HHS Public Access

Author manuscript

Biol Psychiatry. Author manuscript; available in PMC 2019 November 15.

Published in final edited form as:

Biol Psychiatry. 2018 November 15; 84(10): 762–770. doi:10.1016/j.biopsych.2017.12.016.

Genome-wide association study identifies a regulatory variant of *RGMA* associated with opioid dependence in European Americans

Zhongshan Cheng[1], **Hang Zhou**[1], **Richard Sherva**[3], **Lindsay Farrer**[3], **Henry R. Kranzler**[4], and **Joel Gelernter**[1,2,*]

[1]Yale University School of Medicine, Department of Psychiatry, Division of Human Genetics, and VA CT Healthcare Center

[2]Yale University School of Medicine Departments of Genetics and Neuroscience

[3]Boston University Schools of Medicine and Public Health, Departments of Neurology, Ophthalmology, Genetics & Genomics, Epidemiology and Biostatistics

[4]University of Pennsylvania Perelman School of Medicine, Department of Psychiatry, Center for Studies of Addiction and Crescenz Veterans Affairs Medical Center

Abstract

Background—Opioid dependence (OD) is at epidemic levels in the United States. Genetic studies can provide insight into its biology.

Methods—We completed an OD genome-wide association study (GWAS) in 3058 opioid-exposed European-Americans, 1290 of whom met criteria for a DSM-IV diagnosis of OD. Analysis used DSM-IV criterion count.

Results—By meta-analysis of four cohorts, Yale-Penn 1 (n=1388), Yale-Penn 2 (n=996), Yale-Penn 3 (n=98) and SAGE (n=576), we identified a variant on chromosome 15, rs12442183, near *RGMA* (repulsive guidance molecule A (RGMa)), associated with OD ($P=1.3\times10^{-8}$). The association was also genome-wide significant (GWS) in Yale-Penn 1 taken individually, and nominally significant in two of the other three samples. The finding was further supported in a meta-analysis of all available opioid-exposed African-Americans (n=2014, 1106 meeting DSM-IV OD criteria; $P=3.0\times10^{-3}$) from three cohorts; there was nominal significance in two of these samples. Thus, of the seven subsamples examined in two populations, one was GWS and four of six were nominally significant. RGMa is a CNS axon guidance protein. Risk allele rs12442183*T

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

Conflict of Interest: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr. Kranzler has been a consultant or advisory board member for Alkermes, Indivior and Lundbeck. He is a member of the American Society of Clinical Psychopharmacology's Alcohol Clinical Trials Initiative, which was supported in the last three years by AbbVie, Alkermes, Ethypharm, Indivior, Lilly, Lundbeck, Otsuka, Pfizer, Arbor, and Amygdala Neurosciences. The other authors have no biomedical financial interests or potential conflicts of interest.

was correlated with higher expression of a specific *RGMA* transcript variant in frontal cortex ($P=2\times10^{-3}$). After chronic morphine injection, the homologous mouse gene (*Rgma*) was upregulated in C57BL/6J striatum. Co-expression analysis of 1301 brain samples revealed that *RGMA* mRNA expression was associated with that of four genes implicated in other psychiatric disorders, including *GRIN1*.

Conclusion—This is the first study to demonstrate an association of *RGMA* with OD. It provides a new lead into our understanding of OD pathophysiology.

Keywords

Genome-wide association; opioid dependence; *RGMA*; regulatory variant; European-Americans; psychiatric genetics

Introduction

Opioid dependence (OD) is a major cause of medical, legal, and social problems, resulting in increased mortality and decreased productivity. It is presently at epidemic levels in the United States. OD is a genetically complex disorder that is influenced by complex environmental factors (1), such as drug availability, making it difficult to identify genetic factors that underlie the disorder. Genome-wide association study (GWAS) is an unbiased way to search for novel genetic factors contributing to OD. We completed the first GWAS of OD (with DSM-IV criterion count as the phenotype definition, using all available subjects) and found two genome-wide significant single nucleotide polymorphisms (SNPs) in African-Americans (AAs) (2). These two SNPs map to *KCNC1* (potassium voltage-gated channel subfamily C member 1) and *KCNG2* (potassium voltage-gated channel modifier subfamily G member 2), respectively, which are involved in potassium signaling pathways. Pathway-based GWAS analysis also demonstrated an enrichment of genes involved in calcium signaling and long-term potentiation (2). There have been fewer findings in European-Americans (EAs), where the only GWAS to date to identify variants associated to OD in EAs showed an association with *CNIH3* (3). Additionally, we used GWAS to identify a variant close to the *OPRM1* locus that affected opioid dose requirements in AAs (4). These studies demonstrated that the use of an ordinal trait can provide more statistical power than a categorical diagnosis, as it can take account of affection severity.

In the present study, we used an ordinal trait analysis that was limited to OD-exposed individuals in our GWAS. We identified a variant of *RGMA*, which encodes the RGM domain family member A (RGMa, which was associated with OD in EAs at a genome-wide significant (GWS) level. The opioid-exposed EA subjects we studied previously (2) are included in the present study, augmented by additional subjects recruited more recently and we used a new analytic strategy that included only OD-exposed subjects. The finding was then replicated in both EAs and AAs.

Materials and Methods

Subjects, Genotyping and Imputation

Subjects in our three cohorts, Yale-Penn 1-3, were recruited from five eastern US sites to participate in studies of the genetics of drug (opioid or cocaine) or alcohol dependence (see Supplementary Table 1 for the specific relationship between the samples used in the present study and our previous study (2)). These samples contain unrelated individuals and a small proportion of small nuclear families (SNFs) originally collected for linkage studies. All participants provided written informed consent as approved by the institutional review board at each site. Certificates of confidentiality were issued by the National Institute on Drug Abuse and the National Institute on Alcohol Abuse and Alcoholism.

Subjects are designated as Yale-Penn 1, 2, and 3 by the recruitment epoch and microarray platform used for genotyping. Samples for Yale-Penn 1 (n=5540) were genotyped on the Illumina (San Diego, CA, USA) HumanOmni1-Quad v1.0 microarray at the Center for Inherited Disease Research and the Yale Center for Genome Analysis. A total of ~1M SNPs were targeted by microarray in Yale-Penn 1. Samples for Yale-Penn 2 (n=3675) were genotyped with the Illumina HumanCore Exome array, which includes a total of ~0.5M SNPs, including exomic SNPs and tagging SNPs. Genotyping in Yale-Penn 3 (n=592) was performed with the Illumina Multi-Ethnic Genotyping Array, an array targeting ~1.7M genome-wide markers and optimized for GWAS in populations of diverse ancestries. Yale-Penn 2 and Yale-Penn 3 genotyping data were obtained at the Gelernter Lab at Yale. Quality control (QC) for microarrays in each cohort were carried out using PLINK1.9 (5) based on the following criteria: (1) an individual genotype missing rate <2%, (2) a SNP genotype missing rate <2%, (3) a Hardy-Weinberg $P>1\times10^{-6}$, and (4) a minor allele frequency (MAF) >1%. After QC, there were ~0.70M, ~0.27M, and ~1.1M SNPs remaining in Yale-Penn 1-3 respectively, which were subjected to imputation and downstream analysis.

