

Transcriptome association studies of neuropsychiatric traits in African Americans implicate *PRMT7* in schizophrenia

Peter N. Fiorica^{1,2} and Heather E. Wheeler^{2,3,4,5}

¹ Department of Chemistry and Biochemistry, Loyola University Chicago, Chicago, IL, USA

² Department of Biology, Loyola University Chicago, Chicago, IL, USA

³ Program in Bioinformatics, Loyola University Chicago, Chicago, IL, USA

⁴ Department of Computer Science, Loyola University Chicago, Chicago, IL, USA

⁵ Department of Public Health Sciences, Loyola University Chicago, Maywood, IL, USA

ABSTRACT

In the past 15 years, genome-wide association studies (GWAS) have provided novel insight into the genetic architecture of various complex traits; however, this insight has been primarily focused on populations of European descent. This emphasis on European populations has led to individuals of recent African descent being grossly underrepresented in the study of genetics. With African Americans making up less than 2% of participants in neuropsychiatric GWAS, this discrepancy is magnified in diseases such as schizophrenia and bipolar disorder. In this study, we performed GWAS and the gene-based association method PrediXcan for schizophrenia ($n = 2,256$) and bipolar disorder ($n = 1,019$) in African American cohorts. In our PrediXcan analyses, we identified *PRMT7* ($P = 5.5 \times 10^{-6}$, local false sign rate = 0.12) as significantly associated with schizophrenia following an adaptive shrinkage multiple testing adjustment. This association with schizophrenia was confirmed in the much larger, predominantly European, Psychiatric Genomics Consortium. In addition to the *PRMT7* association with schizophrenia, we identified rs10168049 ($P = 1.0 \times 10^{-6}$) as a potential candidate locus for bipolar disorder with highly divergent allele frequencies across populations, highlighting the need for diversity in genetic studies.

Submitted 13 May 2019
Accepted 27 August 2019
Published 26 September 2019

Corresponding author

Heather E. Wheeler,
hwheeler1@luc.edu

Academic editor
Hossein Khiabani

Additional Information and
Declarations can be found on
page 12

DOI 10.7717/peerj.7778

© Copyright
2019 Fiorica and Wheeler

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

Subjects Genetics, Genomics, Psychiatry and Psychology

Keywords GWAS, PrediXcan, Population genetics, Gene expression, Schizophrenia, Bipolar disorder

INTRODUCTION

Individuals of recent African ancestry have been grossly underrepresented in genomic studies. African American participants make up about 2.0% of all genome-wide association studies (GWAS) subjects (*Sirugo, Williams & Tishkoff, 2019*). Specifically, individuals of African ancestry make up only 1.2% of all neuropsychiatric GWAS (*Quansah & McGregor, 2018*). With the advent of polygenic risk scores, accuracy in disease prediction is critical to the development of precision medicine (*Khera et al., 2018*); however, the lack of representative diversity in the study of genomics has impacted the accuracy of genetic risk

prediction across diverse populations. Despite similar incidences of schizophrenia across European and African ancestry populations (*De Candia et al., 2013; Whiteford et al., 2013*), Africans have been predicted to have significantly less disease risk than their European counterparts using current GWAS summary statistics (*Martin et al., 2017*). Inaccuracy in predicting disease risk across populations can lead to further disparities in health and treatment of underrepresented populations. To prevent misclassification of genetic risk, further work in the genetics underlying complex traits in African Americans is needed (*Manrai et al., 2016*). In an attempt to address this discrepancy in genetic risk prediction, we performed a series of genetic association tests for schizophrenia and bipolar disorder in two cohorts of African American individuals (*Manolio et al., 2007; Suarez et al., 2006; Smith et al., 2009*).

Schizophrenia and bipolar disorder are two heritable neuropsychiatric disorders whose genetic components have been attributed to the cumulative effect of thousands of loci across the genome (*Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014; Li et al., 2017; Ikeda et al., 2017b*). Past work shows that the genetic architectures of these two disorders significantly overlap (*Bhalala et al., 2018; Allardyce et al., 2018; The International Schizophrenia Consortium, 2009; Stahl et al., 2019*). Up to this point, the largest GWAS of schizophrenia and bipolar disorder comprise hundreds of thousands of individuals primarily of European descent (*Stahl et al., 2019; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014; Li et al., 2017*). While studies of neuropsychiatric diseases in European and Asian ancestry populations continue to grow, the scarcity of studies in African American populations persists (*Ikeda et al., 2017a*).

To date, one of the largest GWAS of schizophrenia in an African American population was completed by the Genetic Association Information Network (GAIN) (*Manolio et al., 2007*); however, this study found no single nucleotide polymorphisms (SNPs) to be genome-wide significant and offered little insight into the potential function of genes in schizophrenia in African Americans. The GAIN has also performed one of the largest GWAS of bipolar disorder in African Americans (*Manolio et al., 2007; Smith et al., 2009*). Similar to the findings of the GAIN GWAS of schizophrenia, *Smith et al. (2009)* found no SNPs significantly associated with bipolar disorder. In addition to a traditional GWAS using a logistic regression, we performed two gene-level association tests: PrediXcan and MultiXcan (*Gamazon et al., 2015; Barbeira et al., 2019*). PrediXcan offers a series of advantages to SNP-level analyses in detecting genetic association and functionality. First, since tests are being conducted at the gene level, PrediXcan has a lower multiple-testing burden compared to GWAS. Additionally, our understanding of functional pathways are more easily constructed for genes compared to SNPs. By using gene expression as an intermediate phenotype between genetic variation and complex phenotypes, PrediXcan results can help elucidate genetic mechanisms compared to GWAS. We completed our genetic association tests for schizophrenia in a cohort of 2,256 self-identified African American individuals from the genome-wide linkage scan of African American families (*Suarez et al., 2006*) and the GAIN (*Manolio et al., 2007*). For our study of bipolar disorder, we performed these association tests in 1,019 African American individuals from the GAIN. Using these data, we identified one gene significantly associated with schizophrenia and

Table 1 Cohort characteristics. Three separate cohorts were integrated into two main cohorts characterized by phenotype. After merging the two schizophrenia cohorts using PLINK, the number of post-QC SNPs became identical.

dbGaP accession number	phs000021.v3.p2	phs000167.v1.p1	phs000017.v3.p1
Phenotype	Schizophrenia	Schizophrenia	Bipolar disorder
Total individuals	2,220	120	1,045
Cases	1,241	15	359
Controls	979	105	686
Post-QC individuals	2,256	2,256	1,019
Pre-QC SNPs	845,814	909,622	867,411
Post-QC SNPs	742,015	742,015	721,050
Post-imputation SNPs ($r^2 > 0.8$, MAF > 0.01)	12,780,487	12,780,487	12,799,548

tested it for replication in the Psychiatric Genomics Consortium (PGC) GWAS of schizophrenia (*Barbeira et al., 2018*; *Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014*; *Psychiatric GWAS Consortium Bipolar Disorder Working Group, 2011*).

METHODS

Cohorts

The individuals in the cohorts used in this study were all self-identified African Americans. We acquired genotype and phenotype information for these individuals from the National Center for Biotechnology Information database of Genotypes and Phenotypes (dbGaP). The project was confirmed exempt from human subjects federal regulations under exemption number 4 by the Loyola University Chicago Institutional Review Board (project number 2014). Whole Genome genotypes and phenotypic information were acquired from three separate accessions in dbGaP (Table 1). All genotype information for these three accessions were acquired using Affymetrix Genome-Wide Human SNP Array 6.0, covering 934,940 SNPs. In total, our studies included 2,256 and 1,019 individuals for schizophrenia and bipolar disorder, respectively.

Case-control criteria for both phenotypes were determined using the DSM-IV (Diagnostic and Statistical Manual of Mental Disorders) as described previously (*Smith et al., 2009*). Following the update from DSM-IV to DSM-V, 40 individuals previously identified as cases under DSM-IV were removed from the study because they no longer met the case criteria for DSM-V.

