



Published in final edited form as:

*Mol Psychiatry*. 2013 March ; 18(3): 340–346. doi:10.1038/mp.2011.174.

## Enrichment of Cis-Regulatory Gene Expression SNPs and Methylation Quantitative Trait Loci Among Bipolar Disorder Susceptibility Variants

ER Gamazon<sup>1</sup>, JA Badner<sup>2</sup>, L Cheng<sup>3</sup>, C Zhang<sup>3</sup>, D Zhang<sup>4</sup>, NJ Cox<sup>1</sup>, ES Gershon<sup>2</sup>, JR Kelsoe<sup>5</sup>, TA Greenwood<sup>5</sup>, CM Nievergelt<sup>5</sup>, C Chen<sup>3</sup>, R McKinney<sup>5</sup>, PD Shilling<sup>5</sup>, NJ Schork<sup>6</sup>, EN Smith<sup>6</sup>, CS Bloss<sup>6</sup>, JI Nurnberger<sup>7</sup>, HJ Edenberg<sup>8</sup>, T Foroud<sup>8</sup>, DL Koller<sup>7</sup>, WA Scheftner<sup>9</sup>, W Coryell<sup>10</sup>, J Rice<sup>11</sup>, WB Lawson<sup>12</sup>, EA Nwulia<sup>12</sup>, M Hipolito<sup>12</sup>, W Byerley<sup>13</sup>, FJ McMahon<sup>14</sup>, TG Schulze<sup>14,15</sup>, WH Berrettini<sup>16</sup>, JB Potash<sup>17</sup>, PP Zandi<sup>17</sup>, PB Mahon<sup>17</sup>, MG McInnis<sup>18</sup>, S Zöllner<sup>18</sup>, P Zhang<sup>18</sup>, DW Craig<sup>19</sup>, S Szelinger<sup>19</sup>, TB Barrett<sup>20</sup>, and C Liu<sup>3</sup>

<sup>1</sup>Department of Medicine, University of Chicago, Chicago, IL, USA

<sup>2</sup>Department of Psychiatry, University of Chicago, Chicago, IL, USA

<sup>3</sup>Department of Psychiatry, Institute of Human Genetics, University of Illinois at Chicago, Chicago, IL, USA

<sup>4</sup>School of Medicine, University of Zhejiang, Hanzhou, Zhejiang, China

<sup>5</sup>Department of Psychiatry, University of California, San Diego, La Jolla, CA, USA

<sup>6</sup>Scripps Genomic Medicine and Scripps Translational Science Institute, La Jolla, CA, USA

<sup>7</sup>Department of Psychiatry, Indiana University School of Medicine, Indianapolis, IN, USA

<sup>8</sup>Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA

<sup>9</sup>Department of Psychiatry, Rush University, Chicago, IL, USA

<sup>10</sup>Department of Psychiatry, University of Iowa, Iowa City, IA, USA

<sup>11</sup>Division of Biostatistics, Washington University, St Louis, MO, USA

<sup>12</sup>Department of Psychiatry, Howard University, Washington, DC, USA

<sup>13</sup>Department of Psychiatry, University of California, San Francisco, CA, USA

Users may view, print, copy, download and text and data- mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: [http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

Correspondence: Dr C Liu, Department of Psychiatry, Institute of Human Genetics, 900 S Ashland Ave, Room 1006, University of Illinois at Chicago, Chicago, IL 60607-7173, USA. [cliu@psych.uic.edu](mailto:cliu@psych.uic.edu).

**Author Contributions:** E.R.G. and C.L. conceived and designed the study. E.R.G. and J.A.B. performed statistical analyses. E.R.G. and N.J.C. provided analysis tools. E.R.G. and C.L. wrote the manuscript. C.L. coordinated and supervised the study. L.C., C.Z., D.Z., E.S.G., J.R.K., T.A.G., C.M.N., R.M., P.D.S., N.J.S., E.N.S., C.S.B., J.I.N. Jr, H.J.E., T.F., D.L.K., W.A.S., W.C., J.R., W.B., F.J.M., T.G.S., W.H.B., J.B.P., P.P.Z., P.B.M., M.G.M., S.Z., P.Z., D.W.C., S.S., T.B.B., C.L. participated in patient diagnosis and sample collection. All authors edited and approved the manuscript.

