched in functional categories relevant to OpOD, including calcium and cation channel activities; no enrichment was observed for *PNPLA6*-associated genes. Drug repositioning analysis was carried out in the Connectivity Map (CMap) database for *MCOLN1* (*PNPLA6* was not available in CMap) and showed that the opioid agonist drug-induced expression profile is similar to that of *MCOLN1* over-expression, and yielded the highest-ranked expression profile of 83 drug classes. Thus, *MCOLN1* may be a risk gene for OpOD but replication is needed. This knowledge could be helpful in the identification of drug targets for preventing OpOD.

**Correspondence: Joel Gelernter, Department of Psychiatry, Yale University School of Medicine, VA CT 116A2; 950 Campbell Avenue; West Haven, CT 06516; joel.gelernter@yale.edu;

*equal contribution

Potential Conflict of Interest

Dr. Kranzler is a member of the American Society of Clinical Psychopharmacology's Alcohol Clinical Trials Initiative, which was supported in the past three years by AbbVie, Alkermes, Ethypharm, Indivior, Lilly, Lundbeck, Otsuka, Pfizer, Arbor, and Amygdala Neurosciences. Drs. Kranzler and Gelernter are named as inventors on PCT patent application #15/878,640 entitled: "Genotype-guided dosing of opioid agonists," filed January 24, 2018. The other authors have no biomedical financial interests or potential conflicts of interest.

Keywords

Genome-wide association; Opioid overdose; *MCOLNI*; European Americans; African Americans

INTRODUCTION

The United States is experiencing an opioid use and overdose crisis. In 2017, the death toll from drug overdoses hit a record high with more than 72,000 deaths (Centers for Disease Control and Prevention, 2018), and was driven by opioid overdose (OpOD) (Buchanich et al., 2018; Jalal et al., 2018). Clinical symptoms of OpOD include stupor, miosis, and respiratory depression (Boyer, 2012). The best clinical strategy to reverse OpOD requires immediate administration of the opioid antagonist naloxone (Chou et al., 2017).

Genetic factors contribute to opioid dependence (OD) and opioid sensitivity. Several genome-wide association studies (GWAS) have identified SNPs contributing to the risk of OD, including rs60349741, rs62103177, rs12442183 and rs10799590 mapped to *KCNKI*, *KCNG2* (Gelernter et al., 2014), *RGMA* (Cheng et al., 2018) and *CNIH3* (Nelson et al., 2016), respectively. The first two of these risk genes were identified in African Americans (AAs), while the latter two were found in European Americans (EAs). Additionally, rs73568641, ~300kb upstream of *OPRM1*, was associated with therapeutic methadone and morphine dosage in AAs (Smith et al., 2017).

While progress has been made in understanding the genetics of OD risk and opioid dosing, there are no publications that address genetic factors predisposing to OpOD using a genome-wide approach. The diagnosis of a substance use disorder in the previous six months was reported to be the strongest risk factor for OpOD (Webster, 2017), but additional risk factors include the presence of other psychiatric disorders, impaired liver function, impaired vascular and pulmonary function, and a high daily opioid dosage (Nadpara et al., 2018; Zedler et al., 2015). Many of these risk factors are themselves genetically influenced. Variation in opioid metabolism, also genetically influenced, accounts for differential sensitivity to OpOD (Kharasch et al., 2015; White and Irvine, 1999). Thus, knowledge of the specific genetic risk factors for OpOD may guide the clinical care of patients receiving opioids therapeutically and those using the drugs illegally.

We conducted a GWAS of OpOD severity among subjects from the Yale-Penn study of the genetics of substance dependence, in a sample of 5,540 EAs and 3,675 AAs. Of these 9,215 subjects, we analyzed data from 3,477 subjects who were exposed to opioids (used at least once) and responded to questions regarding their overdose histories, resulting in an analysis sample of 2,032 EAs (653 overdose cases) and 1,445 AAs (366 overdose cases).

METHODS AND MATERIALS

Study Participants and the OpOD Trait

We collected samples of EA and AA subjects for genetic studies of substance dependence at five eastern US sites to participate in studies of the genetics of drug (opioid or cocaine) or alcohol dependence (Gelernter et al., 2014). Each participant was carefully phenotyped

using the Semi-Structured Assessment for Drug Dependence and Alcoholism (SSADDA) (Gelernter et al., 2014; Pierucci-Lagha et al., 2007). Subjects comprised two cohorts: Yale-Penn 1 (n=5,540) and Yale-Penn 2 (n=3,675). Each cohort includes some related subjects and they were assessed identically for phenotype but differ by the genotyping array employed. Written informed consent that was approved by the institutional review board at each site was obtained from each participant, and certificates of confidentiality were obtained from the National Institute on Drug Abuse and the National Institute on Alcohol Abuse and Alcoholism.

The OpOD trait was defined for subjects who were exposed to opioids at least once in their lives (n=4,989) and provided information on their OpOD experience (n=3,477); subjects with missing answers for OpOD were excluded from the analysis. The definition of opioid-exposed status was based on the following question in the G1 section of the SSADDA: “Have you ever used any of the following opiate drugs: Heroin, Morphine, Methadone, Fentanyl or P-dope, Codeine, Percodan, Darvon, Dilaudid, Demerol, Percocet, Opium, and Other Opiate?”. All subjects – opioid dependent, exposed, etc. – were ascertained with questions in the G1 section of the full SSADDA interview, in which information regarding illegal and prescription opioid abuse was collected. Severity of OpOD was scored on a 4-point scale as follows: 1: used opioids but no overdose history; 2: had a history of overdose but no consequent medical treatment; 3: overdosed and received medical treatment <3 times; 4: overdosed and received medical treatment ≥ 3 times. A total of 1,019 participants (653 EAs, 366 AAs) had an OpOD severity score ≥ 2, meaning that they all endorsed a history of overdose history (Table 1). Thus, the average overdose rate in opioid-exposed subjects was 29.3% (25.3% and 32.1% in AAs and EAs, respectively).

Genome-wide Association Study (GWAS) Analyses

Yale-Penn 1 samples were genotyped on the Illumina (San Diego, CA, USA) HumanOmni1-Quad v1.0 microarray (total SNPs = 1,140,419) and Yale-Penn 2 samples were genotyped on the Illumina HumanCore Exome array (total SNPs = 550,601). We used PLINK1.9 (Chang et al., 2015) for quality control (QC) in each cohort, applying the following criteria: (1) individual genotype missing rate <2%, (2) SNP genotype missing rate <2%, (3) Hardy-Weinberg $P > 1 \times 10^{-6}$, and (4) minor allele frequency (MAF) >3%. Samples from Yale-Penn 1 and 2 that passed the initial QC were subjected to ancestry analysis by comparison with the 1000 Genomes Project phase 1 reference panel (1000 Genomes Project Consortium et al., 2010). Principal component analysis was performed with Eigensoft (Price et al., 2006) and the resulting first 10 principal components (PCs) served to differentiate AAs and EAs (Sherva et al., 2016). Subsequently, AAs and EAs from each Yale-Penn cohort were subjected to SNP imputation by using Minimac3 implemented in the Michigan Imputation Server (Das et al, 2016) with the 1000 Genomes Project phase 3 reference panel. Dosage data were transformed into best-guess genotypes using PLINK1.9, retaining high-quality data by filtering imputed data with genotype imputation probability (GP) ≥ 0.8, individual genotyping missing rate <5%, SNP MAF >3%, SNP missing call rate <5%, and Hardy-Weinberg $P > 1 \times 10^{-6}$.