GWAS genotypes from all subjects taken separately in each sample group of Yale-Penn 1-3 were compared with the 1000 Genomes Project phase 1 reference panel (6). Eigensoft (7) was used for principal components analysis. The first 10 PC scores served to differentiate European-Americans (EAs) and African-Americans (AAs) through K-means clustering (8). Here, we present GWAS data only for EAs (our previous study presented numerous positive GWAS results in AAs). For each Yale-Penn cohort, genotypes from EAs and AAs were imputed together with Minimac3 implemented in Michigan Imputation Server (https://imputationserver.sph.umich.edu/index.html) with genotyped SNPs and the 1000 Genomes reference phase 3 release that contains haplotypes on 2504 samples (n = 5008 haplotypes for a total of ~81.2M polymorphic markers). The imputation data for Yale-Penn 1-3 were generated in dosage format separately. Under imputation $R^2>0.3$, there was an average of ~47.1M variants included in dosage data from Yale-Penn 1-3. The dosage genotypes for each cohort were transformed into hard call genotypes with PLINK1.9 (5). Using hard-call genotypes to analyze imputed GWAS genotypes has been reported previously (9). In our study, the dosage genotypes for each cohort were transformed into hard call genotypes with PLINK1.9, the rationale for which was to keep only high-quality dosage data by filtering with stricter QC parameters of genotype probability (GP >=0.9) using Plink1.9. The high-

quality hard-call genotyping data are easier to QC with Plink1.9 than dosage data, and the genotype data can be directly used in association tests with GEMMA (Genome-wide Efficient Mixed Model Association) (10) (improving the efficiency of computation), which incorporates and account for the relatedness among our family samples and model the relatedness to increase statistical power. Only genotypes with genotype imputation probability (GP) ≥0.9, an individual genotyping missing rate <5%, MAF >1% and missing call frequency <5% were kept for the association analysis. Following these QC efforts, there were ~13 M, ~9.8 M, and ~5.9 M variants left for Yale-Penn 1-3, respectively.

We also included GWAS data from SAGE (n=4077) (dbGaP study id phs000092.v1.p1) (11, 12), which were genotyped by the Illumina Human 1M array covering ~1 M total SNPs. SAGE genotypes were compared with the HapMap3 reference datasets, comprised of CEU, YRI, and CHB populations, and separated into EA and AA populations by using the first PC scores in principal components analysis with Eigensoft. The raw genotypes were imputed with IMPUTE2 (13). The 1000 Genomes phase 1 reference panel (6) containing phased haplotypes for 1092 individuals was used for imputation. Dosage genotypes were transformed into hard call genotypes with a custom perl script for heterozygous and homozygous genotypes, i.e., dosage probability for heterozygous genotypes were between 0.8 and 1.4 and homozygous genotypes of reference and alternative alleles ranged from 0 to 0.4 and 1.6 to 2, respectively, with others assigned as missing. The QC criteria for the individual genotyping missing rate <5%, MAF >1% and missing call frequency <5% were also utilized to filter GWAS data. After QC, ~14 M variants were retained for meta-analysis with Yale-Penn 1-3.

Analytic Methods

The sum of DSM-IV criterion counts for all EA OD-exposed individuals from the Yale-Penn 1-3 samples were used as an ordinal (quantitative) trait in association analysis (Supplementary figure 1). GEMMA (10) was performed for each cohort separately, with an adjustment for sex, age, the first three PCs of ancestry, and the degree of relatedness among subjects. In our previous study, we also adjusted for the number of other substance dependence criteria, but did not do so here. This was intended to capture more OD information, at the cost of possible confounding by association driven in part by substance use comorbidities. Summary statistics from GEMMA were reformatted for PLINK1.9 for meta-analysis, and all variants were evaluated by matching their chromosomal positions and two alleles among the four GWAS datasets. Variants with conflicting allele information that could not be resolved by being flipped were excluded from the meta-analysis that used PLINK1.9. The inverse variance method implemented in PLINK1.9 was used to generate fixed-effects meta-analysis P-values (meta-P). In the above process, ~8 M distinct variants present in at least in one cohort were reported in the meta-analysis; among these, ~4.7 M variants were present among all four cohorts.

Correlation between rs12442183 Genotype with *RGMA* Expression

The potential correlation between rs12442183 genotype and *RGMA* expression was tested with publicly available microarray data. First, R package ESLiMc (Exon Splicing by Linear Modeling Analysis core) (14) was used to analyze alternative expression of *RGMA*

transcripts among 11 different healthy human tissues, including breast, cerebellum, heart, kidney, liver, muscle, pancreas, prostate, spleen, testes, and thyroid, data from which were obtained from www.affymetrix.com and included in ESLiMc analysis by its author (14). The rationale for using ESLiMc was that if there are no alternative splicing events, an increase in the global expression of a gene should correlate with a higher expression of all of its exons. Linear regression was used by ESLiMc to establish the above relationship between each exon and its corresponding gene. The difference (referred to as gene-exon residual regression score) generated by ESLiMc between the observed exon expression and the predicted exon expression was interpreted as being due to an alternative splicing (14). Next, we tested the correlation between rs12442183 genotype and the expression of *RGMA* transcript variants. We downloaded rs12442183 genotype information from BRAINEAC (15) and reprocessed its corresponding microarray data from GEO (Gene Expression Omnibus) (16) (accession No. GSE60863) (17) with R package ESLiMc. GSE60863 is comprised of 1231 samples originating from 10 brain tissues (occipital cortex, frontal cortex, temporal cortex, hippocampus, intralobular white matter, cerebellar cortex, thalamus, putamen, substantia nigra, and medulla (inferior olivary nucleus)) derived from 134 healthy human brains from UK Biobank (17). These brain samples were confirmed to be free of neurodegenerative disorders and assayed by Affymetrix Human Exon 1.0 ST array. We also used PExFIns (18), an expression quantitative locus (eQTL) analysis software package, to test the association of rs12442183 genotype with *RGMA* mRNA expression in 423 lymphoblastoid cell lines (LCLs) across 6 global populations, including 73 CEU (Utah residents with northern and western European ancestry), 77 CHB (Han Chinese in Beijing, China), 72 JPT (Japanese in Tokyo, Japan), 80 LWK (Luhya in Webuye, Kenya), 42 MEX (Mexican ancestry in Los Angeles) and 79 YRI (Yoruba in Ibadan, Nigeria).

Besides *RGMA*, we also tested the association between rs12442183 genotype and the expression of other genes close to the SNP at both mRNA and exon level (see Integrative cis-eQTL Analysis in Supplementary Information).