Quality control and imputation

Following the download of data from dbGaP, we isolated genotypes and phenotypes of African Americans from those of European Americans or individuals of unidentified ethnicities for each cohort. We merged the PLINK binary files from the two schizophrenia studies. At this point 906,425 SNPs were genotyped in 2,256 individuals in the schizophrenia cohort and 867,411 SNPs and 1,019 individuals in the bipolar disorder cohort. While the schizophrenia cohort included individuals from GAIN and the genome-wide linkage scan of African American families (*Suarez et al., 2006*; *Manolio et al.,*

2007), throughout the rest of the paper, we will refer to the combined cohort as GAIN. In each cohort, we removed SNPs with genotyping call rates less than 99% and those that significantly deviated from Hardy–Weinberg equilibrium ($P < 1 \times 10^{-6}$). We then removed individuals with excess heterozygosity. Individuals greater than three standard deviations from mean heterozygosity were removed from the study. We used EIGENSOFT smartpca (Patterson, Price & Reich, 2006) to generate the first 10 principal components, which were used to confirm self-identified ancestry (Figs. S1 and S2). After this quality control, we had 742,015 SNPs and 2,256 individuals total in the schizophrenia cohort and 721,050 SNPs and 1,019 individuals in the bipolar disorder cohort (Purcell et al., 2007).

From here, the filtered data from each cohort were uploaded to the University of Michigan Imputation Server for genotype imputation (Das et al., 2016). The genotypes for each cohort were imputed using Eagle version 2.3 for phasing and 1000 Genomes Phase 3 version 5 (1000G) as our reference panel (The 1000 Genomes Project Consortium, 2015). After this, we downloaded the imputed data from the Michigan Imputation Server and converted it to PLINK binary format. We then filtered the data by removing SNPs with imputation $r^2 < 0.8$ and minor allele frequency (MAF) < 0.01 . At this point, we were left with 12,780,487 and 12,799,548 SNPs in the schizophrenia and bipolar disorder cohorts, respectively. We explored Consortium on Asthma among African-ancestry Populations in the Americas (CAAPA) as an alternative reference panel for imputation, but 1000G imputed more SNPs meeting our filters while simultaneously imputing SNPs with MAFs identical to those imputed from CAAPA at our chosen imputation r^2 and MAF thresholds (Fig. S3) (Mathias et al., 2016).

Genome-wide association study

Using PLINK, we performed a logistic regression of the phenotype using the first 10 genotypic principal components as covariates to account for population structure. We used a significance threshold of $P < 5 \times 10^{-8}$ to identify significantly associated SNPs. Plots were generated from PLINK results using the web-based tool LocusZoom (Pruim et al., 2011).

PrediXcan

We performed the gene-based association test PrediXcan on both phenotypes, schizophrenia and bipolar disorder, in this study. PrediXcan functions by predicting an individual's genetically regulated gene expression levels using tissue-dependent prediction models trained using reference transcriptome data (Gamazon et al., 2015). For our experiments we tested each phenotype across 55 prediction models. Forty-eight of these models were trained on 48 tissues in GTEx version 7 (Barbeira et al., 2018). Six of these models were generated from monocyte transcriptomes of individuals in the Multi-Ethnic Study of Atherosclerosis (MESA) cohort. The MESA models, the most diverse set of published predictors to date, were built from genotypes and transcriptomes of self-identified African American, Hispanic, and European individuals (Mogil et al., 2018). These models can be found at <http://predictdb.org/>. We also used a model built from dorsolateral prefrontal cortex (DLPFC) data from the CommonMind Consortium

(Huckins et al., 2019). To impute the gene expression levels, the PLINK binary files from each cohort had to be converted to PrediXcan dosage files. To do this, we used the conversion script provided at <https://github.com/hakyimlab/PrediXcan/tree/master/Software>. After predicting a genetically regulated level of expression, we tested each expression level for association with the phenotype of interest. Since PrediXcan does not have a flag for performing a logistic regression with covariates, we performed a logistic regression of the phenotype with the first 10 principal components to generate a residual phenotype in order to account for population structure. We then performed a linear regression with the residual phenotype and gene expression level for each gene.

Following the PrediXcan association tests, we adjusted for multiple testing using the adaptive shrinkage approach implemented in the R package *ashr* (Stephens, 2017). Using this package, we calculated the local false sign rate (lfsr) for each test, which is similar to traditional false discovery rate approaches, but takes into account both the effect sizes and standard errors of each gene-tissue pair ($n = 248,605$). In addition, this empirical Bayes approach uses the assumption that the distribution of actual effects is unimodal with the mode at 0. We set our significance threshold for gene-tissue pairs at $\text{lfsr} < 0.2$.

Due to the dearth of African American neuropsychiatric cohorts, replication could not be completed in an independent African American cohort. To validate our results, we compared our findings to the association results of a meta-analysis of the PGC GWAS summary statistics completed using S-PrediXcan (Barbeira et al., 2018; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014; Psychiatric GWAS Consortium Bipolar Disorder Working Group, 2011). We performed S-PrediXcan in both phenotypes using predictors from GTEx. The Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014) study originally included 36,989 cases and 113,075 controls composed of about 96.5% European and 3.5% Asian individuals. The Psychiatric GWAS Consortium Bipolar Disorder Working Group (2011) study included 7,481 cases and 9,250 controls of primarily European and Asian descent.

MultiXcan

Following imputation of gene expression levels, we performed MultiXcan (Barbeira et al., 2019), a gene-based association test that combines information across multiple tissues while taking their correlation into account. Using the predicted expression levels in 48 tissues across GTEx, we performed MultiXcan on both of the disease phenotypes.

RESULTS

Schizophrenia gene-based association study

To better understand the genetic architecture of schizophrenia in African Americans, we performed transcriptome-wide association studies using prediction models built in 55 tissues. In the GAIN cohort of 2,256 individuals (969 controls and 1,287 cases), we predicted gene expression across 48 tissues in GTEx, six models built from monocytes across MESA, and DLPFC from CommonMind (Barbeira et al., 2018; Wheeler et al., 2016; Mogil et al., 2018; Huckins et al., 2019).

Table 2 Top PrediXcan results for schizophrenia in African Americans (GAIN cohort) by local false sign rate (lfsr). *PRMT7* makes up five of the top eight associated gene-tissue pairs by lfsr. BP are reported as transcription start site for the respective genes.

Gene	Beta	<i>t</i>	<i>P</i>	se(beta)	Tissue (predictor)	lfsr	CHR	BP
<i>PRMT7</i>	0.10	4.56	5.49×10^{-6}	0.022	Heart_Atrial_Appendage	0.119	16	68392457
<i>PRMT7</i>	0.09	4.22	2.50×10^{-5}	0.020	Cells_Transformed_fibroblasts	0.225	16	68392457
<i>PRMT7</i>	0.08	4.00	6.63×10^{-5}	0.019	Heart_Left_Ventricle	0.353	16	68392457
<i>PRMT7</i>	0.11	4.23	2.47×10^{-5}	0.025	Adrenal_Gland	0.365	16	68392457
<i>TBC1D2</i>	0.12	4.24	2.34×10^{-5}	0.028	Brain_Putamen_basal_ganglia	0.468	9	100961311
<i>NPC1</i>	0.07	3.83	1.29×10^{-4}	0.019	Thyroid	0.484	18	21086148
<i>EIF2S2P3</i>	0.08	3.81	1.41×10^{-4}	0.022	Brain_Amygdala	0.558	10	94428502
<i>PRMT7</i>	0.08	3.73	1.97×10^{-4}	0.020	Cells_EBV-transformed_lymphocytes	0.578	16	68392457