We conducted GWAS on the severity score of the OpOD trait in AAs and EAs separately, with adjustment for the relatedness among participants and included age, sex, and the first 10 PCs (to account for population structure using Genome-wide Efficient Mixed Model Association (GEMMA) software (Zhou and Stephens, 2012)) as covariates. Our initial data QC included a procedure to address population stratification, as for our previous studies (Sherva et al., 2016; Zhou et al., 2017). GEMMA can also account for relatedness among samples and can control for population stratification and other confounding factors (Kang et al., 2010; Zhou and Stephens, 2012). We chose a univariate linear mixed model to associate genotype with overdose severity score, the rationale for which is that the overdose severity score represents the continuous nature of opioid overdose severity. Subsequent meta-analysis was carried out using the inverse variance method in PLINK 1.9. The number of SNPs by cohort that passed the QC and were included in the analyses were 10,295,719 in Yale-Penn 1 AA (n=1,020), 6,981,522 in Yale-Penn 1 EA (n=1,181), 7,501,713 in Yale-Penn 2 AA (n=425), and 7,501,713 in Yale-Penn 2 EA (n=851). The significance threshold for GWAS analyses was set at $P = 5 \times 10^{-8}$.

Post-GWAS Prioritization of Possible OpOD Risk Genes with Network Analysis and Drug Repositioning Analysis

Variants close to *MCOLN1* and *PNPLA6* were genome-wide significant (GWS) in AAs; see below for detailed GWAS results. The three GWS SNPs map closely within a region [-1kb, 5kb] relative to the transcription start site (TSS) of *MCOLN1*. They are also ~10 kb distant from the TSS of *PNPLA6*. To evaluate whether *MCOLN1* or *PNPLA6* was more likely to be the risk gene for OpOD (beyond that of the somewhat greater proximity of *MCOLN1* to the GWS markers), we conducted network analysis by querying *MCOLN1* and *PNPLA6* individually and jointly with GeneMANIA (Franz et al., 2018), which provides functional association data taking account of similarities in protein and genetic interactions, pathways, co-expression, co-localization, and protein domains.

Drug repositioning analysis was then carried out for *MCOLN1* with the Connectivity Map (CMap) Database (Subramanian et al., 2017); data for *PNPLA6* were not available in CMap. CMap contains >1 million transcriptional profiles derived from perturbations of multiple cancer-related cell lines that arise from a genetic mutation affecting protein function, genetic perturbation (knockdown or overexpression of a gene), or treatment with a small molecule or a drug. CMap used 8 human cancer-related cell lines: A375 (malignant melanoma), A549 (non-small cell lung carcinoma), HCC515 (non-small cell lung adenocarcinoma), HEPG2 (hepatocellular carcinoma), HT29 (colorectal adenocarcinoma), MCF7 (breast adenocarcinoma), PC3 (prostate adenocarcinoma), and VCAP (metastatic prostate cancer), as well as HA1E (kidney epithelial immortalized non-cancer cell line). Transcriptional profiles with high similarity between two proteins operating in the same pathway, between a drug and its protein target, or between two drugs of similar function but with structural dissimilarity, have the potential to reveal useful and previously unrecognized connections. CMap has provided expression similarity scores for a specific expression profile (here, we use the *MCOLN1* overexpression-induced expression profile) with other drug-induced transcriptional profiles, including consensus transcriptional signatures of 83 drug classes, i.e., transcriptional profiles induced by 2,837 drugs grouped into 83 drug

classes. Expression similarity is evaluated by means of scores that vary from –100 to 100, with –100 the most extreme opposite expression profile and 100 the most extreme similar expression profile. To accomplish this, we downloaded expression similarity scores for *MCOLNI* by selecting the option ‘See all connections’ after querying ‘*MCOLNI*’ directly in the CMap cloud server (<https://clue.io/command?q=/conn>). CMap includes 32 drugs that target mu-, delta-, and kappa-opioid receptors. Only one opioid drug class comprising three opioid receptor agonists is included in the summary statistics of 83 drug classes provided by CMap for *MCOLNI* (described below, “Results”). Accordingly, we also downloaded the individual expression similarity scores for all 2,837 drugs to *MCOLNI* over-expression, and then evaluated expression similarity scores between *MCOLNI* and 29 other opioid-related drugs that were not included in the 83 drug classes. Other drugs with transcriptional signatures opposite that of *MCOLNI* over-expression (negative similarity scores to *MCOLNI*) were also identified.

RESULTS

For the two Yale-Penn cohorts, we analyzed EAs and AAs separately, then meta-analyzed the two cohorts by population (Yale-Penn 1 and 2 EAs, and Yale-Penn 1 and 2 AAs). Then, a trans-ancestry meta-analysis was performed that included all of the EA and AA samples. We identified two GWS signals in AAs, which mapped to the genes *MCOLNI* (also close to *PNPLA6*) and *DDX18* (Figure 1–2 and Figure S1–S2). There were no GWS signals identified in EAs. In the meta-analysis of all subjects, only the signal that mapped to *MCOLNI*, close to *PNPLA6* on chromosome 19, represented by three SNPs (rs115208233, rs116181528, and rs114077267), were GWS (Figure 1). The other SNP that was GWS in AAs, rs369098800, which is ~470 kb from *DDX18*, was only nominally significant in the trans-ancestral meta-analysis (Figures 1 and S1). We found no evidence of inflation ($\lambda \sim 1.0$, Figure 1) both in the meta-analyses of AAs and EAs separately and together.

Of the three AA-specific GWS SNPs, the first two are located in different intronic regions of *MCOLNI* and the last resides in the *MCOLNI* promoter region (477bp upstream of the *MCOLNI*; Figure 2 and 3). The top SNP, rs115208233, is in high linkage disequilibrium (LD) with the promoter SNP rs114077267 and in moderate LD with rs116181528 (rs115208233:rs114077267 $R^2=0.9$; rs115208233:rs116181528 $R^2=0.74$; rs116181528:rs114077267 $R^2=0.67$), which are polymorphic only in AAs, with the minor G allele positively associated with the OpOD severity score (see Figure S3). The MAFs for these SNPs range from 0.04 to 0.06 (Table 2) in that population. The association signals were nominally significant in both Yale-Penn 1 and Yale-Penn 2 AA samples separately (all P values $< 5.0 \times 10^{-4}$, Table 2). The minor alleles are all associated with higher overdose severity scores in AAs.

There were no GWS SNPs in both populations, though one SNP, rs4146322, located in the intronic region of *C4orf22*, showed suggestive evidence of association with OpOD severity in the meta-analysis of AAs and EAs (Figure S2). According to the brain expression quantitative trait locus (eQTL) database of European samples, BRAINEAC (Ramasamy et al., 2014), rs4146322 is an eQTL for both *C4orf22* ($P=7.1 \times 10^{-3}$) and *ANTXR2* ($P=8.9 \times 10^{-3}$) in the putamen and frontal cortex, respectively. In another eQTL database,

GTE_x (GTE_x-Consortium, 2015), rs4146322 is an eQTL for *C4orf22* in testis ($P=7.0\times 10^{-11}$). Another SNP, rs369098800, (which maps near *DDX18*), was GWS in AAs but only nominally significant in the trans-ancestry GWAS of EAs+AAs, but it is not available in BRAINEAC or GTE_x databases.