Effect of Morphine Injection on Mouse *Rgma* Expression

The publicly available microarray experiment dataset GSE7762 was designed to determine the impact of genetic background on the transcriptional effects of morphine in mouse striatum (19). It used different morphine treatments, including a single morphine injection of 20 mg/kg, repeated (3 times daily for 5 days) morphine injections of 10-40 mg/kg, and saline, in four inbred mouse strains (129P3/J, DBA/2J, C57BL/6J and SWR/J). We used the R package GEOquery (20) to download the normalized GSE7762 microarray data from GEO and analyze the expression of *Rgma* (the mouse gene homologous to *RGMA*). One-way ANOVA implemented in MATLAB (Statistics Toolbox Release 2015b, MathWorks, Inc., Natick, Massachusetts, United States) was used to test the effect of morphine on *Rgma* expression in different mouse strains. Statistical significance was set at $P<0.05$.

Co-expression Analysis of *RGMA*

A novel bioinformatics pipeline was developed to search and prioritize genes co-expressed with *RGMA* in an unbiased way genome-wide. The correlations identified between *RGMA* and the identified co-expressed genes were subject to validation in brain tissue samples

(n=1301) downloaded from GEO (see *RGMA* Co-Expression Analysis in Supplementary Information).

Results

One GWS SNP was identified: rs12442183 (MAF 0.41 in EAs), which is ~110 kb upstream of *RGMA*, in Yale-Penn 1 ($P=4.7\times10^{-8}$, n=1388). Meta-analysis of this sample with the other three samples (Yale-Penn 2 and 3, and SAGE) improved the GWS association signal (meta-$P=1\times10^{-8}$, n=3058) (Figure 1 and Table 1). The robust association signals around the top SNP (Figure 1) and the observed negligible inflation factor in the QQ plot (Supplementary figure 2) support the validity of the GWAS result. As noted above, the association was GWS in Yale-Penn 1 taken individually. We also observed P <0.05 in both Yale-Penn 2 and Yale-Penn 3: the risk allele T of rs12442183 was significantly correlated with increased OD criterion counts in all cohorts, except for SAGE (Table 1 and Figure 2). SAGE has little variation in OD criterion count and >50% of individuals have a criterion count of 0, yielding low power in that sample.

The association of the lead SNP rs12442183 with OD was tested further in opioid-exposed AAs from three cohorts (Table 1), including Yale-Penn 1 (n=1506), Yale-Penn 2 (n=245) and SAGE (n=263) (MAF 0.33 in AAs). Yale-Penn 3 was not included because of the small number of AAs (n=12). Rs12442183 was significantly ($P=1.2\times10^{-2}$) or marginally significantly ($P=8\times10^{-2}$) associated with OD in Yale-Penn 1 and Yale-Penn 2, respectively. Possibly due to the same low variation of OD criterion counts in SAGE AAs, the association of rs12442183 with OD was not significant. However, by meta-analysis of the 2014 AA samples, the association of rs12442183 with OD was replicated (meta-$P=3\times10^{-3}$) (Table 1 and Figure 2).

The validity of the association of rs12442183 with OD in EAs and potentially also in AAs was demonstrated by the observation that 5 out of 7 cohorts (Table 1 and Figure 2) show at least nominal significance at this marker. Rs12442183 and rs1872052 (hg19 position chr15: 93752878, ~4.9 kb downstream of rs12442183) are the only two of the top 10 signals in the EA GWAS that were also nominally associated in AAs (Supplementary Table 2). Many of the variants that were associated in EAs were not present or were uncommon in AAs, so there was no relevant information for them. The meta-analysis P-value for rs12442183 was also significant (3.0×10^{-8}) across all cohorts (EA and AA), but the combined P-value is higher (i.e., less significant) than the European-specific P-value, which may be due to the different haplotype structures for EAs and AAs.

We next considered the effects of other substance use disorder criteria, including those for cocaine dependence (CD), alcohol dependence (AD), and nicotine dependence (ND), on the rs12442183 association signal. When we included CD, AD and ND criterion counts as covariates along with age, sex and first three PCs, there was decreased evidence of association of rs12442183 with OD in Yale-Penn 1-3 EAs, where nominal P-values were 3.7×10^{-7}, 0.17, and 0.08, respectively (meta-analysis $P=3.91\times10^{-7}$). Thus, the effect of CD, AD and ND on the observed association of rs12442183 with OD was observable, but limited.

The association of a previously implicated OD locus, *CNIH3* (3), was evaluated in the current GWAS meta-analysis of EAs. The lead SNP for *CNIH3*, rs10799590, was identified in a GWAS that compared opioid-dependent daily drug injectors with opioid misusers who never progressed to daily injection (5), whereas in our OD GWAS, the phenotype was OD criterion count. With that in mind, we note that the meta-analysis P-value for rs10799590 with OD criterion count was 0.038 in EAs.

The regulatory relationship between rs12442183 and *RGMA* expression was studied using publicly available microarray data. *RGMA* encodes a glycosylphosphatidylinositol-anchored glycoprotein (RGMa). There are three RGMa isoforms (Supplementary figure 3) that are generated by six *RGMA* transcript variants (Figure 3A), including NM_001166283 (variant 1), NM_001166286 (variant 2), NM_001166287 (variant 3), NM_020211 (variant 4), NM_001166288 (variant 5) and NM_001166289 (variant 6). Transcript variant 1 and 4 encode RGMa isoforms 1 and 3, respectively, and all other transcripts produce the same RGMa isoform 2. To identify alternative expression of *RGMA* among 11 healthy human tissues, four *RGMA* exons, including ENSE00001325239, ENSE00001319132, ENSE00001304160 and ENSE00001532613, were analyzed via R package ESLiMc (Figure 3A). ENSE00001325239 and ENSE00001319132 reside in the overlapping exonic regions in all 6 transcripts of *RGMA*, i.e., the expression signals of the two exons represent the total mRNA of *RGMA*. ENSE00001532613 is a specific exon representing transcript variant 4 (NM_020211, encoding RGMa isoform 3). Similarly, exon ENSE00001304160, specifically represents transcript variant 2 (NM_001166286, encoding RGMa isoform 2). Only ENSE00001532613, representing transcript variant 4, was variably expressed among the 11 tissues studied, and cerebellum was the only brain tissue with both the lowest expression and with the lowest gene-exon residual regression score for *RGMA* (Figure 3A and 3B). Further correlation analysis between rs12442183 genotype and the gene-exon expression residual scores of the four exons revealed that only ENSE00001532613 (representative of transcript variant 4) was differentially expressed ($P=2\times10^{-3}$) among rs12442183*T risk allele carriers in frontal cortex (Figure 3C). Among the other nine brain tissues, there were no significant correlations between rs12442183 genotype and gene-exon residual scores (data not shown). We also tested whether rs12442183 was an eQTL in LCLs across 6 global populations, including CEU, YRI, LWK, MEX, CHB and JPT, but found no significant associations (all linear model P-values >0.05), indicating that rs12442183 may be a specific eQTL for *RGMA* transcript variant 4 (encoding RGMa isoform 3) in frontal cortex.