<sup>14</sup>Genetic Basis of Mood and Anxiety Disorders Unit, National Institute of Mental Health Intramural Research Program, National Institutes of Health, US Department of Health and Human Services, Bethesda, MD, USA

<sup>15</sup>Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Mannheim, Germany

<sup>16</sup>Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, USA

<sup>17</sup>Department of Psychiatry, Johns Hopkins School of Medicine, Baltimore, MD, USA

<sup>18</sup>Department of Psychiatry, University of Michigan, Ann Arbor, MI, USA

<sup>19</sup>Neurogenomics Division, The Translational Genomics Research Institute, Phoenix, AZ, USA

<sup>20</sup>Department of Psychiatry, Portland VA Medical Center, Portland, OR, USA

## Abstract

We conducted a systematic study of top susceptibility variants from a genome-wide association (GWA) study of Bipolar Disorder to gain insight into the functional consequences of genetic variation influencing disease risk. We report here the results of experiments to explore the effects of these susceptibility variants on DNA methylation and mRNA expression in human cerebellum samples. Among the top susceptibility variants, we identified an enrichment of *cis* regulatory loci on mRNA expression (eQTLs), and a significant excess of quantitative trait loci for DNA CpG methylation, hereafter referred to as mQTLs. Bipolar Disorder susceptibility variants that *cis*-regulate both cerebellar expression and methylation of the same gene are a very small proportion of Bipolar Disorder susceptibility variants. This finding suggests that mQTLs and eQTLs provide orthogonal ways of functionally annotating genetic variation within the context of studies of pathophysiology in brain. No lymphocyte mQTL enrichment was found, suggesting that mQTL enrichment was specific to the cerebellum, in contrast to eQTLs. Separately, we found that using mQTL information to restrict the number of SNPs studied enhances our ability to detect a significant association. With this restriction *a priori* informed by the observed functional enrichment, we identified a significant association (*rs12618769*,  $P_{\text{bonferroni}} < 0.05$ ) from two other GWA studies (TGen+GAIN; 2,191 cases and 1,434 controls) of Bipolar Disorder, which we replicated in an independent GWA study (WTCCC). Collectively, our findings highlight the importance of integrating functional annotation of genetic variants for gene expression and DNA methylation to advance biological understanding of Bipolar Disorder.

---

Epigenetic phenomena <sup>1</sup>, including DNA methylation, histone modification, and RNA-induced gene silencing, are important mechanisms of transcriptional regulation. Rapid advances in the development of high-throughput assays facilitate correspondingly dramatic advances in elucidating how such processes mediate gene expression and other phenotypes <sup>2,3</sup>. Enhanced understanding of epigenetic changes promises to shed light on disease mechanisms and clarify the consequences of epigenetic disruptions on pathophysiology <sup>4</sup>. Single-gene disruptions of the epigenetic machinery have long been recognized to cause neurodevelopmental disorders; for example, mutations in the *MECP2* gene which encodes a protein that binds to methylated DNA are the cause of Rett syndrome <sup>5</sup>. Little, however, is known regarding the epigenetic component of common

complex diseases. Recent studies from our group <sup>6, 7</sup> and others <sup>8</sup> on the utility of expression quantitative trait loci (eQTLs) to improve our understanding of complex traits genetics and to enhance discovery from genome-wide association (GWA) studies prompted us to evaluate to what extent genetic and epigenetic regulation of gene expression may contribute to susceptibility to neuropsychiatric disorders.

Focusing on Bipolar Disorder, we were particularly interested in the question of whether mapping methylation levels and transcript levels to genomic loci as methylation quantitative trait loci (mQTLs) <sup>9</sup> and expression quantitative trait loci (eQTLs) <sup>10</sup> in brain respectively, may provide information on the functional consequences of otherwise anonymous susceptibility variants identified by GWA studies on Bipolar Disorder <sup>11-13</sup>, assist in the identification of novel susceptibility loci, and provide a framework for interpreting and prioritizing results from GWA studies of neuropsychiatric disorders. For an investigation of disease epigenetics, Bipolar Disorder provides a particularly striking example, as dynamic epigenetic fluctuations in gene regulation may contribute to fluctuation of disease course <sup>14, 15</sup>. In particular, a systematic genome-wide investigation of genetic control of DNA methylation allowed us to test such genomic loci for effect on disease risk.