SNP rs73568641, ~300kb upstream of *OPRM1*, associated with therapeutic opioid dose in AAs (Smith et al., 2017), was not significant in the OpOD analysis in the AA part of the sample ($P=0.54$). Two other SNPs associated with opioid dependence in AAs (Gelernter et al., 2014), rs62103177 (mapped to *KCNQ2*) and rs60349741 (mapped to *KCNKI*), were also not significantly associated with OpOD ($P=0.80$ and 0.28 , respectively). Similarly, rs12442183 close to *RGMA* and associated with OD in EAs (Cheng et al., 2018), was not associated with OpOD in EAs ($P=0.56$).

Based on functional prediction from Haploreg4 (Ward and Kellis, 2016) and as described above, the three OpOD GWS SNPs reside in the predicted enhancer regions, and rs114077267 is also in the promoter region of *MCOLN1*. Also, according to annotation data from Haploreg4 (Ward and Kellis, 2016), this promoter region is active in 23 tissues, including brain, blood, muscle, and heart. The risk allele, rs114077267*T, is predicted to generate motifs with high binding affinity for the transcription factors GFI1 (transcription repressor), GATA (transcription activator), and HOXA7 (transcription repressor). Also, as noted above, these three GWS SNPs are about 10 kb from another gene, *PNPLA6*. Per the UCSC Genome Browser (Kent et al., 2002), the 3' region of *MCOLN1* is only 144 bp upstream of the TSS of *PNPLA6* (Figure 3), suggesting that common genetic elements could regulate the expression of both loci.

The potential involvement of *MCOLN1*, *PNPLA6*, or both of these loci in OpOD was evaluated further by performing network analysis for biologically relevant genes by querying *MCOLN1* and *PNPLA6* jointly (Figure 4a) or separately (Figure 4b and 4c) in the GeneMANIA database (Franz et al., 2018). This showed that the expression of *MCOLN1* is correlated with that of *PNPLA6*. Further network analysis with GeneMANIA revealed that MCOLN1, the protein encoded by *MCOLN1*, displays protein-protein interaction with MCOLN3 (encoded by *MCOLN3*, a paralog of *MCOLN1*) and SLC35E1 (a putative transporter). PNPLA6, the protein product of *PNPLA6*, demonstrates a very different protein-protein interaction pattern from that of MCOLN1 (Figures 4a–4c); proteins that interact with PNPLA6 do not have strong associations with MCOLN1. In addition, in functional enrichment analysis, most genes associated with MCOLN1 in network analysis have similar protein domains and are involved in calcium and other metal ion and cation channel activities (Figure 4d). In contrast, no functional categories were enriched for *PNPLA6*-associated genes.

The possible involvement of *MCOLN1* in OpOD was further evaluated by performing a drug repositioning analysis with CMap (see Methods). The consensus expression profile of the drug class “opioid receptor agonists” provided the transcriptional signature most similar to that of an *MCOLN1* over-expression profile (Figure 5a, expression similarity score= 99). The opioid receptor agonist class is comprised of FIT (fentanyl isothiocyanate, an irreversible opioid receptor agonist targeting the delta opioid receptor), Ioperamide (a

peripheral opioid receptor agonist affecting multiple intestinal opioid receptors), and ICI-199441 (a kappa opioid receptor agonist), and their expression similarity scores relative to *MCOLN1* over-expression are 99, 95 and 88, respectively (Figure 5b). Other than these three opioidergic drugs allocated to drug classes in CMap, only 3 of the 32 opioidergic drugs included are allocated to a group. The other four top drug classes ranked by their expression similarity scores (>96) to an *MCOLN1* over-expression induced profile include tricyclic antidepressants, serotonin receptor antagonists, dopamine receptor antagonists, and calcium receptor inhibitors. In addition, as the opioid receptor antagonist naloxone can reverse OpOD, we specifically evaluated a possible relationship between *MCOLN1* over-expression and the naloxone-stimulated transcriptome signature, including a set of overexpressed and downregulated genes. Transcriptional profiles with high similarity between what is observed in the case of *MCOLN1* over-expression and what is observed with opioidergic drugs might represent useful and previously unrecognized connections; the *MCOLN1* induced expression profile yielded an expression similarity score of 87 to that of naloxone, and naloxone was ranked sixth among 32 opioid drugs based on their similarity to transcriptional signatures of *MCOLN1* over-expression. Overall, the average similarity score to that of *MCOLN1* over-expression for all 32 available opioid drugs was 62 ± 27 . Other opioid drugs that target mu-, kappa-, and delta-opioid receptors also demonstrated high similarity for perturbed transcriptional signatures with that of *MCOLN1* over-expression (all expression similarity scores >90), including the three opioid receptor agonists, and naltrindole and naltriben, which are opioid receptor antagonists. That is, a total of 5 opioid-related drugs shared similarities in the drug-induced expression pattern to *MCOLN1*, each scoring >90. Furthermore, we identified BRD-A80383043, a glutamate receptor agonist, as the only drug that can induce a transcriptional profile opposite that of an *MCOLN1* over-expression stimulated profile (expression similarity score = -94).

DISCUSSION

We conducted a GWAS for OpOD among AAs and EAs, and found a GWS signal specific to AAs, which spans a total of only 7,231 basepairs, mapped to *MCOLN1* and ~10 kb from *PNPLA6* (Figure 3). Among the three GWS SNPs, rs114077267 is potentially functional, residing in the promoter region of *MCOLN1*, suggesting that *MCOLN1* is a more likely risk gene for OpOD than *PNPLA6*. This conclusion is further supported by a network analysis and a drug repositioning analysis, in which (1) genes correlated with *MCOLN1* but not *PNPLA6* were enriched in OpOD-related functional categories, including regulation by Ca²⁺, and (2) transcriptional signatures induced by opioid receptor agonists were very similar to the *MCOLN1* over-expression stimulated profile. Since the pattern of *MCOLN1* overexpression is similar to that of opioid drug induced expression, we infer that they have related biology.

The association of *MCOLN1* with OpOD could be related to its connection to opioid receptor signaling. Previous research showed that MCOLN1 (Mucolipin-1) is a receptor protein with potential cationic channel activities (Bassi et al., 2000; Montell, 2001). Loss-of-function mutations in *MCOLN1* cause mucopolipidosis type IV, a severe lysosomal storage disorder (Bassi et al., 2000). MCOLN1 has also been reported to be an endolysosomal iron release channel (Dong et al., 2008), and knockdown of *MCOLN1* in human retinal

pigmented epithelial 1 cells promotes the production of reactive oxygen species by Fe²⁺ trapped in lysosomes (Coblentz et al., 2014). In our study, network analysis demonstrated that *MCOLNI*-associated genes were enriched in cation and metal ion channel activities (Figure 4d). Furthermore, in a drug repositioning analysis, the *MCOLNI* overexpression-induced transcription signature was highly similar to that of three opioid receptor agonists' induced transcriptional profiles (Figure 5). These three opioid receptor agonists – the only ones included in the defined CMap drug classes – are FIT (fentanyl isothiocyanate, irreversible delta opioid receptor agonist), Ioperamide (peripheral opioid receptor agonist), and ICI-199441 (kappa opioid receptor agonist). We also found that three other opioid receptor antagonists not included in CMap classes – naloxone, naltrindole, and naltriben – induced transcriptional profiles similar to an *MCOLNI* over-expression induced transcriptional signature, which suggests that both opioid receptor antagonists and agonists are functionally related to *MCOLNI*. Further, a glutamate receptor agonist, BRD-A80383043, displayed a transcriptional signature opposite to that of the *MCOLNI* over-expression induced profile. Glutamate is the major excitatory neurotransmitter in the brain and plays an integral role in opioid addiction (Peters and De Vries, 2012) and OpOD (White and Irvine, 1999). As previously shown, the glutamate receptor antagonists JNJ16259685 and LY341495 increase the antinociceptive efficacy of opioids (Fischer et al., 2008). We suspect that *MCOLNI* overexpression may be associated with OpOD severity, the underlying mechanism being, potentially, its ability to enhance the antinociceptive efficacy of opioids. This is supported by drug repositioning analyses that revealed similarities between the drug-induced expression pattern of 5 opioid-related drugs and *MCOLNI* overexpression, while a glutamate receptor agonist drug demonstrated an expression profile opposite to *MCOLNI* overexpression.