We also tested whether the observed association between rs12442183 and OD was driven by the SNP's regulatory effect on other genes in the region (see Integrative cis-eQTL Analysis in Supplementary Information). With the independent data from the R package ESLiM, only *RGMA* and *MCTP2* showed alternative splicing in healthy brain tissue. We further correlated the genotype of rs12442183 with available exon expression of *RGMA*, *CHD2*, *FAM174B* and *MCTP2* from BRAINEAC. There was a significant correlation of rs12442183 genotype with specific exons of *RGMA* and *MCTP2*. However, the expression of *MCTP2* is very low among 27 human tissues.

We re-analyzed publicly available data from a microarray experiment (GEO accession No. GSE7762) designed to test the effects of different morphine treatments on gene expression

in mouse striatum tissue. By comparison with control and chronic morphine treatment, we found that the expression of mouse *Rgma* was significantly decreased (all ANOVA P-values<0.01) after acute morphine challenge in all 4 mouse strains, except for C57BL/6J. Compared to acute morphine challenge, significant upregulation of *Rgma* was found in C57BL/6J mouse striatum after chronic morphine injection (ANOVA P-value<0.01). These results demonstrate that *Rgma* was differentially regulated in mouse striatum when comparing C57BL/6J with other mouse strains.

The expression of six genes previously associated with psychiatric disorders and two potential regulators for *RGMA* were correlated with that of *RGMA* in brain tissue samples (n=1301). These genes are *GRIN1* (glutamate receptor, ionotropic, N-methyl D-aspartate 1), *KCNIP3* (Kv channel interacting protein 3, calsenilin), *PTCHD1* (patched domain containing 1), *FXYD6* (FXYD domain containing ion transport regulator 6), *DLG4* (discs, large homolog 4) and *NRG1* (neuregulin 1), as well as two potential regulators (*TCF3* and *TCF4*). *GRIN1*, *FXYD6*, *NRG1* and *DLG4* have been associated with schizophrenia (21-24), *KCNIP3* with Alzheimer's disease (25), *PTCHD1* with autism (26), and *NRG1* with psychosis (27). The E-box (5′-CANNTG-3′) motif of the two potential regulators, *TCF3* and *TCF4*, was found in the promoter region of *RGMA*, *GRIN1* and *PTCHD1*. We combined normalized expression data (z-scores) of 1301 brain tissue samples to test the pairwise correlation among the six genes. As shown in Supplementary figure 4A, the expressions of *GRIN1*, *KCNIP3*, *FXYD6*, *PTCHD1*, *RGMA*, *TCF3* and *TCF4* were highly inter-correlated. For *RGMA*, there was a strong positive correlation with *GRIN1* (Pearson correlation coefficient r=0.40 and $P=7.0\times10^{-50}$) and *KCNIP3* (r=0.38 and $P=1.35\times10^{-45}$) (Supplementary figure 4B), with the latter two also significantly inter-correlated (Supplementary figure 4A: r=0.73 and $P=1\times10^{-200}$).

Discussion

We report here results from a GWAS of opioid criterion count in EA subjects. There was a GWS association of an *RGMA* variant, rs12442183, in Yale-Penn 1. There was nominally significant evidence for association of the same variant in both Yale-Penn 2 and Yale-Penn 3, and when all EA samples (including SAGE, where no evidence of the association was obtained) were meta-analyzed together, the significance of the association finding increased. Because the minor allele frequency of rs12442183 was similar for the EA and AA populations, we further evaluated the association of rs12442183 with OD in the Yale-Penn AA sample. The nominally significant association in the AA meta-analysis (n=2014, meta-$P=3\times10^{-3}$) further supports the association. (AA replication of results from EA samples is not universally applicable, but in this case, we found evidence that we believe is of value.) Cis-eQTL analysis of *RGMA* in brain showed the rs12442183*T risk allele to be associated with higher expression of *RGMA* transcript variant 4 (encoding RGMa isoform 3) in frontal cortex. This observation is consistent with the upregulation of *Rgma* in C57BL/J6 mice treated with morphine chronically, which is not the case in three other mouse strains: DBA/2J, 129P3/J and SWR/J. C57BL/6J mice have a greater preference for oral self-administration of morphine than three other strains (19). Further co-expression analysis in large brain tissue expression datasets demonstrated significant correlation between the expression of *RGMA* and that of four genes that have been shown to affect risk for other

psychiatric disorders, *GRIN1* (schizophrenia) (23), *KCNIP3* (Alzheimer's disease) (25), *FXYD6* (schizophrenia) (21) and *PTCHD1* (autism) (26), as well as two transcription factors *TCF3* and *TCF4*. The latter gene TCF4 is a schizophrenia-risk gene (28). These *RGMA* co-expressed genes are critical to normal brain function, supporting the potential involvement of *RGMA* in related biological processes.

The protein product of *RGMA*, RGMa, is a glycosylphosphatidylinositol-anchored glycoprotein that exists in membrane-bound and soluble forms (29), both of which are potent regulators of cell death and inhibitors of nerve growth. RGMa was first reported to be a developmental neurite growth inhibitor (30). It has since emerged as an important factor inhibiting neuronal regeneration and functional recovery (31, 32). RGMa inhibits neurite growth by binding to its neuronal receptor, neogenin, a member of the immunoglobulin (Ig) superfamily. Neogenin overexpression and RGMa down expression in the developing embryonic neural tube induce apoptosis (33). RGMa may be involved in neuroplasticity consistent with increased secretion of RGMa in stroke- or injury-damaged human brain tissue (34), in the plaque regions of Alzheimer's disease patients (35), in the substantia nigra of Parkinson's disease patients (36), and in multiple sclerosis (MS) patients (37). Considering the example of progressive MS (38), RGMa was found in active and chronic lesions, and in normal-appearing gray and white matter, and was expressed by cellular meningeal infiltrates in autopsy material. Levels of soluble RGMa in the cerebrospinal fluid were decreased in progressive MS patients for whom treatment resulted in functional improvement. Antibody-mediated neutralization of RGMa or downregulation of RGMa can result in neuroprotection and enhanced functional recovery (38). Whether RGMa plays similar roles in OD patients and thereby modulates addiction risk warrants further study.

The mouse experiment that investigated the effects of morphine on the striatal transcriptome (19) identified 661 morphine-responsive genes. *Rgma* was one of the top 98 genes that were down-regulated with *acute* morphine treatment. Consistent downregulation of *Rgma* was observed among 3 mouse strains, including DBA/2J, 129P3/J and SWR/J, but not C57BL/6J. However, *Rgma* was specifically upregulated in C57BL/6J mice (in contrast to the other 3 mouse strains) following *chronic* morphine treatment. The different responses of these mouse strains after chronic or acute morphine treatment may reflect genetic differences, as C57BL/6J mice have been reported to have the greatest preference for orally self-administered morphine of the strains studied (19). Further, in the current study, risk rs12442183*T-allele carriers demonstrated a higher expression of one *RGMA* transcript in frontal cortex. As downregulation of RGMa can result in neuroprotection and enhanced functional recovery (38) and increased RGMa in brain is associated with risk of multiple neurological disorders, including Alzheimer's disease, Parkinson's disease, and multiple sclerosis, downregulation of RGMa may have a protective effect on neurons after acute morphine treatment, while upregulation of RGMa may induce brain vulnerability with chronic morphine treatment.