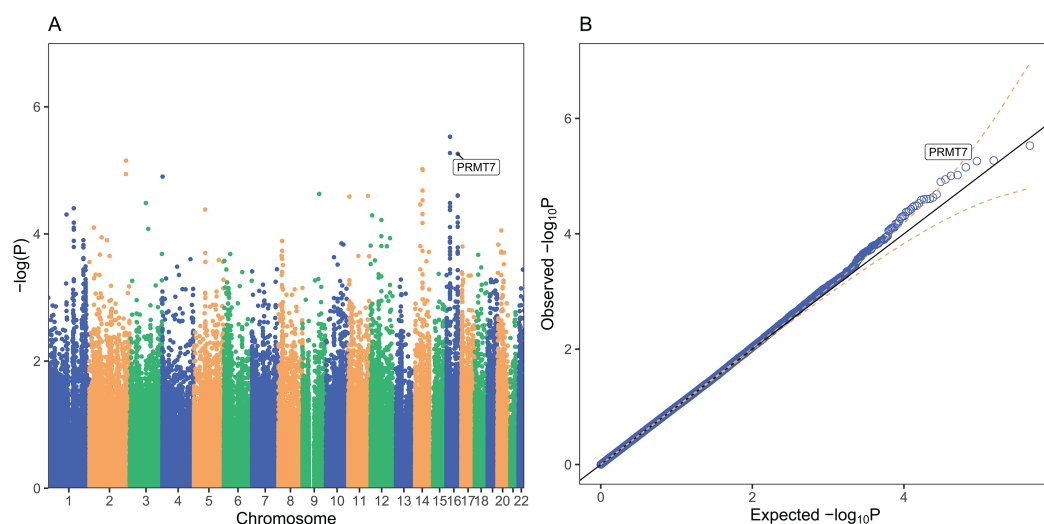


Figure 1 PrediXcan association results for schizophrenia in GAIN African Americans. Each point on the Manhattan (A) and Quantile-Quantile (B) plots represents one gene-tissue test for association with schizophrenia using GTEx version 7 gene expression prediction models. *PRMT7* expression in atrial appendage of the heart is labeled in both plots since it had the lowest lfsr of all tissues (lfsr = 0.119). Predicted *RP11-646C24.5* expression in pancreas and sigmoid colon associations are represented as the two points with lower *P*-values than *PRMT7*, respectively, but lfsr was greater than 0.8 for each association. Unlike more traditional false discovery rate approaches such as Benjamini-Hochberg, the gene with the lowest *P*-value may not be the gene with the lowest lfsr especially if the standard error of the effect size estimate is high (Stephens, 2017). [Full-size !\[\]\(b5d7dedcc48d5bfd2c56b334ed39e34f_img.jpg\) DOI: 10.7717/peerj.7778/fig-1](https://doi.org/10.7717/peerj.7778/fig-1)

PRMT7 was the most significantly associated gene with an lfsr of 0.119 and a *P*-value of 5.49×10^{-6} in the atrial appendage of the heart (Table 2; Figs. 1 and 2). Increased predicted expression of *PRMT7* associated with schizophrenia in 32 of 33 tissues in GTEx tested (Figs. 2 and 3). Effect sizes were also positive for *PRMT7* associations with schizophrenia in 42 of 42 tissues tested in our S-PrediXcan application to the PGC data (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014) (Fig. 3). Of the 42 tissues tested, 30 associations were statistically significant ($P < 0.0012$) after Bonferroni adjustment for the number of tissues tested and all 42 tissues had $P < 0.05$ (Table S1). We found no significant gene-tissue associations using the MESA or DLPFC models.

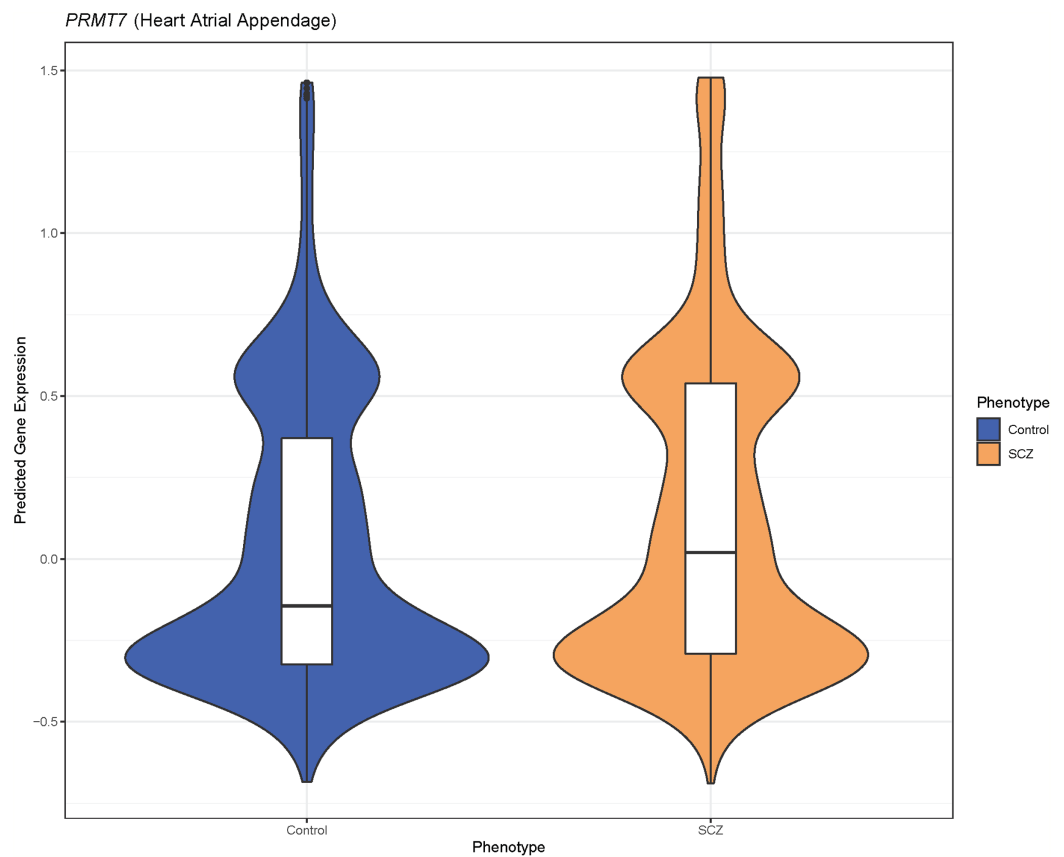


Figure 2 Predicted *PRMT7* expression is higher in schizophrenia cases than controls in GAIN. The violin plot represents the differences in density of predicted gene expression levels of *PRMT7* between cases (SCZ) and controls in heart atrial appendage from GTEx ($P = 5.49 \times 10^{-6}$).

Full-size  DOI: [10.7717/peerj.7778/fig-2](https://doi.org/10.7717/peerj.7778/fig-2)

While *PRMT7* in atrial appendage had the lowest lfsr across all models, *RP11-646C24.5* had a lower P -value (Fig. 1), but high lfsr in both pancreas (lfsr = 0.860) and sigmoid colon (lfsr = 0.851). Notably, the standard error in both of these tissues was over twice the size of that of *PRMT7*. Unlike more traditional false discovery rate approaches such as Benjamini–Hochberg, both effect size and standard error are used in an empirical Bayesian framework to calculate lfsr and thus the gene with the lowest P -value may not be the gene with the lowest lfsr (Stephens, 2017). We found no significant associations with the MultiXcan, cross-tissue model.

Bipolar disorder gene-based association study

To develop a better understanding of the genetic mechanisms governing bipolar disorder in African Americans, we performed PrediXcan in a cohort of 1,019 individuals (671 controls and 348 Cases). Similar to our gene-based association study of schizophrenia, we performed our tests across the same 55 gene expression prediction models in our bipolar disorder study.

In the GAIN cohort of 1,019 African American individuals, no genes were identified to be significantly associated with bipolar disorder. Increased predicted expression of