Additionally, in considering the context of our OpOD findings, we looked up in the OpOD GWAS data SNPs that were previously associated with opioid dosing in AAs (Smith et al., 2017) and with opioid dependence in AAs (Gelernter et al., 2014) and EAs (Cheng et al., 2018), and did not find even nominal association in either case. The OpOD trait is different biologically from both opioid dosing and opioid dependence *per se*; these are not “failures to replicate.”

Our finding requires independent validation. Because one of our top SNPs, rs114077267, from the AA OpOD GWAS, maps to the promoter region of *MCOLNI* and is AA specific, we checked the extent of promoter SNPs (MAF \geq 0.03) at *MCOLNI* and compared their allele frequencies among different populations, including African (AFR), Admixed American (AMR), Asian (ASN) and European (EUR) populations (Figure S4). Compared to the other three populations, the AFR population has more SNPs (MAF $>$ 0.03) in the promoter region of *MCOLNI* and, thus there is a greater potential to detect true associations. Among EUR and ASN, there are few SNPs in the *MCOLNI* promoter. Thus, another AA population would be required for replication, although it may be possible to identify consistent evidence (with other SNPs mapped to the same region) in other populations.

Our study has potential limitations. *MCOLNI* may be associated with quantity or frequency of exposure to opioids. However, we do not have accurate measures of quantity and

frequency of opioid exposure, so we cannot test this directly. Potential polygenic risk score (PRS) analyses between OpOD and other independent psychiatric and SUD traits should be carried out in the future to identify pleiotropy, but at this point our sample is not adequately powered for this step. This is a critical problem for the research field: namely, that recruitment of more subjects to increase power is absolutely essential.

Our OpOD GWAS provides statistically significant support for *MCOLNI* as a risk gene for OpOD, with the hypothesis that *MCOLNI* is the effect locus supported by both network analysis and drug repositioning analysis. An *MCOLNI* overexpression-induced profile was highly similar to opioid receptor agonist-induced expression profiles. The apparent involvement of *MCOLNI* in the activity of the opioidergic system warrants evaluation in cellular and animal models. In conclusion, our GWAS supports *MCOLNI* as a risk gene for OpOD. Considering all of the data, the effect of *MCOLNI* over-expression on the opioid receptor system could be a mechanism for its association with OpOD. In any case, the findings suggest a novel biology underlying OpOD.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This study was supported by National Institutes of Health grants R01 DA12690, R01 AA11330, R01 AA017535, and the VA Connecticut Healthcare Center and Philadelphia VA MIRECCs. Genotyping services for a part of our GWAS were provided by the Center for Inherited Disease Research and Yale University (Center for Genome Analysis), which is fully funded by Federal contract N01-HG-65403 from the NIH to The Johns Hopkins University. Ann Marie Lacobelle, MS and Christa Robinson provided technical support.

REFERENCES

- 1000 Genomes Project Consortium, Abecasis GR, Altshuler, Auton A, Brooks LD, Durbin RM, Gibbs RA, Hurles ME, McVean GA (2010) A map of human genome variation from population-scale sequencing. *Nature* 467:1061–1073. [PubMed: 20981092]
- Bassi MT, Manzoni M, Monti E, Pizzo MT, Ballabio A, Borsani G (2000) Cloning of the gene encoding a novel integral membrane protein, mucopolipidin and identification of the two major founder mutations causing mucopolipidosis type IV. *Am J Hum Genet* 67:1110–1120. [PubMed: 11013137]
- Boyer EW (2012) Management of opioid analgesic overdose. *N Engl J Med* 367:146–155. [PubMed: 22784117]
- Buchanich JM, Balmert LC, Williams KE, Burke DS (2018) The effect of incomplete death certificates on estimates of unintentional opioid-related overdose deaths in the United States, 1999–2015. *Public Health Rep* 133:423–431. [PubMed: 29945473]
- Centers for Disease Control and Prevention (2018) Provisional Counts of Drug Overdose Death. National-Center-for-Health-Statistics.
- Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 4:7. [PubMed: 25722852]
- Cheng Z, Zhou H, Sherva R, Farrer LA, Kranzler HR, Gelernter J (2018) Genome-wide association study identifies a regulatory variant of *RGMA* associated with opioid dependence in European Americans. *Biological psychiatry* 84:762–770. [PubMed: 29478698]
- Chou R, Korthuis PT, McCarty D, Coffin PO, Griffin JC, Davis-O'Reilly C, Grusing S, Daya M (2017) Management of suspected opioid overdose with naloxone in out-of-hospital settings: a systematic review. *Ann Intern Med* 167:867–875. [PubMed: 29181532]