In the analysis of several brain tissue samples, *RGMA* expression was associated with the expression of four other genes related to psychiatric traits, *GRIN1*, *KCNIP3*, *FXYD6* and *PTCHD1*. *GRIN1* was the gene that was most highly co-expressed (r=0.40) with *RGMA* in brain. It plays a key role in synaptic plasticity, synaptogenesis, excitotoxicity, memory

acquisition and learning (39). *GRIN1* can tether the μ-opioid receptor (the main biological target of opioid drugs) and subsequently immobilize the μ-opioid receptor within lipid rafts (40). The overexpression of *GRIN1* increases the amount of μ-opioid receptor in lipid rafts and the magnitude of receptor signaling. The significant correlation between the expression of other genes (*PTCHD1*, *FXYD6*, and *KCNIP3*) and that of *RGMA*, may reflect their critical function in the brain (21, 26, 41). These findings raise the question of whether, in response to opioid stimulation in brain, there is crosstalk between *RGMA* and *GRIN1*, and possibly *KCNIP3*, *FXYD6* and *PTCHD1* as well.

Different proportions of prescribed vs. illegal users of opioids in the Yale-Penn and the SAGE samples could have affected our results and may help to understand the different results from these samples. Neither the SSADDA (Yale-Penn) nor the SSAGA (SAGE) differentiates prescribed from illegal-recreational use of opioids, and therefore some subjects who endorsed DSM-IV OD criteria probably did so despite using opioids therapeutically. However, a majority (59.5%) of individuals in our samples who met criteria for OD endorsed at least five of the seven diagnostic criteria (mean criterion count = 6.6, Supplementary figure 1), suggesting that the GWS finding is attributable to moderate-to-severe OD, rather than being contributed to substantially by individuals who used opioids as prescribed, but required an increased dosage for pain control or who experienced withdrawal symptoms when opioid treatment was discontinued. The much greater proportion of participants in the Yale-Penn samples than SAGE who met the maximum number of (i.e., seven) OD criteria supports that interpretation. Also, compared to our (Yale-Penn) samples, the SAGE samples were recruited for alcoholism (42), cocaine dependence (43), or nicotine dependence (44), rather than OD. Therefore, a much higher percentage of opioid-exposed individuals in SAGE are “controls”: viz. individuals who used opioids fewer than 11 times lifetime and would have skipped out of the remainder of the SSAGA opioid section and had a null criterion count (Supplementary figure 1). Additionally, the relative percentage of opioid-exposed individuals varies considerably across our three Yale-Penn EA samples because of the different recruiting epochs, although the assessment protocol did not change: the earlier Yale-Penn samples were more focused on OD recruitment than the later ones.

In summary, we report a GWS association of rs12442183, at the *RGMA* locus, with OD. In contrast to our earlier study, the present study includes only opioid-exposed subjects selected from an expanded sample; and our analyses for this study did not co-vary for the criterion count for other substance use disorder traits. We believe that the former two differences were key in our identification of the present very robust association finding. The finding was GWS in our largest sample taken separately, with nominally significant associations of the same variant in two of three other available samples of EAs. Further support was obtained in AA samples. Our current data support a model in which the regulatory variant rs12442183 affects expression of *RGMA* transcript variant 4 in frontal cortex, and the consequent variation in RGMa isoform 3 may alter the response to opioids, a potential underlying mechanism for its association with OD. This mechanism and locus merit further study, which could increase our understanding of the genetics and pathophysiology of OD.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was supported by National Institutes of Health grants RC2 DA028909, R01 DA12690, R01 DA12849, R01 DA18432, R01 AA11330, R01 AA017535, and the VA Connecticut Healthcare Center and Philadelphia VA MIRECCs. We appreciate the work in recruitment and assessment by James Poling, PhD at Yale University School of Medicine and the APT Foundation; Roger Weiss, MD at McLean Hospital; by Kathleen Brady, MD/PhD and Raymond Anton, MD at the Medical University of South Carolina; and David Oslin, MD at the University of Pennsylvania. 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. We thank Ann Marie Lacobelle, MS and Christa Robinson, BS, who provided technical assistance. We thank Yaira Nunez and Michelle Slivinsky who led the quality control effort for the Semi-Structured Assessment for Drug Dependence and Alcoholism interviews and phenotyping for the study sample. We thank John Farrell, PhD, Section of Biomedical Genetics, Boston University School of Medicine, who provided database management assistance. We also thank three anonymous reviewers for their constructive suggestions on the manuscript.

References

1. Mistry CJ, Bawor M, Desai D, Marsh DC, Samaan Z. Genetics of Opioid Dependence: A Review of the Genetic Contribution to Opioid Dependence. Curr Psychiatry Rev. 2014; 10:156–167. [PubMed: 25242908]
2. 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; 76:66–74. [PubMed: 24143882]
3. 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. [PubMed: 26239289]
4. Smith AH, Jensen KP, Li J, Nunez Y, Farrer LA, Hakonarson H, et al. Genome-wide association study of therapeutic opioid dosing identifies a novel locus upstream of OPRM1. Mol Psychiatry. 2017; 22:346–352. [PubMed: 28115739]
5. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015; 4:7. [PubMed: 25722852]
6. Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, et al. 1000-Genomes-Project-Consortium. A map of human genome variation from population-scale sequencing. Nature. 2010; 467:1061–1073. [PubMed: 20981092]
7. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006; 38:904–909. [PubMed: 16862161]
8. Sherva R, Wang Q, Kranzler H, Zhao H, Koesterer R, Herman A, et al. Genome-wide association study of cannabis dependence severity, novel risk variants, and shared genetic risks. JAMA Psychiatry. 2016; 73:472–480. [PubMed: 27028160]
9. Radmanesh F, Devan WJ, Anderson CD, Rosand J, Falcone GJ. Alzheimer's Disease Neuroimaging I. Accuracy of imputation to infer unobserved APOE epsilon alleles in genome-wide genotyping data. Eur J Hum Genet. 2014; 22:1239–1242. [PubMed: 24448547]
10. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nat Genet. 2012; 44:821–824. [PubMed: 22706312]
11. Bierut LJ, Strickland JR, Thompson JR, Afful SE, Cottler LB. Drug use and dependence in cocaine dependent subjects, community-based individuals, and their siblings. Drug Alcohol Depen. 2008; 95:14–22.
12. Edenberg HJ. The collaborative study on the genetics of alcoholism: an update. Alcohol Res Health. 2002; 26:214–218. [PubMed: 12875050]