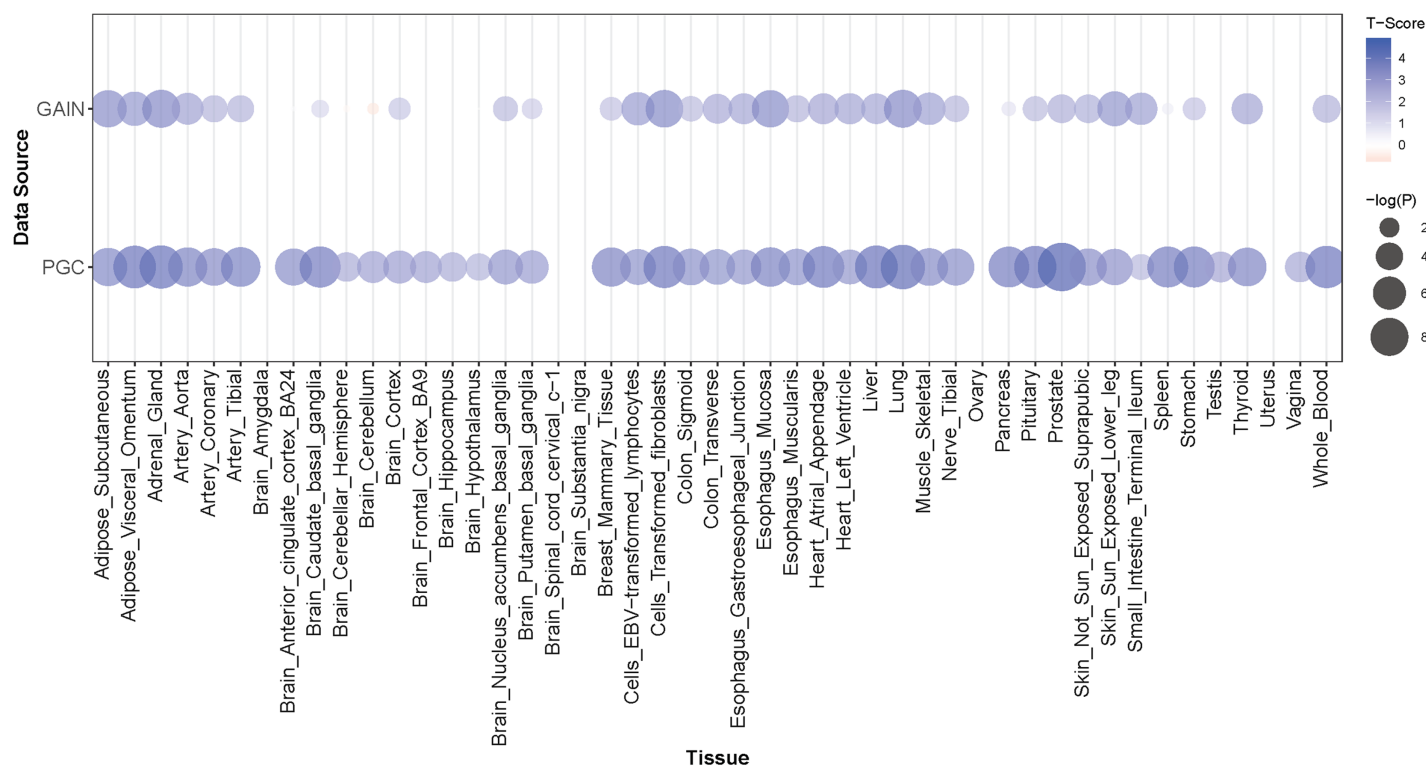


Figure 3 *PRMT7* PrediXcan discovery (GAIN) and replication (PGC) results across tissue models. In each bubble plot, the radius of the bubble is representative of the significance of *PRMT7* association with SCZ. The color of the bubble represents the test statistic with blue representing a positive direction of effect and red representing a negative direction of effect. [Full-size !\[\]\(5fd6ef84f97f42d7f8b34275f1b65312_img.jpg\) DOI: 10.7717/peerj.7778/fig-3](https://doi.org/10.7717/peerj.7778/fig-3)

GREM2 in testis was the most associated ($P = 2.20 \times 10^{-5}$) gene-tissue pair with bipolar disorder (Fig. 4). *KCNMB3* had the lowest *lfsr* at 0.919. We also found no significant associations with the MultiXcan, cross-tissue model.

Schizophrenia SNP-level association test

We performed a GWAS across greater than 12 million SNPs following imputation to help elucidate the role specific SNPs play in the genetics of schizophrenia in African Americans. We used the first 10 principal components as covariates for our logistic regression in order to adjust for population stratification in the cohort. In our SNP-level GWAS, we found no significantly associated SNPs; however, one of the most associated SNPs, rs8063446 ($P = 2.66 \times 10^{-6}$), is located at the *PRMT7* locus (Fig. 5) While not genome-wide significant, the most associated SNP in our study was rs112845369 ($P = 1.094 \times 10^{-6}$) on chromosome 15.

Bipolar disorder SNP-level association test

We also performed a logistic GWAS in over 12 million SNPs in an attempt to understand the role specific SNPs play in the genetics of bipolar disorder. We similarly used the first 10 principal components to adjust for population stratification in the bipolar disorder cohort. Similar to our findings in our schizophrenia GWAS, we identified no SNPs

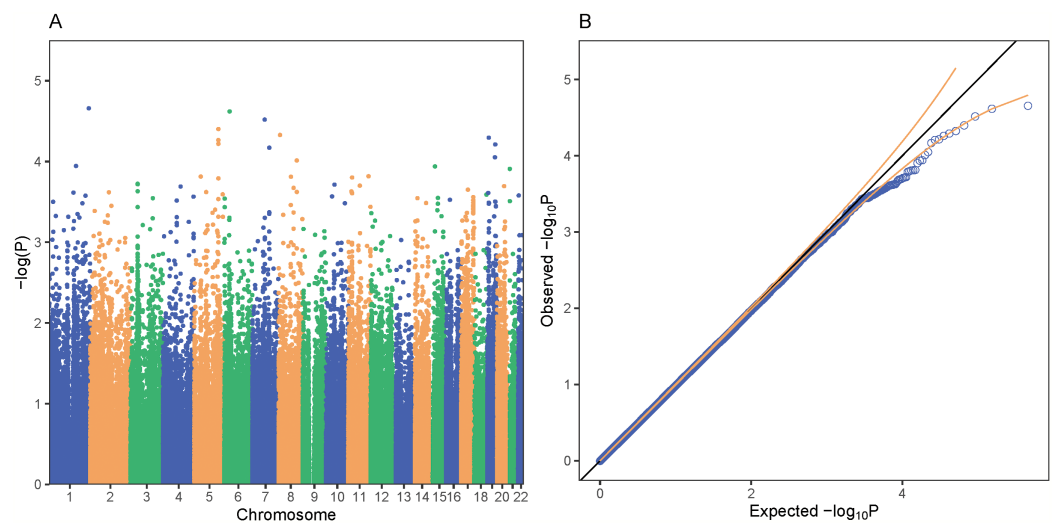


Figure 4 PrediXcan association results for bipolar disorder in GAIN. Each point on the Manhattan (A) and Quantile-Quantile (B) plots represents a gene-tissue association test for our study of bipolar disorder using GTEx models. *GREM2* on chromosome 1 was the gene most associated with bipolar disorder in our study. All of the gene associations tests across 48 tissues in GTEx are plotted in (A) and (B). A total of 95% confidence intervals depicted by gold dotted lines (B).

Full-size DOI: [10.7717/peerj.7778/fig-4](https://doi.org/10.7717/peerj.7778/fig-4)

significantly associated with bipolar disorder after adjusting for multiple tests. rs10168049 on chromosome 2 was the most associated SNP ($P = 1.04 \times 10^{-6}$). While not significantly associated with bipolar disorder, rs10168049 has a MAF of 0.465 in African 1000G populations compared to those of European and East Asian populations with minor allele frequencies of 0.042 and 0.097, respectively (Fig. 6) (Sherry, 2001; The 1000 Genomes Project Consortium, 2015). All PrediXcan and GWAS summary statistics for both diseases are available at <https://github.com/WheelerLab/Neuropsychiatric-Phenotypes>.

DISCUSSION

We performed gene-level (PrediXcan) and SNP-level association studies for schizophrenia and bipolar disorder in African Americans from GAIN (Suarez et al., 2006; Manolio et al., 2007). We used summary statistics from the predominantly European PGC to replicate our findings (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014; Psychiatric GWAS Consortium Bipolar Disorder Working Group, 2011).

A potential role for *PRMT7* in schizophrenia

PRMT7 was significantly associated with schizophrenia in our PrediXcan analyses and one of the most associated SNPs in our GWAS is just upstream of the gene (Figs. 1 and 5). Schizophrenia was associated with increased expression of *PRMT7* in 32 of 33 tissues in which it was predicted in the GAIN cohort. When S-PrediXcan was applied to the PGC summary stats, increased expression of *PRMT7* was associated with schizophrenia in all 42 tissues in which expression was predicted (Fig. 3). *PRMT7* made up five of the eight most associated gene-tissue pairs (Table 2). *PRMT7* has previously been associated with SBIDDS syndrome, an intellectual disability syndrome (Agolini et al., 2018). Its function in