- Coblentz J, St Croix C, Kiselyov K (2014) Loss of TRPML1 promotes production of reactive oxygen species: is oxidative damage a factor in mucopolipidosis type IV? *Biochem J* 457:361–368. [PubMed: 24192042]
- Dong XP, Cheng X, Mills E, Delling M, Wang F, Kurz T, Xu H (2008) The type IV mucopolipidosis-associated protein TRPML1 is an endolysosomal iron release channel. *Nature* 455:992–996. [PubMed: 18794901]
- Encode_Project_Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* 489:57–74. [PubMed: 22955616]
- Franz M, Rodriguez H, Lopes C, Zuberi K, Montojo J, Bader GD, Morris Q (2018) GeneMANIA update 2018. *Nucleic Acids Res* 46:W60–W64. [PubMed: 29912392]
- Gelernter J, Kranzler HR, Sherva R, Koesterer R, Almasy L, Zhao H, Farrer LA (2014) Genome-wide association study of opioid dependence: multiple associations mapped to calcium and potassium pathways. *Biological psychiatry* 76:66–74. [PubMed: 24143882]
- GTEX-Consortium (2015) Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* 348:648–660. [PubMed: 25954001]
- Jalal H, Buchanich JM, Roberts MS, Balmert LC, Zhang K, Burke DS (2018) Changing dynamics of the drug overdose epidemic in the United States from 1979 through 2016. *Science* 361.
- Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, Sabatti C, Eskin E (2010) Variance component model to account for sample structure in genome-wide association studies. *Nat Genet* 42:348–354. [PubMed: 20208533]
- Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D (2002) The human genome browser at UCSC. *Genome Research* 12:996–1006. [PubMed: 12045153]
- Kharasch ED, Regina KJ, Blood J, Friedel C (2015) Methadone pharmacogenetics: CYP2B6 polymorphisms determine plasma concentrations, clearance, and metabolism. *Anesthesiology* 123:1142–1153. [PubMed: 26389554]
- Montell C (2001) Physiology, phylogeny, and functions of the TRP superfamily of cation channels. *Science's STKE : signal transduction knowledge environment* 2001:re1.
- Nadpara PA, Joyce AR, Murrelle EL, Carroll NW, Carroll NV, Barnard M, Zedler BK (2018) Risk factors for serious prescription opioid-induced respiratory depression or overdose: comparison of commercially insured and veterans health affairs populations. *Pain medicine (Malden, Mass)* 19:79–96.
- Nelson EC, Agrawal A, Heath AC, Bogdan R, Sherva R, Zhang B, Al-Hasani R, Bruchas MR, Chou YL, Demers CH, Carey CE, Conley ED, Fakira AK, Farrer LA, Goate A, Gordon S, Henders AK, Hesselbrock V, Kapoor M, Lynskey MT, Madden PA, Moron JA, Rice JP, Saccone NL, Schwab SG, Shand FL, Todorov AA, Wallace L, Wang T, Wray NR, Zhou X, Degenhardt L, Martin NG, Hariri AR, Kranzler HR, Gelernter J, Bierut LJ, Clark DJ, Montgomery GW (2016) Evidence of CNH3 involvement in opioid dependence. *Mol Psychiatry* 21:608–614. [PubMed: 26239289]
- Peters J, De Vries TJ (2012) Glutamate mechanisms underlying opiate memories. *Cold Spring Harb Perspect Med* 2:a012088. [PubMed: 22951449]
- Pierucci-Lagha A, Gelernter J, Chan G, Arias A, Cubells JF, Farrer L, Kranzler HR (2007) Reliability of DSM-IV diagnostic criteria using the semi-structured assessment for drug dependence and alcoholism (SSADDA). *Drug Alcohol Depen* 91:85–90.
- Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38:904–909. [PubMed: 16862161]
- Ramasamy A, Trabzuni D, Guelfi S, Varghese V, Smith C, Walker R, De T, Coin L, de Silva R, Cookson MR, Singleton AB, Hardy J, Ryten M, Weale ME, Consortium UBE, Consor NABE (2014) Genetic variability in the regulation of gene expression in ten regions of the human brain. *Nat Neurosci* 17:1418–1428. [PubMed: 25174004]
- Sherva R, Wang Q, Kranzler H, Zhao H, Koesterer R, Herman A, Farrer LA, Gelernter J (2016) Genome-wide association study of cannabis dependence severity, novel risk variants, and shared genetic risks. *JAMA Psychiatry* 73:472–480. [PubMed: 27028160]

- Smith AH, Jensen KP, Li J, Nunez Y, Farrer LA, Hakonarson H, Cook-Sather SD, Kranzler HR, Gelernter J (2017) Genome-wide association study of therapeutic opioid dosing identifies a novel locus upstream of OPRM1. *Mol Psychiatry* 22:346–352. [PubMed: 28115739]
- Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, Gould J, Davis JF, Tubelli AA, Asiedu JK, Lahr DL, Hirschman JE, Liu Z, Donahue M, Julian B, Khan M, Wadden D, Smith IC, Lam D, Liberzon A, Toder C, Bagul M, Orzechowski M, Enache OM, Piccioni F, Johnson SA, Lyons NJ, Berger AH, Shamji AF, Brooks AN, Vrcic A, Flynn C, Rosains J, Takeda DY, Hu R, Davison D, Lamb J, Ardlie K, Hogstrom L, Greenside P, Gray NS, Clemons PA, Silver S, Wu X, Zhao WN, Read-Button W, Wu X, Haggarty SJ, Ronco LV, Boehm JS, Schreiber SL, Doench JG, Bittker JA, Root DE, Wong B, Golub TR (2017) A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell* 171:1437–1452 e1417. [PubMed: 29195078]
- Ward LD, Kellis M (2016) HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res* 44:D877–881. [PubMed: 26657631]
- Webster LR (2017) Risk factors for opioid-use disorder and overdose. *Anesth Analg* 125:1741–1748. [PubMed: 29049118]
- White JM, Irvine RJ (1999) Mechanisms of fatal opioid overdose. *Addiction* 94:961–972. [PubMed: 10707430]
- Zedler B, Xie L, Wang L, Joyce A, Vick C, Brigham J, Kariburyo F, Baser O, Murrelle L (2015) Development of a risk index for serious prescription opioid-induced respiratory depression or overdose in veterans' health administration patients. *Pain medicine (Malden, Mass)* 16:1566–1579.
- Zhou H, Polimanti R, Yang BZ, Wang Q, Han S, Sherva R, Nunez YZ, Zhao H, Farrer LA, Kranzler HR, Gelernter J (2017) Genetic risk variants associated with comorbid alcohol dependence and major depression. *JAMA Psychiatry* 74:1234–1241. [PubMed: 29071344]
- Zhou X, Stephens M (2012) Genome-wide efficient mixed-model analysis for association studies. *Nat Genet* 44:821–824. [PubMed: 22706312]