## Materials and Methods

### Samples

DNA and RNA of human cerebellum samples from 153 individuals of European ancestry were obtained from the Stanley Medical Research Institute (SMRI). As previously described <sup>9</sup>, the specimens came from patients with psychiatric disorders and normal controls. Information on age, gender, ethnicity, brain pH, smoking and alcohol use, suicide status, postmortem interval (PMI) is available for the samples included in this study.

### Genotyping

DNA was extracted from cerebellar tissues obtained from SMRI. Sample genotyping was conducted using Genome-Wide Human SNP Array 5.0, at Translational Genomics Research Institute (Tgen). BRLMM-p (Affymetrix) was used as the genotype calling algorithm. SNPs with call rates less than 99% were excluded from the analysis. SNPs showing departure from Hardy-Weinberg equilibrium (HWE) were filtered ( $p < 0.001$ ). Of the remaining SNPs, only SNPs showing minor allele frequency (MAF) of at least 10% were carried forward for further analysis. These quality control filters yielded nearly 239,000 SNPs. We expanded on the earlier study <sup>9</sup> by performing imputation using MaCH v1.0 <sup>16, 17</sup> to increase the density of interrogated SNPs. Only SNPs with MAF of at least 10% and  $R^2$  (the squared correlation between imputed and true genotypes)  $> 0.3$ . 886,338 autosomal SNPs were analyzed for QTL mapping.

To test for the presence of population structure, principal components analysis (PCA) as implemented in Eigenstrat was performed. To identify novel relationships between the samples, pairwise identity by state analysis was done using PLINK.

### Microarray expression phenotyping

RNAs were extracted from the cerebellum cortex of 132 samples using the RNeasy Mini kit (Qiagen, Valencia, CA). The concentration and A260/A280 ratio were measured on the NanoDrop spectrophotometer. The ratio of 28S to 18S rRNA and RNA Integrity Number (RIN) were measured using an RNA LabChip kit on the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA). In order to avoid use of seriously degraded samples, only RNA samples with a RIN > 7 were used for the expression profiling. Affymetrix Human Gene 1.0 ST array was used for whole genome transcriptome profiling at the NIH Neuroscience Microarray Consortium at the facility at Yale University.

### DNA Methylation Phenotyping

DNA methylation was assayed using the Illumina Infinium HumanMethylation27 BeadChips (Illumina Inc., San Diego, CA) platform; the assay was performed at the Genomics Core Facility at Northwestern University. 27,578 CpG sites were probed, almost all of which (approximately 97.7%) are within 1500 bp of a transcription start site (TSS). Methylation level at each interrogated CpG site was defined as: methylated probe intensity / (sum of methylated and unmethylated probe intensities) Therefore, methylation level is a quantitative trait whose values range from 0 to 1.

### Expression and Methylation Data Preprocessing

Affymetrix Expression Console was used to process raw data. All probes, which contain common SNPs (MAF > 0.01) based on the 1000 Genome Project data and HapMap data, or could be mapped to multiple genomic regions, were excluded from the analysis. ComBat<sup>18</sup> was used to correct for batch effects within the methylation and expression array data. For later analysis of each technical replicate pair, the data were averaged for the replicated samples to obtain a single datum. In order to remove the effects of known and unknown covariates on the data, surrogate variable analysis (SVA)<sup>19</sup> was applied and the identified surrogate variables were regressed out. We have found that these methods perform well in removing the effects of batch and known covariates such as pH. To fit a normal distribution, quantile normalization was used for both expression and methylation residuals.

### Association QTL Analyses

Imputed genotype dosage data was analyzed for association with methylation and expression phenotypes using PLINK<sup>20</sup>. Linear regression of the phenotype with dosage of the minor allele for each SNP was performed. *Cis* regions were defined as 4 Mb of the probe site. This definition allowed us to compare our findings in cerebellum to our earlier observations in lymphocytes<sup>7, 21</sup>. *Trans* regions included the rest of the genome. The significance thresholds were based on the estimated number of probes for the *cis* and *trans*-analysis. For the *cis* analysis, the significance threshold was 0.01. For the *trans*-analysis, 8597 and 25834 probes were analyzed for the methylation and expression data respectively. The *trans*-analysis significance threshold was set to 0.05 divided by the number of probes.