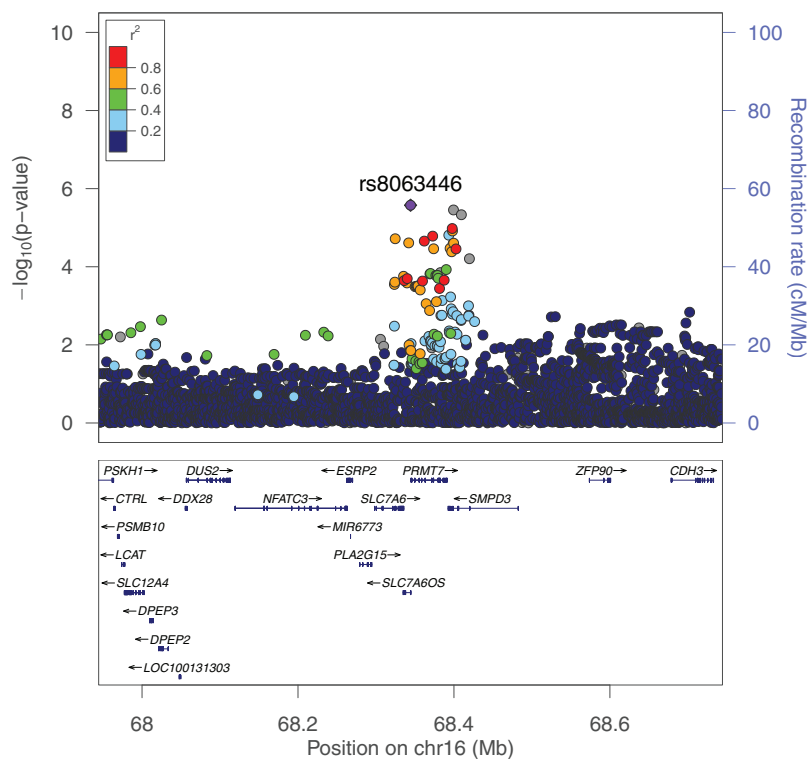


Figure 5 LocusZoom plot of the *PRMT7* locus in the GAIN GWAS for schizophrenia. rs8063446 is found in *SLC7A6OS* and 514 bp upstream of *PRMT7*. In our PrediXcan analyses, we found that increased predicted expression of *PRMT7* is associated with schizophrenia. rs8063446 is located in a linkage disequilibrium (LD) block with other SNPs associated with schizophrenia when plotted using 1000G AFR LD Population. [Full-size](#) DOI: 10.7717/peerj.7778/fig-5

the disorder remains unclear, but *PRMT7* has a functional role in neuronal differentiation, which could be a potential mechanism to explore further (Dhar et al., 2012).

While not found to be significantly associated with schizophrenia in brain tissues, the association of *PRMT7* in adrenal gland and other vascular system organs highlights the sharing of eQTLs across tissues. Recently, genes in both colon and adrenal gland were identified to be significantly associated with schizophrenia (Gamazon et al., 2019). Gamazon et al. highlight the opportunity to better understand the genetic mechanisms of neuropsychiatric diseases outside of the context of the central nervous system. A larger sample size would be needed to elucidate expression correlations and potential co-expression networks underlying African American neuropsychiatric traits.

Growing need for diversity in GWAS and genetic prediction models

In our GWAS of bipolar disorder, rs10168049 was the most significantly associated SNP. This SNP has not been implicated in previous studies, and its higher MAF in African populations compared to European and Asian populations demonstrates how GWAS in African ancestry populations may tag key loci missed in European-only studies (Fig. 6). In addition to not being identified in any published GWAS, rs10168049 is not present in any of the 55 prediction models we used to impute gene expression levels

chr2:122078406 G/A

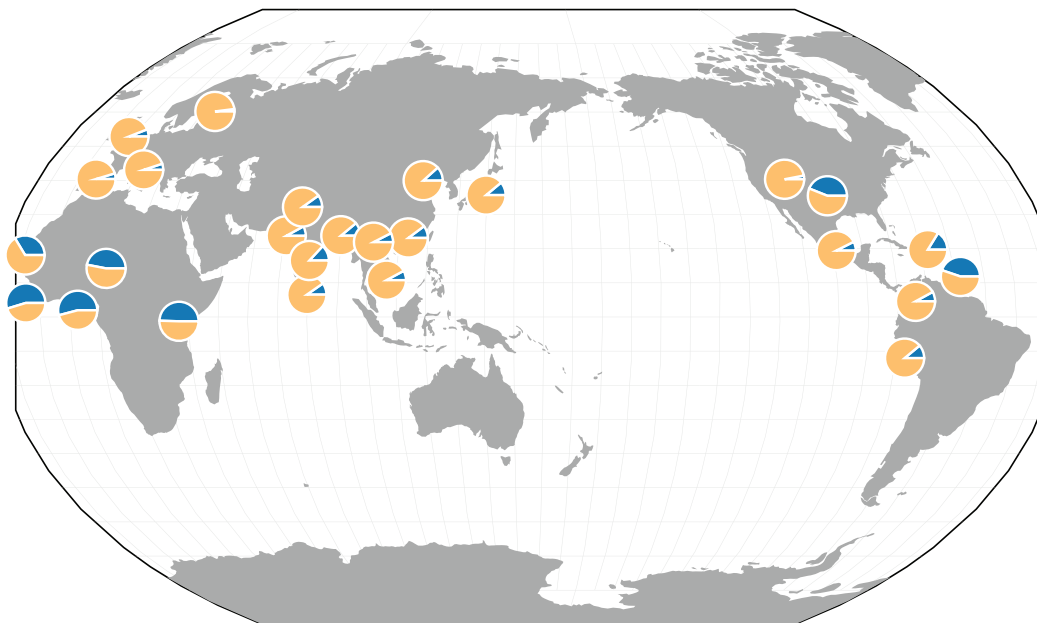


Figure 6 rs10168049 frequency across 1000G populations. Representative minor allele frequencies (MAFs) of rs10168049, which associated with bipolar disorder in GAIN ($P = 1.04 \times 10^{-6}$) in different populations from 1000G. The global MAF of this SNP in 1000G is 0.187; however, the MAF reaches up to 0.541 in the YRI (Yoruba people in Ibadan, Nigeria) population of 1000G. This figure was generated using the Geography of Genetic Variants Browser ([Marcus & Novembre, 2017](#)).

Full-size DOI: [10.7717/peerj.7778/fig-6](https://doi.org/10.7717/peerj.7778/fig-6)

([MacArthur et al., 2017](#)). As a result, this SNP did not contribute to predicted expression levels in our gene-based association tests. The association of this SNP with bipolar disorder needs to be replicated in larger studies; however, the lack of associations in other GWAS in the GWAS Catalog ([MacArthur et al., 2017](#)) may also be a result of ascertainment bias due to the field's focus on European populations ([Lachance & Tishkoff, 2014](#)).

Version 7 of the GTEx predictors we used contain data exclusively from individuals of European descent. These models are not optimal for predicting expression in African American cohorts ([Mogil et al., 2018](#); [Mikhaylova & Thornton, 2019](#)). While they offer power driven by sample size, they do not include models with African ancestry-specific alleles that might affect susceptibility to neuropsychiatric traits in this population.

The use of prediction models from monocytes in MESA offers advantages with respect to similar ancestry, but at the loss of nearly half of the sample size of many GTEx tissues. To ideally predict expression in African American cohorts, prediction models built in more tissues from African ancestry reference transcriptomes are needed. Moreover, future ancestry-specific models will not only increase accuracy of expression prediction, but they will also create opportunities for different methods, such as local ancestry mapping, to be applied to expression prediction by accounting for recent admixture within African American cohorts ([Zhong, Perera & Gamazon, 2019](#)).

CONCLUSION

Information from this study provides promising insight into the genetic architecture of gene expression underlying two neuropsychiatric disorders in African Americans. The results of our study were curbed by the small sample size of GAIN. With just over 2,100 individuals in our study of schizophrenia and 1,000 individuals in our study bipolar disorder, our findings were limited in power compared to larger European studies nearly two orders of magnitude greater in size (*Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014; Stahl et al., 2019*).

The size and diversity of our prediction models further hindered our ability to identify novel genes associated with these disorders. The MESA models, the most diverse of our predictors, were still limited in tissue type (monocytes only) and size at 233 African American individuals, 352 Hispanic individuals, and 578 European individuals (*Mogil et al., 2018*). The GTEx and CommonMind predictors we used, while generated from a larger sample size than the MESA African American cohort for many tissues, were made from exclusively European individuals (*GTEx Consortium, 2017; Huckins et al., 2019*). These key limitations highlight the need to increase the number of individuals from diverse populations in the study of neuropsychiatric genomics. To best characterize the molecular mechanisms that govern complex traits in diverse populations, diverse models, reference panels, and study subjects need to be included in genomics research.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was supported by the National Institutes of Health National Human Genome Research Institute Academic Research Enhancement Award R15 HG009569 (PI: Heather E. Wheeler), the Loyola University Chicago Carbon Undergraduate Research Fellowship (Peter N. Fiorica), and the Loyola University Chicago Mulcahy Scholarship (Peter N. Fiorica). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:

National Institutes of Health National Human Genome Research Institute Academic Research Enhancement Award: R15 HG009569.

Loyola University Chicago Carbon Undergraduate Research Fellowship.

Loyola University Chicago Mulcahy Scholarship.

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Peter N. Fiorica conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft.

- Heather E. Wheeler conceived and designed the experiments, analyzed the data, contributed reagents/materials/analysis tools, authored or reviewed drafts of the paper, approved the final draft.