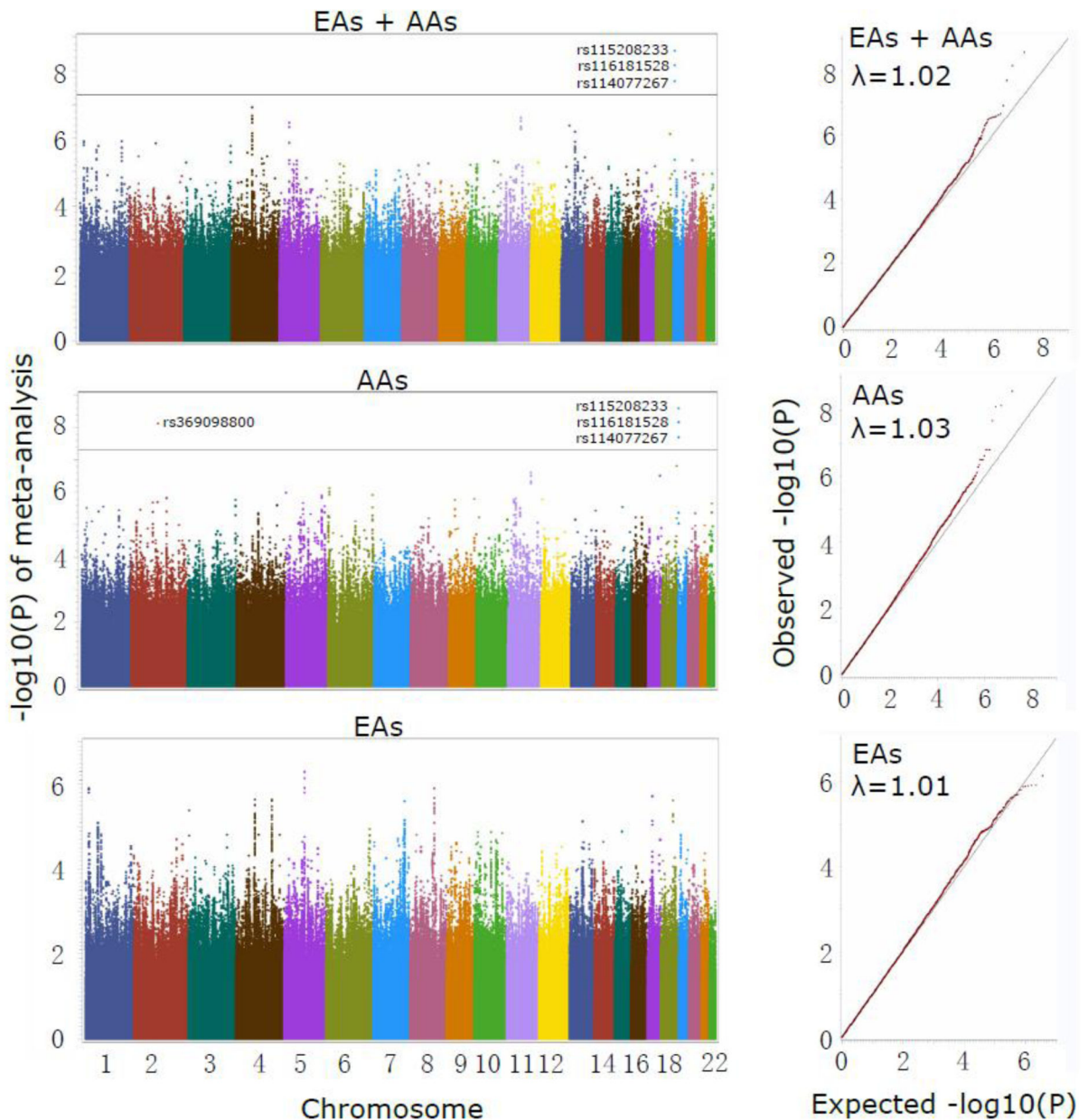


Figure 1. Genome-wide association studies of opioid-overdose severity among four cohorts, including Yale-Penn1 and Yale-Penn2, each of which included both African Americans (AAs) and European Americans (EAs). The three panels display the Manhattan plots (left) and qq plots (right) for the meta-analysis of both population groups, and the AA and EA cohorts separately. Four SNPs reached genome-wide significance; three SNPs (rs115208233, rs116181528 and rs114077267) on chromosome 19 are specific to the AA population and map to *MCOLN1*, close to *PNPLA6*. Another SNP (rs369098800) is on chromosome 2, ~471kb away from *DDX18*. On chromosome 4, a suggestive significant association peak close to gene *C4orf22* was seen in the meta-analysis of EA and AA.

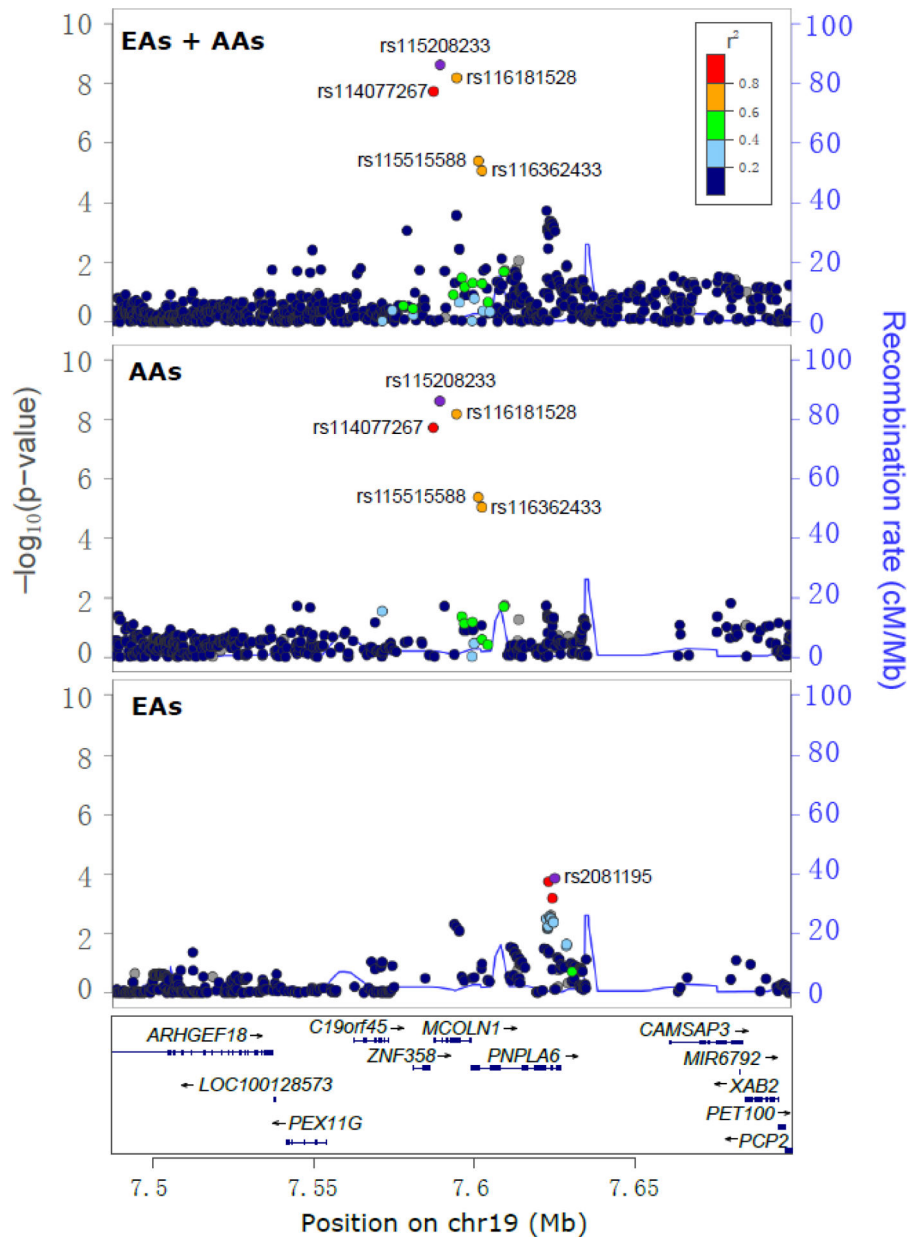


Figure 2.

Regional Manhattan plots displaying one independent signal associated with opiod overdose (OpOD) severity on the locus *MCOLN1*. Three panels illustrate the association signals that emerged from the meta-analysis of all European Americans (EAs) and African Americans (AAs) and in the meta-analysis of EAs and AAs separately. Three genome-wide significant SNPs specific to the AA population map to *MCOLN1*: rs115208233, rs116181528 and rs114077267. Among these, rs115208233 is the lead SNP and rs114077267 is in the promoter region of *MCOLN1*. Two other AA-specific SNPs, rs116362433 (an intronic SNP of *MCOLN1*) and rs115515588 (a promoter SNP of *PNPLA6*), were nominally associated with OpOD severity (all P values $\sim 5 \times 10^{-6}$). One SNP, rs2081195, located in the intronic region of *PNPLA6*, was nominally significantly

associated with OpOD severity in EAs ($P=1.4\times 10^{-4}$). Note: the SNPs are colored to reflect the linkage disequilibrium (R^2) with the lead SNP (rs115208233 for EAs and EAs+AAs; rs2081195 for EAs) based on the African (for AAs and EAs+AAs) or European (for EAs) population data from the 1000 Genomes Project. The light blue line and right Y-axis indicate the observed recombination rate estimated from HapMap samples. Only association signals existing in 2 cohorts are plotted.