13. 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:e1000529. [PubMed: 19543373]
14. Risueno A, Roson-Burgo B, Dolnik A, Hernandez-Rivas JM, Bullinger L, De Las Rivas J. A robust estimation of exon expression to identify alternative spliced genes applied to human tissues and cancer samples. BMC Genomics. 2014; 15:879. [PubMed: 25297679]
15. Hernandez DG, Nalls MA, Moore M, Chong S, Dillman A, Trabzuni D, et al. Integration of GWAS SNPs and tissue specific expression profiling reveal discrete eQTLs for human traits in blood and brain. Neurobiol Dis. 2012; 47:20–28. [PubMed: 22433082]
16. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets-update. Nucleic Acids Res. 2013; 41:D991–D995. [PubMed: 23193258]
17. 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; 17:1418–1428. [PubMed: 25174004]
18. Cheng Z, Chu H, Fan Y, Li C, Song YQ, Zhou J, et al. PExFInS: an integrative post-GWAS explorer for functional indels and SNPs. Sci Rep. 2015; 5:17302. [PubMed: 26612672]
19. Korostynski M, Piechota M, Kaminska D, Solecki W, Przewlocki R. Morphine effects on striatal transcriptome in mice. Genome Biol. 2007; 8
20. Sean D, Meltzer PS. GEOquery: a bridge between the gene expression omnibus (GEO) and BioConductor. Bioinformatics. 2007; 23:1846–1847. [PubMed: 17496320]
21. Choudhury K, McQuillin A, Puri V, Pimm J, Datta S, Thirumalai S, et al. A genetic association study of chromosome 11q22-24 in two different samples implicates the FXYD6 gene, encoding phosphohippolin, in susceptibility to schizophrenia. Am J Hum Genet. 2007; 80:664–672. [PubMed: 17357072]
22. Jonsson EG, Saetre P, Vares M, Andreou D, Larsson K, Timm S, et al. DTNBP1, NRG1, DAOA, DAO and GRM3 polymorphisms and schizophrenia: an association study. Neuropsychobiology. 2009; 59:142–150. [PubMed: 19439994]
23. Demontis D, Nyegaard M, Buttenschon HN, Hedemand A, Pedersen CB, Grove J, et al. Association of GRIN1 and GRIN2A-D with schizophrenia and genetic interaction with maternal herpes simplex virus-2 infection affecting disease risk. Am J Med Genet B Neuropsychiatr Genet. 2011; 156B:913–922. [PubMed: 21919190]
24. Tsai SJ, Hong CJ, Cheng CY, Liao DL, Liou YJ. Association study of polymorphisms in post-synaptic density protein 95 (PSD-95) with schizophrenia. J Neural Transm (Vienna). 2007; 114:423–426. [PubMed: 17093888]
25. Pereira AC, Gray JD, Kogan JF, Davidson RL, Rubin TG, Okamoto M, et al. Age and Alzheimer/'s disease gene expression profiles reversed by the glutamate modulator riluzole. Mol Psychiatry. 2017; 22:296–305. [PubMed: 27021815]
26. Torrico B, Fernandez-Castillo N, Hervas A, Mila M, Salgado M, Rueda I, et al. Contribution of common and rare variants of the PTCHD1 gene to autism spectrum disorders and intellectual disability. Eur J Hum Genet. 2015; 23:1694–1701. [PubMed: 25782667]
27. Okochi T, Kishi T, Ikeda M, Kitajima T, Kinoshita Y, Kawashima K, et al. Genetic association analysis of NRG1 with methamphetamine-induced psychosis in a Japanese population. Prog Neuropsychopharmacol Biol Psychiatry. 2009; 33:903–905. [PubMed: 19394386]
28. Steinberg S, de Jong S, Andreassen OA, Werge T, Borglum AD, Mors O, et al. Common variants at VRK2 and TCF4 conferring risk of schizophrenia. Hum Mol Genet. 2011; 20:4076–4081. [PubMed: 21791550]
29. Tassew NG, Charish J, Seidah NG, Monnier PP. SKI-1 and Furin generate multiple RGMa fragments that regulate axonal growth. Dev Cell. 2012; 22:391–402. [PubMed: 22340500]
30. Monnier PP, Sierra A, Macchi P, Deitinghoff L, Andersen JS, Mann M, et al. RGM is a repulsive guidance molecule for retinal axons. Nature. 2002; 419:392–395. [PubMed: 12353034]
31. Hata K, Fujitani M, Yasuda Y, Doya H, Saito T, Yamagishi S, et al. RGMa inhibition promotes axonal growth and recovery after spinal cord injury. JCB. 2006; 173:47–58. [PubMed: 16585268]

32. Kyoto A, Hata K, Yamashita T. Synapse formation of the cortico-spinal axons is enhanced by RGMa inhibition after spinal cord injury. Brain Res. 2007; 1186:74–86. [PubMed: 17996222]
33. Matsunaga E, Tauszig-Delamasure S, Monnier PP, Mueller BK, Strittmatter SM, Mehlen P, et al. RGM and its receptor neogenin regulate neuronal survival. Nat Cell Biol. 2004; 6:749–755. [PubMed: 15258591]
34. Schwab JM, Monnier PP, Schluesener HJ, Conrad S, Beschorner R, Chen L, et al. Central nervous system injury-induced repulsive guidance molecule expression in the adult human brain. Arch Neurol. 2005; 62:1561–1568. [PubMed: 16216939]
35. Satoh J, Tabunoki H, Ishida T, Saito Y, Arima K. Accumulation of a repulsive axonal guidance molecule RGMa in amyloid plaques: a possible hallmark of regenerative failure in Alzheimer's disease brains. Neuropathol Appl Neurobiol. 2013; 39:109–120. [PubMed: 22582881]
36. Bossers K, Meerhoff G, Balesar R, van Dongen JW, Kruse CG, Swaab DF, et al. Analysis of gene expression in Parkinson's disease: possible involvement of neurotrophic support and axon guidance in dopaminergic cell death. Brain Pathol. 2009; 19:91–107. [PubMed: 18462474]
37. Nohra R, Beyeen AD, Guo JP, Khademi M, Sundqvist E, Hedreul M, et al. RGMA and IL21R show association with experimental inflammation and multiple sclerosis. Genes Immun. 2010; 11:279–293. [PubMed: 20072140]
38. Demicheva E, Cui YF, Bardwell P, Barghorn S, Kron M, Meyer AH, et al. Targeting repulsive guidance molecule a to promote regeneration and neuroprotection in multiple sclerosis. Cell Rep. 2015; 10:1887–1898. [PubMed: 25801027]
39. Mundo E, Tharmalingham S, Neves-Pereira M, Dalton EJ, Macciardi F, Parikh SV, et al. Evidence that the N-methyl-D-aspartate subunit 1 receptor gene (GRIN1) confers susceptibility to bipolar disorder. Mol Psychiatry. 2003; 8:241–245. [PubMed: 12610658]
40. Ge X, Qiu Y, Loh HH, Law PY. GRIN1 regulates micro-opioid receptor activities by tethering the receptor and G protein in the lipid raft. J Biol Chem. 2009; 284:36521–36534. [PubMed: 19861419]
41. Jo DG, Lee JY, Hong YM, Song S, Mook-Jung I, Koh JY, et al. Induction of pro-apoptotic calsenilin/DREAM/KChIP3 in Alzheimer's disease and cultured neurons after amyloid-beta exposure. J Neurochem. 2004; 88:604–611. [PubMed: 14720210]
42. Wang JC, Foroud T, Hinrichs AL, Le NX, Bertelsen S, Budde JP, et al. A genome-wide association study of alcohol-dependence symptom counts in extended pedigrees identifies C15orf53. Mol Psychiatry. 2013; 18:1218–1224. [PubMed: 23089632]
43. Bierut LJ, Strickland JR, Thompson JR, Afful SE, Cottler LB. Drug use and dependence in cocaine dependent subjects, community-based individuals, and their siblings. Drug Alcohol Depen. 2008; 95:14–22.
44. Luo ZH, Alvarado GF, Hatsukami DK, Johnson EO, Bierut LJ, Breslau N. Race differences in nicotine dependence in the Collaborative Genetic study of Nicotine Dependence (COGEND). Nicotine Tob Res. 2008; 10:1223–1230. [PubMed: 18629733]
45. Ward LD, Kellis M. HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. Nucleic Acids Res. 2016; 44:D877–881. [PubMed: 26657631]