Human Ethics

The following information was supplied relating to ethical approvals (i.e., approving body and any reference numbers):

This project was approved by the Loyola University Chicago Institutional Review Board (project number 2014).

Data Availability

The following information was supplied regarding data availability:

All datasets analyzed are available from the NCBI dbGaP at <https://www.ncbi.nlm.nih.gov/gap/> using the accession numbers [phs000021.v3.p2](https://www.ncbi.nlm.nih.gov/gap/acc.cgi?acc=phs000021.v3.p2), [phs000167.v1.p1](https://www.ncbi.nlm.nih.gov/gap/acc.cgi?acc=phs000167.v1.p1), and [phs000017.v3.p1](https://www.ncbi.nlm.nih.gov/gap/acc.cgi?acc=phs000017.v3.p1). Gene prediction models are available from PredictDB at <http://predictdb.org/>.

All scripts and summary statistics can be found at <https://github.com/WheelerLab/Neuropsychiatric-Phenotypes/>.

Funding support for the companion studies, Genome-Wide Association Study of Schizophrenia (GAIN) and Molecular Genetics of Schizophrenia—nonGAIN Sample (MGS_nonGAIN), was provided by Genomics Research Branch at NIMH (see below) and the genotyping and analysis of samples was provided through the Genetic Association Information Network (GAIN) and under the MGS U01s: MH79469 and MH79470. Assistance with data cleaning was provided by the National Center for Biotechnology Information. The MGS dataset(s) used for the analyses described in this manuscript were obtained from the database of Genotype and Phenotype (dbGaP) found at <https://www.ncbi.nlm.nih.gov/gap/> through dbGaP accession numbers [phs000021.v2.p1](https://www.ncbi.nlm.nih.gov/gap/acc.cgi?acc=phs000021.v2.p1) (GAIN) and [phs000167.v1.p1](https://www.ncbi.nlm.nih.gov/gap/acc.cgi?acc=phs000167.v1.p1) (nonGAIN). Samples and associated phenotype data for the MGS GWAS study were collected under the following grants: NIMH Schizophrenia Genetics Initiative U01s: MH46276 (CR Cloninger), MH46289 (C Kaufmann), and MH46318 (MT Tsuang); and MGS Part 1 (MGS1) and Part 2 (MGS2) R01s: MH67257 (NG Buccola), MH59588 (BJ Mowry), MH59571 (PV Gejman), MH59565 (Robert Freedman), MH59587 (F Amin), MH60870 (WF Byerley), MH59566 (DW Black), MH59586 (JM Silverman), MH61675 (DF Levinson), and MH60879 (CR Cloninger). Further details of collection sites, individuals, and institutions may be found in data supplemental Table 1 of Sanders et al. (2008; PMID: 18198266) and at the study dbGaP pages.

Funding support for the Whole Genome Association Study of Bipolar Disorder was provided by the National Institute of Mental Health (NIMH) and the genotyping of samples was provided through the Genetic Association Information Network (GAIN). The datasets used for the analyses described in this manuscript were obtained from the database of Genotypes and Phenotypes (dbGaP) found at <http://www.ncbi.nlm.nih.gov/gap/> through dbGaP accession number [phs000017.v3.p1](https://www.ncbi.nlm.nih.gov/gap/acc.cgi?acc=phs000017.v3.p1). Samples and associated phenotype data for the Collaborative Genomic Study of Bipolar Disorder were provided by The

NIMH Genetics Initiative for Bipolar Disorder. Data and biomaterials were collected in four projects that participated in NIMH Bipolar Disorder Genetics Initiative. From 1991 to 1998, the Principal Investigators and Co-Investigators were: Indiana University, Indianapolis, IN, U01 MH46282, John Nurnberger, M.D., Ph.D., Marvin Miller, M.D., and Elizabeth Bowman, M.D.; Washington University, St. Louis, MO, U01 MH46280, Theodore Reich, M.D., Allison Goate, Ph.D., and John Rice, Ph.D.; Johns Hopkins University, Baltimore, MD U01 MH46274, J. Raymond DePaulo, Jr, M.D., Sylvia Simpson, M.D., MPH, and Colin Stine, Ph.D.; NIMH Intramural Research Program, Clinical Neurogenetics Branch, Bethesda, MD, Elliot Gershon, M.D., Diane Kazuba, B.A., and Elizabeth Maxwell, M.S.W. Data and biomaterials were collected as part of 10 projects that participated in the NIMH Bipolar Disorder Genetics Initiative. From 1999 to 2003, the Principal Investigators and Co-Investigators were: Indiana University, Indianapolis, IN, R01 MH59545, John Nurnberger, M.D., Ph.D., Marvin J. Miller, M.D., Elizabeth S. Bowman, M.D., N. Leela Rau, M.D., P. Ryan Moe, M.D., Nalini Samavedy, M.D., Rif El-Mallakh, M.D. (at University of Louisville), Hussein Manji, M.D. (at Wayne State University), Debra A. Glitz, M.D. (at Wayne State University), Eric T. Meyer, M.S., Carrie Smiley, R.N., Tatiana Foroud, Ph.D., Leah Flury, M.S., Danielle M. Dick, Ph.D., Howard Edenberg, Ph.D.; Washington University, St. Louis, MO, R01 MH059534, John Rice, Ph.D., Theodore Reich, M.D., Allison Goate, Ph.D., Laura Bierut, M.D.; Johns Hopkins University, Baltimore, MD, R01 MH59533, Melvin McInnis M.D., J. Raymond DePaulo, Jr, M.D., Dean F. MacKinnon, M.D., Francis M. Mondimore, M.D., James B. Potash, M.D., Peter P. Zandi, Ph.D., Dimitrios Avramopoulos, and Jennifer Payne; University of Pennsylvania, PA, R01 MH59553, Wade Berrettini M.D., Ph.D.; University of California at Irvine, CA, R01 MH60068, William Byerley M.D., and Mark Vawter M.D.; University of Iowa, IA, R01 MH059548, William Coryell M.D., and Raymond Crowe M.D.; University of Chicago, IL, R01 MH59535, Elliot Gershon, M.D., Judith Badner Ph.D., Francis McMahon M.D., Chunyu Liu Ph.D., Alan Sanders M.D., Maria Caserta, Steven Dinwiddie M.D., Tu Nguyen, Donna Harakal; University of California at San Diego, CA, R01 MH59567, John Kelsoe, M.D., Rebecca McKinney, B.A.; Rush University, IL, R01 MH059556, William Scheftner M.D., Howard M. Kravitz, D.O., M.P.H., Diana Marta, B.S., Annette Vaughn-Brown, MSN, RN, and Laurie Bederow, MA; NIMH Intramural Research Program, Bethesda, MD, 1Z01MH002810-01, Francis J. McMahon, M.D., Layla Kassem, PsyD, Sevilla Detera-Wadleigh, Ph.D, Lisa Austin, Ph.D, Dennis L. Murphy, M.D.

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.7778#supplemental-information>.

REFERENCES

- Agolini E, Dentici ML, Bellacchio E, Alesi V, Radio FC, Torella A, Musacchia F, Tartaglia M, Dallapiccola B, Nigro V, Digilio MC, Novelli A. 2018. Expanding the clinical and molecular spectrum of PRMT7 mutations: 3 additional patients and review. *Clinical Genetics* 93(3):675–681 DOI 10.1111/cge.13137.