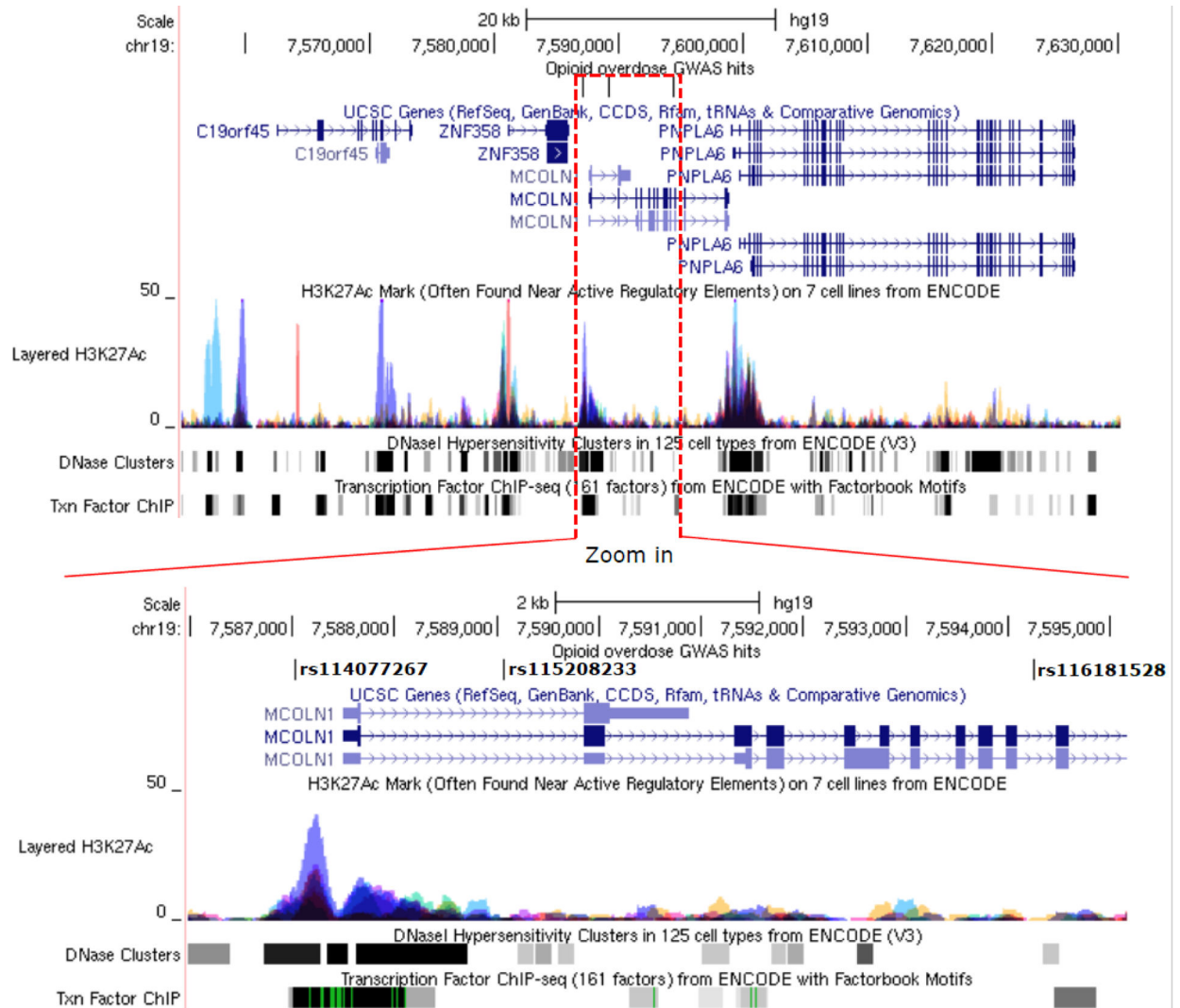


Figure 3. Genomic and functional annotation for rs114077267, rs115208233, and rs116181528. The upper panel denotes the chromosomal region of the three genome-wide significant (GWS) SNPs. In addition, three regulatory features, including H3K27Ac marks, DNase clusters, and transcription factor ChIP-seq binding sites, are seen in the region. A subset of the region covered by the red dashed line is magnified in the lower panel, with the GWS SNPs in bold. The signals for the H3K27Ac mark, DNase I hypersensitivity cluster, and transcriptional factor binding site are annotated according to experimental data from the ENCODE Consortium (Encode_Project_Consortium, 2012). The height and color in the H3K27Ac marks represent signal intensity and data source from different cell lines. The gray box indicates the extent of the hypersensitive region or cluster of transcriptional factor occupancy. The shade is proportional to the maximum signal strength observed in any cell line. The green line indicates the highest scoring site of an identified canonical motif for the corresponding factor.

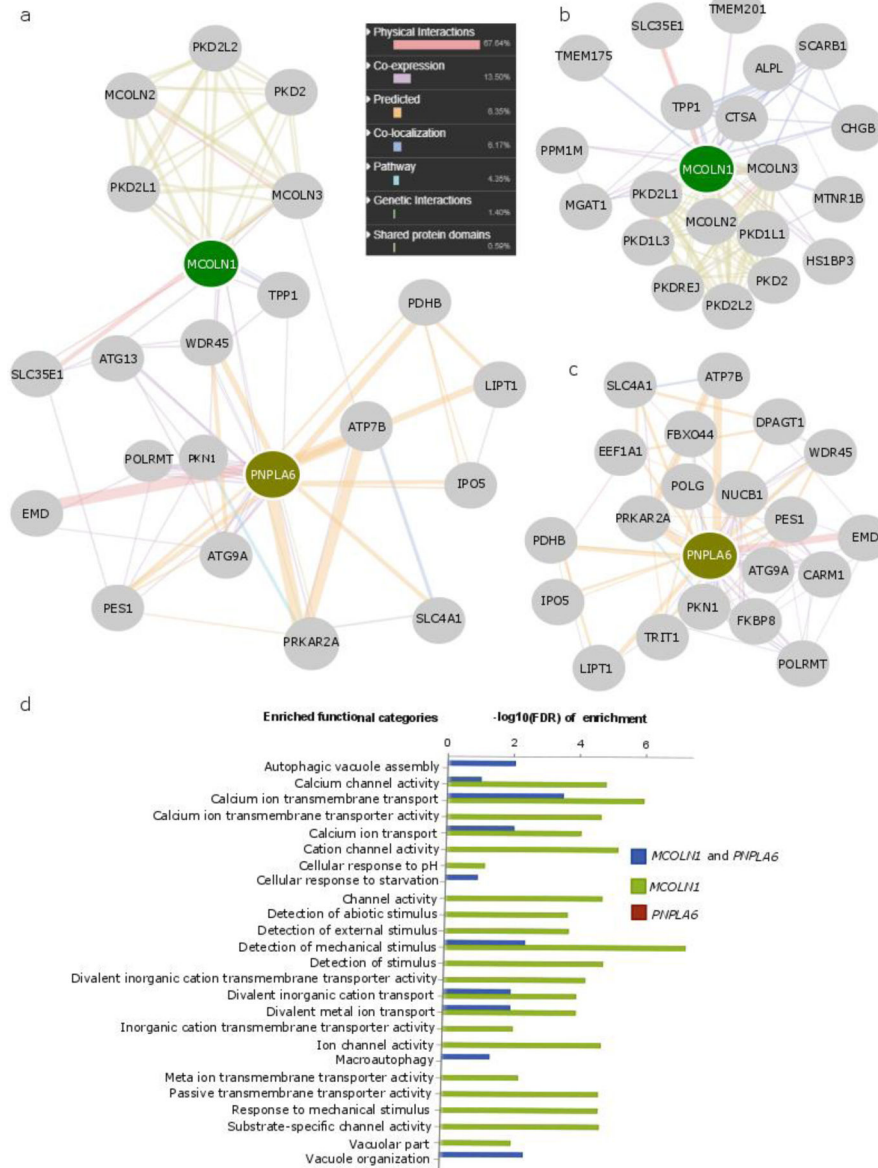


Figure 4. Network analysis prioritizes *MCOLN1* but not *PNPLA6* as a potential opioid overdose risk gene. GeneMANIA was applied to obtain biologically related genes that are associated with *MCOLN1* and *PNPLA6* when queried jointly (a) or individually (b and c), the enriched functional categories of which were also illustrated (d). The above networks display different types of association between *MCOLN1* or *PNPLA6* and their biologically related genes, in which the color and size of each line linking two genes indicates the type and strength of the association. These associations comprise protein and genetic interactions, pathways, co-expression, co-localization, and protein domain similarity. The enrichment *P* values ($-\log_{10}(\text{FDR})$) provided by GeneMANIA are demonstrated for different functional categories. Most enriched functional categories are observed among *MCOLN1*-associated

genes, and no functional categories are enriched in *PNPLA6*-associated genes. (FDR: false discovery rate)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