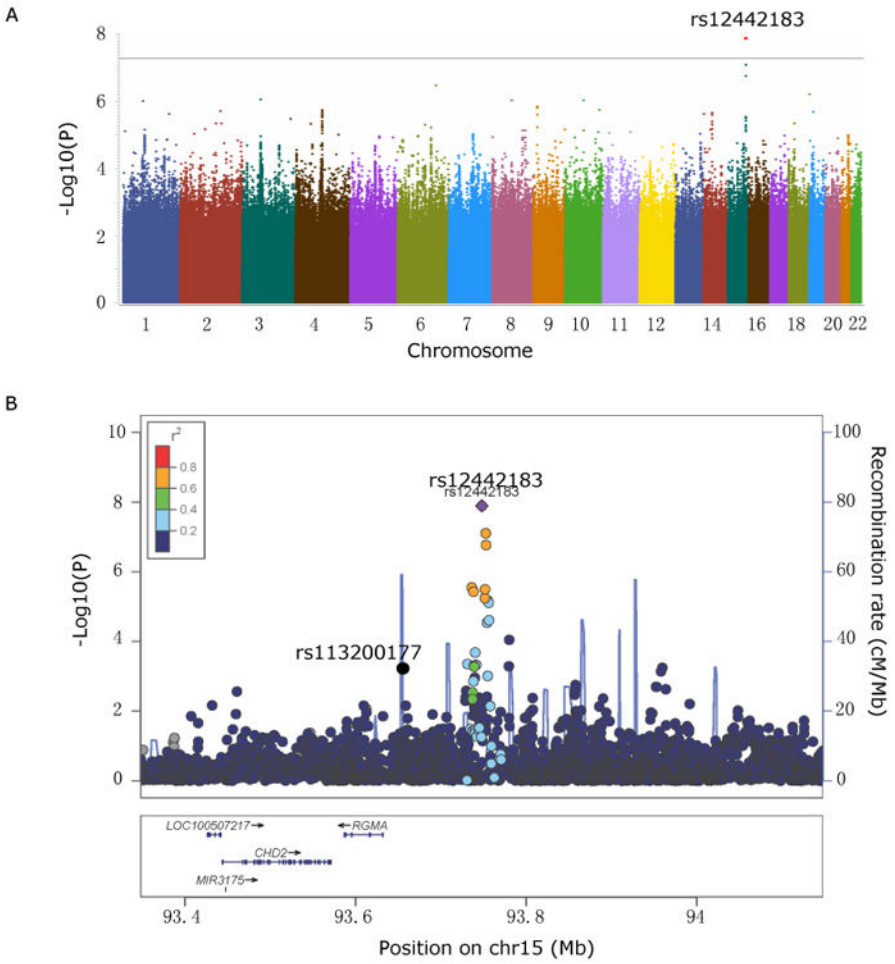

Figure 1. Manhattan plots showing genome-wide association signals of opioid dependence in opioid-exposed European-Americans by meta-analysis

(**A**) Manhattan plot with one significant variant, rs12442183, on chromosome 15, in meta-analysis of four OD cohorts. The line in the plot represents the genome-wide significance cutoff (5×10^{-8}). (**B**) Regional Manhattan plot demonstrates rs12442183 is close to gene *RGMA* (and regulates *RGMA* expression; see Figure 3 and text). The light blue line and right Y-axis show the observed recombination rate in the 1000 Genomes Project European samples (EUR, hg19). The SNPs are colored in accordance to R^2 with rs12442183 (in purple), except for rs113200177, which is specifically highlighted in dark. rs113200177 (T/TG, meta-P=2×10^{-3}) is located in the promoter of *RGMA* and is predicted to be able to destroy the E-box motif (5′-CANNTG-3′) of transcription factor 3 or 4 (*TCF3* or *TCF4*) by Haploreg4 (45). By adjusting the effect of the indel in conditional association analysis for rs12442183, the association signal of rs12442183 became stronger in Yale-Penn 1 taken individually (P=1.9×10^{-8}).

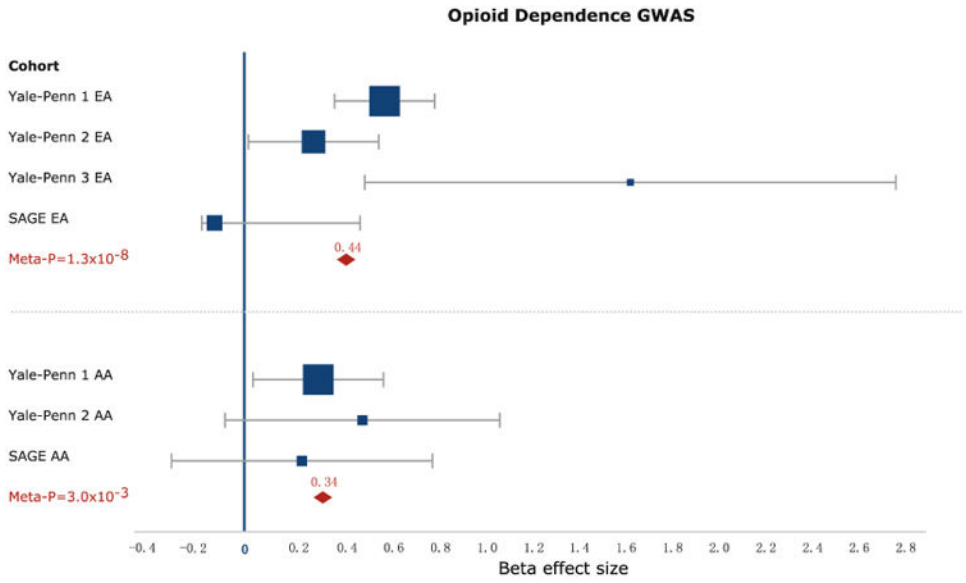

Figure 2. Forest plot for rs12442183 risk allele T in opioid dependence GWAS

Effect Beta and its 95% confidence interval are plotted with squares (proportional to sample size in each cohort) and horizontal lines with whiskers, respectively. The vertical line (beta=0) indicates no effect. Names of cohorts are shown on the left. Meta-analysis P-values and their corresponding effect betas (red diamonds) for European-American (EA) and African-American (AA) populations are also provided. The forest plot was generated by DistillerSR Forest Plot Generator from Evidence Partners (https://www.evidencepartners.com/resources/forest-plot-generator).