## Simulation Analysis for Enrichment Test

To determine the level of (*cis*) mQTL (or eQTL or joint eQTL-mQTL) enrichment in the set of top susceptibility variants (defined in terms of p-value threshold) from a GWAS, we generated 1000 sets, randomly drawn without replacement from a control set (e.g., Affymetrix GeneChip 500K Mapping Array Set, the study platform, for the WTCCC GWAS), which match the minor allele frequency distribution of the SNP set of interest. For each random set, the (*cis*) mQTL (or eQTL or joint eQTL-mQTL) count was determined, yielding the distribution under the null. The observed (*cis*) mQTL (or eQTL or joint eQTL-mQTL) count enabled us to calculate an empirical p-value for the level of enrichment by considering the proportion of random sets with mQTL (or eQTL or joint eQTL-mQTL) count that matches or exceeds the observed count. We tested the dependence of our enrichment results on the definition of *cis* by doing these analyses when *cis* is defined as within 1 Mb of the gene.

## Genome-wide Association Studies

We utilized the results from several GWA studies of Bipolar Disorder. Our TGen sample set consists of 1,190 Bipolar Disorder cases from the Bipolar Genome Study (BiGS) and 401 controls<sup>12</sup>. The GAIN sample set, from an earlier study conducted under the GAIN initiative, consists of 1,001 Bipolar Disorder cases and 1,033 controls of European ancestry<sup>11</sup>. The WTCCC study of seven common diseases, including Bipolar Disorder, had an experimental design of 2000 cases for each disease and 3000 shared controls<sup>13</sup>.

## Results

### A catalog of mQTLs

We re-analyzed our published mQTL dataset<sup>9</sup> by including imputed genotype data (see Materials and Methods), and extended the *cis*-region to 4 Mb. Analogous to the traditional classification of eQTLs, we sought to classify methylation QTLs (mQTLs) into *cis* and *trans* regulators of methylation phenotypes. SNPs within the *cis*-region of a variably methylated CpG site were included in our analysis. We identified 5974 distinct genes with *cis* mQTLs at association  $p < 0.01$ , and 4143 distinct genes with *cis* mQTLs at a more stringent  $p < 0.001$ . Supplemental Figure 1 illustrates the genomic locations of the identified *cis* mQTLs. To further the investigation of these mQTLs, we have deposited this dataset into SCAN (<http://www.scandb.org>)<sup>6</sup>, a publicly available genomic resource containing initially the results of eQTL studies conducted in lymphoblastoid cell lines.

### Top Bipolar Disorder susceptibility SNPs from GWA studies are enriched for mQTLs

We sought to test the hypothesis that the top associations from the WTCCC Bipolar Disorder study are enriched for variants that affect methylation levels. This follows our earlier study on the use of eQTL enrichment analysis and eQTL-based annotation to enhance discovery of reproducible signals from GWA studies<sup>6</sup>. That initial study was performed on (transformed) lymphocytes. We thus wanted to test a secondary hypothesis, whether SNPs that exert an influence on methylation levels in brain, the presumed tissue of relevance for Bipolar Disorder, are enriched in SNPs showing strong associations with Bipolar Disorder.

Figure 1 illustrates the result of our enrichment analysis applied to the WTCCC study of Bipolar Disorder. First, we took the top SNP associations (defined as  $p < 0.001$ ;  $n=935$ ) and asked whether there was an excess of *cis* mQTLs relative to null expectation. For the definition of mQTLs, we used the threshold of  $p < 0.01$ ; a subsequent analysis using a more stringent  $p < 0.001$  yielded similar results (so that the enrichment results appear robust to the definition of mQTL). To generate the null distribution, we conducted simulation studies, as previously described<sup>6</sup>. See Materials and Methods for details of the simulation procedure. At the given mQTL threshold, we identified 132 such mQTLs among the top Bipolar Disorder associated SNPs (shown in Figure 1 as a black dot). From the simulated datasets ( $N=1000$ ) each comprising of minor allele frequency-matched SNPs, the expected *cis* mQTL count was found to be 108.9 with standard deviation of 9.7. The enrichment  $p$  value is 0.01.

We then tested for *cis* mQTL enrichment for all SNP-disease associations satisfying  $p < 0.0001$ . Such a test would establish robustness of our observation to the definition of “top SNP associations” used in the analysis. We found that the *cis* mQTL enrichment held robustly ( $p = 0.01$ ) at the more stringent SNP-disease association cutoff.

We also tested for enrichment of *trans* mQTLs among the top associated SNPs. At this initial mQTL threshold, there was no evidence for any significant excess. In 1000 simulated sets, the expected count of *trans* mQTLs was found to be 201.8 ( $SD=12.4$ ); the observed count, however, was 199.