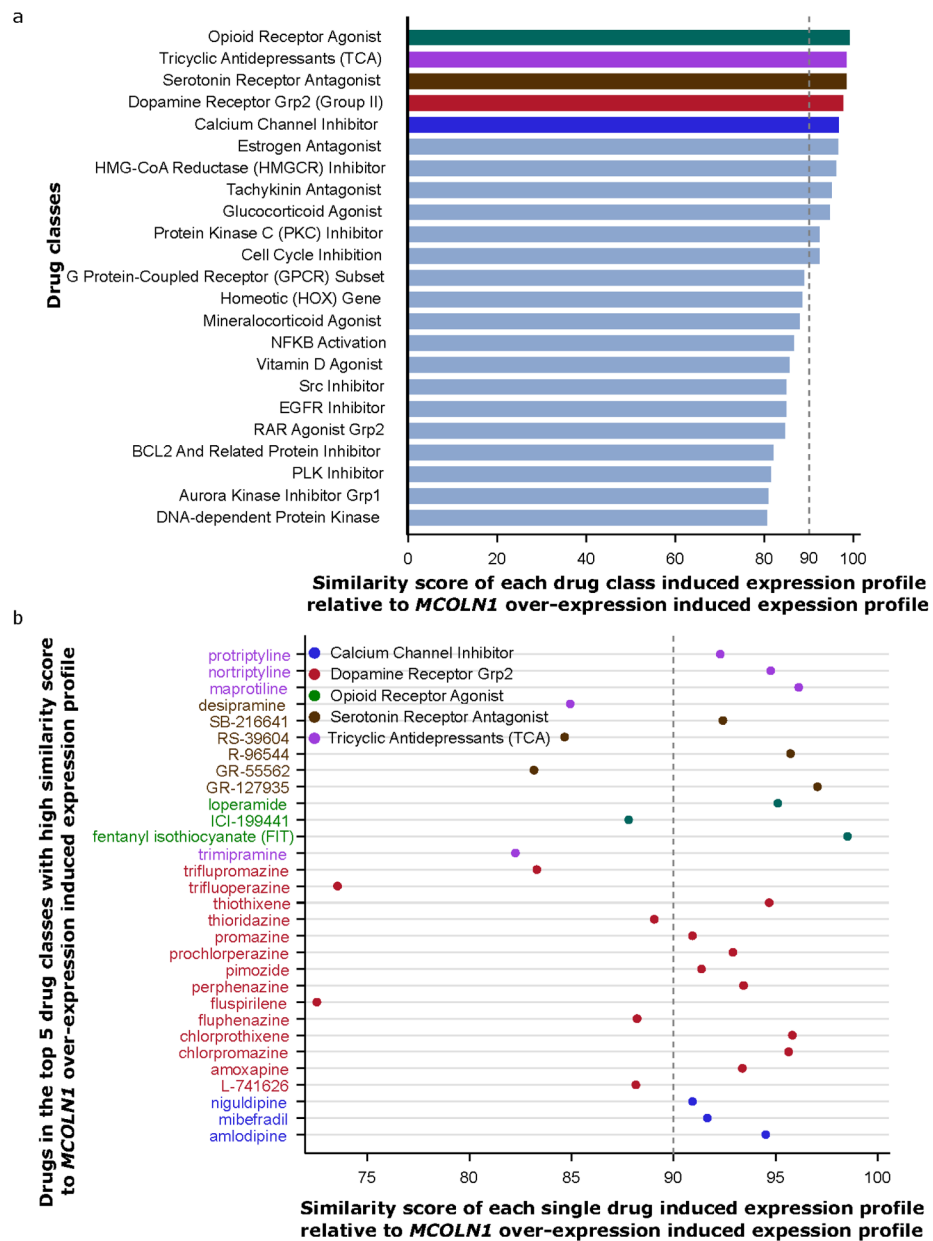


Figure 5. Drug repositioning analysis supports the involvement of *MCOLN1* in the opioid receptor system. (a) Among 83 drug classes, opioid receptor agonist-induced expression profiles display the highest expression similarity score with the *MCOLN1* over-expression induced profile in CMap. Expression similarity scores are generated by comparing the *MCOLN1* over-expression induced expression profile to the signatures of 83 classes of drug-induced transcription. (b) Single drugs from these top 5 drug classes (highlighted in part a) with expression similarity score ≥ 80 are shown.

Table 1.

Sample counts and characteristics for the severity of the opioid overdose trait.

Opioid Overdose Coding	African American					European American					Total N
	YP1 N	YP2 N	N	Mean Age	YP1+YP2 Male (%)	YP1 N	YP2 N	N	Mean Age	YP1+YP2 Male (%)	
1	741	338	1079	41.9	0.71	777	602	1379	36.2	0.67	2458
2	116	35	151	46.1	0.83	125	75	200	37.1	0.69	351
3	97	29	126	43.7	0.93	191	114	305	37.2	0.68	431
4	66	23	89	45.2	0.83	88	60	148	37.2	0.72	237
Overdose	279	87	366	45	0.86	404	249	653	37.2	0.69	1019
Subtotal	1020	425	1445	42.7	0.74	1181	851	2032	36.5	0.68	3477

Note: YP1 and YP2 represent the cohorts, Yale-Penn 1 and Yale-Penn 2, respectively. Mean age was calculated for the combined sample of the Yale-Penn 1 and Yale-Penn 2 cohorts.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

Characteristics of the three genome-wide significant SNPs located on *MCOLN1*.

Chr	Position	SNP	A1	A2	MAF	Beta	SE	p-value	Cohort
19	7587017	rs114077267	T	C	0.055	3.41E-01	8.88E-02	1.16E-04	YP1 AA
					0.041	6.85E-01	1.51E-01	5.09E-06	YP2 AA
						4.30E-01		1.97E-08	Meta AA YP1+YP2
19	7589061	rs115208233	G	C	0.053	3.61E-01	8.98E-02	5.62E-05	YP1 AA
					0.039	7.67E-01	1.55E-01	6.84E-07	YP2 AA
						4.63E-01		2.51E-09	Meta AA YP1+YP2
19	7594247	rs116181528	A	C	0.061	3.23E-01	8.40E-02	1.17E-04	YP1 AA
					0.048	7.01E-01	1.43E-01	8.89E-07	YP2 AA
						4.20E-01		6.85E-09	Meta AA YP1+YP2

Note: AA, African American. Chr, chromosome. SE, standard error. YP1 and YP2 represent the cohorts Yale-Penn 1 and Yale-Penn 2, respectively. Allele1 is the minor allele.