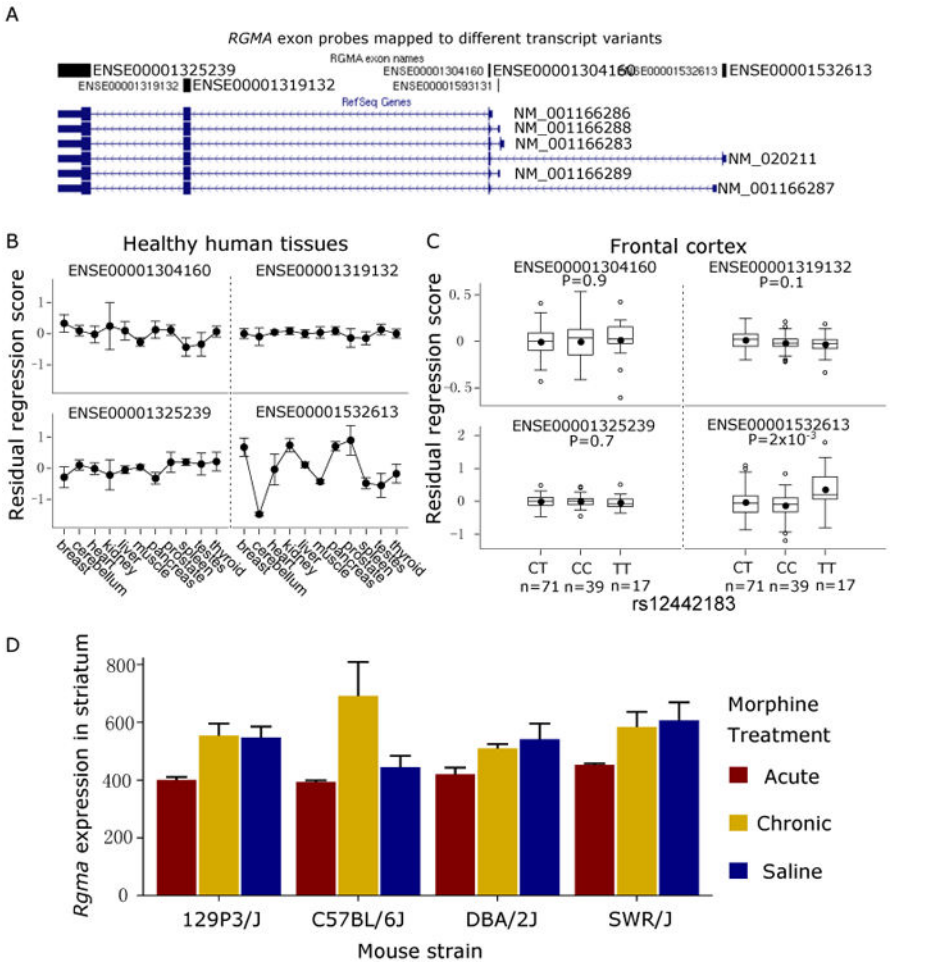

Figure 3. rs12442183 is a regulatory variant of *RGMA* and *Rgma* is an acute opioid response gene in mouse

(**A**) Six transcripts of *RGMA* and four exons used for alternative splicing analysis are mapped together. The relationship between the expression of these four exons and *RGMA* gene expression was determined by using a linear regression model implemented in R package ESLiMc. The difference between the observed exon expression and the predicted exon expression is named as residual regression score. The alternative splicing events of *RGMA* across 11 healthy human tissues are displayed in (**B**), with dot and error-bar represent mean and standard deviation of 3 replicates, respectively. Boxplots (**C**) are used to illustrate the association between rs12442183 genotype and the residual regression scores of four exons in brain tissue frontal cortex in European samples. Dot and line within each boxplot represent mean and median, respectively, and box region indicates the range from first quantile to third quantile. Up and down whiskers, as well as dots outside box region represent maximum value, minimum value, and outliers, respectively. (**D**) The expression of *Rgma* (homologous gene in mouse) among different mouse strains acutely after morphine injection (a single injection, 20 mg/kg) in comparison with chronic injections (repeated morphine administration, 10-40 mg/kg, 3 times daily for 5 days). X-axis shows the four inbred mouse strains (129P3/J, DBA/2J, C57BL/6J and SWR/J). Y-axis demonstrates the mRNA expression of mouse gene *Rgma*. The expression data were derived from NCBI Gene

Expression Omnibus (GEO) dataset GSE7762. The standard deviation is at the top of each bar.

Table 1

Meta-analysis of rs12442183 Risk Allele T in OD Exposed Individuals

Population	Cohort[1]	Total Subjects No.	Exposed Subject No.	Criterion Counts[2]	MAF[3]	Effect Beta (95% CI) [4]	Imputation R[2]	P
European-Americans (EA)	Yale-Penn 1	1768	1388	5.0 ± 2.8	0.41	0.60 (0.39 - 0.82)	0.98	4.7×10^{-8}
	Yale-Penn 2	1725	996	4.7 ± 2.9	0.38	0.30 (0.02 - 0.58)	0.82	3.8×10^{-2}
	Yale-Penn 3	403	98	3.6 ± 3.2	0.41	1.6 (0.52 - 2.8)	0.81	5.3×10^{-3}
	SAGE	2752	576	2.0 ± 2.8	0.38	-0.14 (-0.18 - 0.50)	0.98	3.9×10^{-1} *
	Meta-analysis					**0.44**		**1.3×10^{-8}**
African-Americans (AA)	Yale-Penn 1	3289	1209	3.8 ± 3.1	0.33	0.14(0.04 - 0.6)	0.98	2.3×10^{-2}
	Yale-Penn 2	1744	245	3.4± 3.1	0.3	0.50 (-0.08 - 1.1)	0.82	8.0×10^{-2}
	Yale-Penn 3	139	12	-	-	-	-	-
	SAGE	1312	252	2.5 ± 3.0	0.34	0.25 (-0.31 - 0.81)	0.98	3.7×10^{-1}
	Meta-analysis					**0.34**		**3.0×10^{-3}**

[1] Illumina microarray HumanOmni1-Quad v1.0, HumanCore Exome array, and Human 1M were used to genotype samples for Yale-Penn 1-3, respectively.

[2] Mean ± standard deviation for criterion counts of ordinal traits.

[3] Minor allele frequency.

[4] 95% confidence interval of continuous effect beta based on likelihood ratio test by using the software Genome-wide Efficient Mixed Model Association (GEMMA).

* indicates the opposite effect direction.