In the WTCCC study, *cis* mQTLs comprised 14% of the most significant associations ( $p < 0.001$ ). We asked whether a similar level of *cis* mQTL enrichment could be observed among the top SNPs in the GAIN and TGen studies of Bipolar Disorder<sup>11, 12</sup>. At the given association threshold, a similar percentage, 14.4%, of the top SNPs from the GAIN study were found to be *cis* mQTLs. Consistent with this trend, the top variants from the TGen study show a significant overlap with local mQTLs (20%). Furthermore, as in the WTCCC study, we observed no enrichment of *trans* mQTLs (enrichment  $p > 0.50$ ) among the susceptibility variants from either the GAIN or TGen study.

Given the reported enrichment of *cis* mQTLs, we applied mQTL annotation to the top loci from GWAS of Bipolar Disorder, as represented in the NHGRI GWAS repository<sup>22</sup>, to assess the functional relevance of these reported susceptibility loci (see Table 1).

### Overlap of methylation QTLs and expression QTLs in the top Bipolar Disorder SNPs

We then sought to determine the extent of overlap between the mQTLs and the eQTLs in the top Bipolar Disorder associated SNPs. First, we evaluated for (*cis*) eQTL enrichment in cerebellum among the top Bipolar Disorder associated SNPs. As expected from our earlier study<sup>7</sup> (which was done, however, in transformed lymphocytes), we found a significant excess of locally acting eQTLs in cerebellum among the top SNPs. Again, from 1000 simulated sets, the observed count was 222 while the expected count was 167.9 with standard deviation of 10.9; no simulated set yielded a (*cis*) eQTL count that matched the observed count for an enrichment  $p$  value  $< 0.001$ . In contrast, we found no evidence for excess of *trans*-acting eQTLs (enrichment  $p = 0.36$ ) in cerebellum among the top Bipolar

Disorder associated SNPs. The enrichment we report here does not appear to depend on the 4 Mb definition of *cis* since the use of a lower threshold (1 Mb) preserves the enrichment robustly (enrichment  $p < 0.001$ ).

Given this enrichment for eQTLs, we asked whether the enrichment for *cis* mQTLs could be explained by our observation on the excess of *cis* eQTLs relative to expectation in cerebellum. Of the 132 *cis* mQTLs among the top Bipolar Disorder associated SNPs, 55 were found to be mQTLs that are *not cis* eQTLs. The remaining 77 SNPs are *cis* eQTLs. However, when we required that *cis* mQTLs and *cis* eQTLs have the same target genes in the SNP-gene pair, of 169 identified mQTL-gene pairs (resulting from several of the 132 *cis* mQTLs having multiple methylation gene targets), we identified only 8 pairs with the same SNP-gene combination as *cis* eQTL. A *cis* mQTL SNP for a given gene may be a *cis* eQTL of a different gene target. Of these 8 pairs, 5 are for one gene, *DLG5* (for *cis* joint eQTLs-mQTLs: *rs10762763*, *rs11002306*, *rs1866437*, *rs2579177*, and *rs2579178*, the first two of which are intronic to the gene). The overlap of the target genes of the *cis* mQTL and of the *cis* eQTL is not greater than expected (hypergeometric distribution,  $p = 0.25$ ).

To evaluate the significance of the observed number of joint *cis* eQTLs-mQTLs (for the same target gene) among the top Bipolar Disorder associated SNPs, we conducted simulations (see Materials and Methods for details of the simulation procedure to empirically generate the null distribution to test the significance of the observed count of such joint *cis* eQTLs-mQTLs). The results of the simulation analysis on joint *cis* eQTLs-mQTLs show that such SNPs altering expression level and DNA methylation of the same proximal transcript constitute a small proportion of either *cis* mQTLs or *cis* eQTLs, and that the count of such joint eQTLs-mQTLs falls well within the range (enrichment  $p = 0.75$ ) that can occur by chance based on our simulation analysis. 75% of all simulated sets matched or exceeded the actual count of joint (*cis*) eQTLs-mQTLs.

Notably, SNPs affecting both methylation and expression (allowing the respective target genes to be different) are significantly enriched (enrichment  $p < 0.001$ ) among the top Bipolar Disorder associated SNPs.