- Allardyce J, Leonenko G, Hamshere M, Pardiñas AF, Forty L, Knott S, Gordon-Smith K, Porteous DJ, Haywood C, Di Florio A, Jones L, McIntosh AM, Owen MJ, Holmans P, Walters JTR, Craddock N, Jones I, O'Donovan MC, Escott-Price V. 2018. Association between schizophrenia-related polygenic liability and the occurrence and level of mood-incongruent psychotic symptoms in bipolar disorder. *JAMA Psychiatry* 75(1):28–35 DOI 10.1001/jamapsychiatry.2017.3485.
- Barbeira AN, Dickinson SP, Bonazzola R, Zheng J, Wheeler HE, Torres JM, Torstenson ES, Shah KP, Garcia T, Edwards TL, Stahl EA, Huckins LM, Consortium G, Nicolae DL, Cox NJ, Im HK. 2018. Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nature Communications* 9(1):1825 DOI 10.1038/s41467-018-03621-1.
- Barbeira AN, Pividori MD, Zheng J, Wheeler HE, Nicolae DL, Im HK. 2019. Integrating predicted transcriptome from multiple tissues improves association detection. *PLOS Genetics* 15(1):e1007889 DOI 10.1371/journal.pgen.1007889.
- Bhalala OG, Nath AP, UK Brain Expression Consortium, Inouye M, Sibley CR. 2018. Identification of expression quantitative trait loci associated with schizophrenia and affective disorders in normal brain tissue. *PLOS Genetics* 14(8):e1007607 DOI 10.1371/journal.pgen.1007607.
- Das S, Forer L, Schön herr S, Sidore C, Locke AE, Kwong A, Vrieze SI, Chew EY, Levy S, McGue M, Schlessinger D, Stambolian D, Loh P-R, Iacono WG, Swaroop A, Scott LJ, Cucca F, Kronenberg F, Boehnke M, Abecasis GR, Fuchsberger C. 2016. Next-generation genotype imputation service and methods. *Nature Genetics* 48(10):1284–1287 DOI 10.1038/ng.3656.
- De Candia TR, Lee SH, Yang J, Browning BL, Gejman PV, Levinson DF, Mowry BJ, Hewitt JK, Goddard ME, O'Donovan MC, Purcell SM, Posthuma D, Visscher PM, Wray NR, Keller MC. 2013. Additive genetic variation in schizophrenia risk is shared by populations of African and European descent. *American Journal of Human Genetics* 93(3):463–470 DOI 10.1016/j.ajhg.2013.07.007.
- Dhar SS, Lee SH, Kan PY, Voigt P, Ma L, Shi X, Reinberg D, Lee MG. 2012. Trans-tail regulation of MLL₄-catalyzed H₃K₄ methylation by H₄R₃ symmetric dimethylation is mediated by a tandem PHD of MLL₄. *Genes & Development* 26(24):2749–2762 DOI 10.1101/gad.203356.112.
- Gamazon ER, Wheeler HE, Shah KP, Mozaffari SV, Aquino-Michaels K, Carroll RJ, Eyster AE, Denny JC, Nicolae DL, Cox NJ, Im HK. 2015. A gene-based association method for mapping traits using reference transcriptome data. *Nature Genetics* 47(9):1091–1098 DOI 10.1038/ng.3367.
- Gamazon ER, Zwinderman AH, Cox NJ, Denys D, Derks EM. 2019. Multi-tissue transcriptome analyses identify genetic mechanisms underlying neuropsychiatric traits. *Nature Genetics* 51(6):933–940 DOI 10.1038/s41588-019-0409-8.
- GTEx Consortium. 2017. Genetic effects on gene expression across human tissues. *Nature* 550(7675):204–213 DOI 10.1038/nature24277.
- Huckins LM, Dobbyn A, Ruderfer DM, Hoffman G, Wang W, Pardinas AF, Rajagopal VM, Als TD, Hoang HT, Girdhar K, Boocock J, Roussos P, Fromer M, Kramer R, Domenici E, Gamazon E, Purcell SM, Demontis D, Borglum AD, Walters J, O'Donovan M, Sullivan P, Owen M, Devlin B, Sieberts SK, Cox N, Im HK, Sklar P, Stahl EA, CommonMind Consortium, the Schizophrenia Working Group of the Psychiatric Genomics Consortium, iPSYCH-GEMS Schizophrenia Working Group. 2019. Gene expression imputation across multiple brain regions reveals schizophrenia risk throughout development. *Nature Genetics* 51:659–674 DOI 10.1038/s41588-019-0364-4.

- Ikeda M, Saito T, Kondo K, Iwata N. 2017a.** Genome-wide association studies of bipolar disorder: a systematic review of recent findings and their clinical implications. *Psychiatry and Clinical Neurosciences* 72(2):52–63 DOI 10.1111/pcn.12611.
- Ikeda M, Takahashi A, Kamatani Y, Okahisa Y, Kunugi H, Mori N, Sasaki T, Ohmori T, Okamoto Y, Kawasaki H, Shimodera S, Kato T, Yoneda H, Yoshimura R, Iyo M, Matsuda K, Akiyama M, Ashikawa K, Kashiwase K, Tokunaga K, Kondo K, Saito T, Shimasaki A, Kawase K, Kitajima T, Matsuo K, Itokawa M, Someya T, Inada T, Hashimoto R, Inoue T, Akiyama K, Tani H, Arai H, Kanba S, Ozaki N, Kusumi I, Yoshikawa T, Kubo M, Iwata N. 2017b.** A genome-wide association study identifies two novel susceptibility loci and trans population polygenicity associated with bipolar disorder. *Molecular Psychiatry* 23(3):639–647 DOI 10.1038/mp.2016.259.
- Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, Natarajan P, Lander ES, Lubitz SA, Ellinor PT, Kathiresan S. 2018.** Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nature Genetics* 50(9):1219–1224 DOI 10.1038/s41588-018-0183-z.
- Lachance J, Tishkoff SA. 2014.** SNP ascertainment bias in population genetic analyses: why it is important, and how to correct it. *BioEssays* 35(9):780–786 DOI 10.1002/bies.201300014.
- Li Z, Chen J, Yu H, He L, Xu Y, Zhang D, Yi Q, Li C, Li X, Shen J, Song Z, Ji W, Wang M, Zhou J, Chen B, Liu Y, Wang J, Wang P, Yang P, Wang Q, Feng G, Liu B, Sun W, Li B, He G, Li W, Wan C, Xu Q, Li W, Wen Z, Liu K, Huang F, Li J, Ripke S, Yue W, Sullivan PF, O'Donovan MC, Shi Y. 2017.** Genome-wide association analysis identifies 30 new susceptibility loci for schizophrenia. *Nature Genetics* 49(11):1576–1583 DOI 10.1038/ng.3973.
- MacArthur J, Bowler E, Cerezo M, Gil L, Hall P, Hastings E, Junkins H, McMahon A, Milano A, Morales J, Pendlington ZM, Welter D, Burdett T, Hindorf L, Flicek P, Cunningham F, Parkinson H. 2017.** The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Research* 45(D1):D896–D901 DOI 10.1093/nar/gkw1133.
- Manolio TA, Rodriguez LL, Brooks L, Abecasis G, Ballinger D, Daly M, Donnelly P, Faraone SV, Frazer K, Gabriel S, Gejman P, Guttmacher A, Harris EL, Insel T, Kelsoe JR, Lander E, McCowin N, Mailman MD, Nabel E, Ostell J, Pugh E, Sherry S, Sullivan PF, Thompson JF, Warram J, Wholley D, Milos PM, Collins FS. 2007.** New models of collaboration in genome-wide association studies: the Genetic Association Information Network. *Nature Genetics* 39(9):1045–1051 DOI 10.1038/ng2127.
- Manrai AK, Funke BH, Rehm HL, Olesen MS, Maron BA, Szolovits P, Margulies DM, Loscalzo J, Kohane IS. 2016.** Genetic misdiagnoses and the potential for health disparities. *New England Journal of Medicine* 375(7):655–665 DOI 10.1056/nejmsa1507092.
- Marcus JH, Novembre J. 2017.** Visualizing the geography of genetic variants. *Bioinformatics* 33(4):594–595 DOI 10.1093/bioinformatics/btw643.
- Martin AR, Gignoux CR, Walters RK, Wojcik GL, Neale BM, Gravel S, Daly MJ, Bustamante CD, Kenny EE. 2017.** Human demographic history impacts genetic risk prediction across diverse populations. *American Journal of Human Genetics* 100(4):635–649 DOI 10.1016/j.ajhg.2017.03.004.
- Mathias RA, Taub MA, Gignoux CR, Fu W, Musharoff S, O'Connor TD, Vergara C, Torgerson DG, Pino-Yanes M, Shringarpure SS, Huang L, Rafaels N, Boorgula MP, Johnston HR, Ortega VE, Levin AM, Song W, Torres R, Padhukasahasram B, Eng C, Mejia-Mejia D-A, Ferguson T, Qin ZS, Scott AF, Yazdanbakhsh M, Wilson JG, Marrugo J, Lange LA, Kumar R, Avila PC, Williams LK, Watson H, Ware LB, Olopade C, Olopade O, Oliveira R, Ober C, Nicolae DL, Meyers D, Mayorga A, Knight-Madden J, Hartert T,**

- Hansel NN, Foreman MG, Ford JG, Faruque MU, Dunston GM, Caraballo L, Burchard EG, Bleecker E, Araujo MI, Herrera-Paz EF, Gietzen K, Grus WE, Bamshad M, Bustamante CD, Kenny EE, Hernandez RD, Beaty TH, Ruczinski I, Akey J, Barnes KC. 2016. A continuum of admixture in the Western Hemisphere revealed by the African Diaspora genome. *Nature Communications* 7(1):12522 DOI 10.1038/ncomms12522.
- Mikhaylova AV, Thornton TA. 2019. Accuracy of gene expression prediction from genotype data with PrediXcan varies across and within continental populations. *Frontiers in Genetics* 10:261 DOI 10.3389/fgene.2019.00261.
- Mogil LS, Andaleon A, Badalamenti A, Dickinson SP, Guo X, Rotter JI, Johnson WC, Im HK, Liu Y, Wheeler HE. 2018. Genetic architecture of gene expression traits across diverse populations. *PLOS Genetics* 14(8):e1007586 DOI 10.1371/journal.pgen.1007586.
- Patterson N, Price AL, Reich D. 2006. Population structure and eigenanalysis. *PLOS Genetics* 2(12):e190 DOI 10.1371/journal.pgen.0020190.
- Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer CJ, Frishman D. 2011. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* 26(18):2336–2337 DOI 10.1093/bioinformatics/btq419.
- Psychiatric GWAS Consortium Bipolar Disorder Working Group. 2011. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nature Genetics* 43(10):977–983 DOI 10.1038/ng.943.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, De Bakker PIW, Daly MJ, Sham PC. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* 81(3):559–575 DOI 10.1086/519795.
- Quansah E, McGregor NW. 2018. Towards diversity in genomics: the emergence of neurogenomics in Africa? *Genomics* 110(1):1–9 DOI 10.1016/j.ygeno.2017.07.009.
- Schizophrenia Working Group of the Psychiatric Genomics Consortium. 2014. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 511(7510):421–427 DOI 10.1038/nature13595.
- Sherry ST. 2001. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Research* 29(1):308–311 DOI 10.1093/nar/29.1.308.
- Sirugo G, Williams SM, Tishkoff SA. 2019. The missing diversity in human genetic studies. *Cell* 177(1):26–31 DOI 10.1016/j.cell.2019.02.048.
- Smith EN, Bloss CS, Badner JA, Barrett T, Belmonte PL, Berrettini W, Byerley W, Coryell W, Craig D, Edenberg HJ, Eskin E, Foroud T, Gershon E, Greenwood TA, Hipolito M, Koller DL, Lawson WB, Liu C, Lohoff F, McInnis MG, McMahon FJ, Mirel DB, Murray SS, Nievergelt C, Nurnberger J, Nwulia EA, Paschall J, Potash JB, Rice J, Schulze TG, Scheftner W, Panganiban C, Zaitlen N, Zandi PP, Zöllner S, Schork NJ, Kelsoe JR. 2009. Genome-wide association study of bipolar disorder in European American and African American individuals. *Molecular Psychiatry* 14(8):755–763 DOI 10.1038/mp.2009.43.
- Stahl EA, Breen G, Forstner AJ, McQuillan A, Ripke S, Trubetskoy V, Mattheisen M, Wang Y, Coleman JRI, Gaspar HA, de Leeuw CA, Steinberg S, Pavlides JMW, Trzaskowski M, Byrne EM, Pers TH, Holmans PA, Richards AL, Abbott L, Agerbo E, Akil H, Albani D, Alliey-Rodriguez N, Als TD, Anjorin A, Antilla V, Awasthi S, Badner JA, Bækvad-Hansen M, Barchas JD, Bass N, Bauer M, Belliveau R, Bergen SE, Pedersen CB, Bøen E, Boks MP, Boocock J, Budde M, Bunney W, Burmeister M, Bybjerg-Grauholm J, Byerley W, Casas M, Cerrato F, Cervantes P, Chambert K, Charney AW, Chen D, Churchhouse C, Clarke T.-K, Coryell W, Craig DW, Cruceanu C, Curtis D, Czerski PM,

- Dale AM, de Jong S, Degenhardt F, Del-Favero J, DePaulo JR, Djurovic S, Dobbyn AL, Dumont A, Elvsåshagen T, Escott-Price V, Fan CC, Fischer SB, Flickinger M, Foroud TM, Forty L, Frank J, Fraser C, Freimer NB, Frisé L, Gade K, Gage D, Garnham J, Giambartolomei C, Pedersen MG, Goldstein J, Gordon SD, Gordon-Smith K, Green EK, Green MJ, Greenwood TA, Grove J, Guan W, Guzman-Parra J, Hamshere ML, Hautzinger M, Heilbronner U, Herms S, Hipolito M, Hoffmann P, Holland D, Huckins L, Jamain S, Johnson JS, Juréus A, Kandaswamy R, Karlsson R, Kennedy JL, Kittel-Schneider S, Knowles JA, Kogevinas M, Koller AC, Kupka R, Lavebratt C, Lawrence J, Lawson WB, Leber M, Lee PH, Levy SE, Li JZ, Liu C, Lucae S, Maaser A, MacIntyre DJ, Mahon PB, Maier W, Martinsson L, McCarroll S, McGuffin P, McInnis MG, McKay JD, Medeiros H, Medland SE, Meng F, Milani L, Montgomery GW, Morris DW, Mühleisen TW, Mullins N, Nguyen H, Nievergelt CM, Adolfsson AN, Nwulia EA, O'Donovan C, Loohuis LMO, Ori APS, Oruc L, Ösby U, Perlis RH, Perry A, Pfennig A, Potash JB, Purcell SM, Regeer EJ, Reif A, Reinbold CS, Rice JP, Rivas F, Rivera M, Roussos P, Ruderfer DM, Ryu E, Sánchez-Mora C, Schatzberg AF, Scheftner WA, Schork NJ, Shannon Weickert C, Shekhtman T, Shilling PD, Sigurdsson E, Slaney C, Smeland OB, Sobell JL, Söholm Hansen C, Spijker AT, St Clair D, Steffens M, Strauss JS, Streit F, Strohmaier J, Szlinger S, Thompson RC, Thorgeirsson TE, Treutlein J, Vedder H, Wang W, Watson SJ, Weickert TW, Witt SH, Xi S, Xu W, Young AH, Zandi P, Zhang P, Zöllner S, Adolfsson R, Agartz I, Alda M, Backlund L, Baune BT, Bellivier F, Berrettini WH, Biernacka JM, Blackwood DHR, Boehnke M, et al. 2019. Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nature Genetics* 51(5):793–803 DOI 10.1038/s41588-019-0397-8.
- Stephens M. 2017. False discovery rates: a new deal. *Biostatistics* 18(2):275–294 DOI 10.1093/biostatistics/kxw041.
- Suarez BK, Duan J, Sanders AR, Hinrichs AL, Jin CH, Hou C, Buccola NG, Hale N, Weilbaecher AN, Nertney DA, Olincy A, Green S, Schaffer AW, Smith CJ, Hannah DE, Rice JP, Cox NJ, Martinez M, Mowry BJ, Amin F, Silverman JM, Black DW, Byerley WF, Crowe RR, Freedman R, Cloninger CR, Levinson DF, Gejman PV. 2006. Genomewide linkage scan of 409 European-ancestry and African American families with schizophrenia: suggestive evidence of linkage at 8p23.3-p21.2 and 11p13.1-q14.1 in the combined sample. *American Journal of Human Genetics* 78(2):315–333 DOI 10.1086/500272.
- The 1000 Genomes Project Consortium. 2015. A global reference for human genetic variation. *Nature* 526(7571):68–74 DOI 10.1038/nature15393.
- The International Schizophrenia Consortium. 2009. Common polygenic variation contributes to risk of schizophrenia that overlaps with bipolar disorder. *Nature* 460(7256):748–752 DOI 10.1038/nature08185.
- Wheeler HE, Shah KP, Brenner J, Garcia T, Aquino-Michaels K, Cox NJ, Nicolae DL, Im HK, GTEx Consortium. 2016. Survey of the heritability and sparse architecture of gene expression traits across human tissues. *PLOS Genetics* 12(11):e1006423 DOI 10.1371/journal.pgen.1006423.
- Whiteford HA, Degenhardt L, Rehm J, Baxter AJ, Ferrari AJ, Erskine HE, Charlson FJ, Norman RE, Flaxman AD, Johns N, Burstein R, Murray CJ, Vos T. 2013. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *Lancet* 382(9904):1575–1586 DOI 10.1016/s0140-6736(13)61611-6.
- Zhong Y, Perera MA, Gamazon ER. 2019. On using local ancestry to characterize the genetic architecture of human phenotypes: genetic regulation of gene expression in multiethnic or admixed populations as a model. *American Journal of Human Genetics* 104:1–19 DOI 10.1016/j.ajhg.2019.04.009.