Thus, we conclude that the enrichment of cerebellum mQTLs and that of eQTLs in the Bipolar Disorder data are independent features. The variants that are both eQTLs and mQTLs for the same proximal target transcript are not driving the observed enrichment. They should thus be considered separately in terms of biological functionality.

### Tissue specificity of mQTLs

To determine whether there is tissue specificity regarding the contribution of mQTLs to disease, we also assessed the level of overlap between mQTLs in cerebellum and in LCLs among the top associations with Bipolar Disorder. We compared our findings with the (publicly available) results from a recent association analysis of methylation levels<sup>23</sup> in the HapMap lymphoblastoid cell lines to look for evidence of cross-tissue common regulators. The methylation profiles for these cell lines were assayed using the same platform as our human cerebellum samples. We found among the top SNP associations ( $p < 0.001$ ) from the WTCCC study no overlap between mQTLs identified in brain with the *cis* mQTLs identified

in the LCLs. Indeed, none of the mQTLs identified in LCLs overlapped with the most significant SNPs ( $p < 0.001$ ) from the WTCCC study.

### The use of mQTL information enhances ability to detect significant genotypic association with Bipolar Disorder

We used the results of GWA studies of Bipolar Disorder (GAIN and TGen)<sup>11, 12</sup> in order to evaluate whether incorporating *a priori* functional (mQTL) information into GWA studies would have utility for increasing power to detect significant associations. The meta-analysis of the original GWA studies yielded no genome-wide significant findings<sup>11, 12</sup>. But restricting the association analysis to *cis* mQTLs ( $p < 0.001$ ,  $n=59959$ ) identified a SNP, *rs12618769*, that acts locally to regulate methylation patterns for *INPP4A* ( $p = 0.0009$ ) and exhibits significant association (TGen+GAIN,  $p = 8.318e-07$ , OR=1.43, T allele) with Bipolar Disorder that can survive Bonferroni correction for 59959 tests. We also found the SNP was a *cis* eQTL for *UNC50* ( $p = 0.009$ ). In the WTCCC study<sup>13</sup>, the SNP showed a replicated association at  $p = 0.02$  (OR=1.14, T allele). If we restrict the association analyses to *cis* mQTLs-*cis* eQTLs (i.e., SNPs affecting both methylation and expression regardless of the target gene), the SNP *rs12618769* continues to be significant with the number of tests reduced to 18173.

## Discussions

In this study, an approach to the investigation of the genetic basis of Bipolar Disorder was pursued. Particularly, we sought to integrate the results of genome-wide analyses of genetic regulation of DNA methylation and of gene expression to inform on the biological relevance of the results from genome-wide association studies of Bipolar Disorder.

A recent publication has revealed the impact of genetic control of gene expression in brain on Schizophrenia<sup>24</sup>. Thus, our effort was aimed *primarily* at understanding the impact of genetic control of DNA methylation on Bipolar Disorder, a common complex disease for which epigenetic mechanisms have been proposed as critical for pathophysiology. The proposed epigenetic connection is consistent with such observations as reported clinical differences between males and females and the fluctuation (i.e., clinical remissions and relapses) of disease course<sup>15</sup>. Our study implicates SNP-mediated regulation of epigenetic modification in the etiology of Bipolar Disorder. We found that susceptibility loci of Bipolar Disorder are enriched for genetic variation with substantial impact on local methylation patterns.

We were also *secondarily* interested in integrating studies of genetic regulation of DNA methylation and of transcriptional expression, on a genome-wide scale, into a coherent framework for characterizing the underlying biological relevance of disease susceptibility loci. We found that of the most significant associations with Bipolar Disorder, SNPs that control both methylation and the expression of the same proximal gene are a small minority. This is consistent with a recent report that concludes that mQTLs and eQTLs largely capture independent traits<sup>25</sup>. This finding lends substantial support to our thesis that, within the context of studies of brain pathophysiology, genetic variation controlling gene expression and genetic regulation of methylation present largely independent sources of variability in



disease susceptibility. While a SNP may regulate both methylation and gene expression in brain, it is likely to do so with different target genes. DNA methylation may contribute to disease risk other than through influencing gene expression, for example by affecting DNA stability. Alternatively, methylation effects on expression may be more difficult to detect due to methodological limitations; for example, RNA is much more susceptible to degradation than DNA.

The result of our *cis* eQTL enrichment analysis strongly demonstrates the influence of genetic regulation of gene expression in brain on Bipolar Disorder. In contrast to the enrichment method (which tests a polygenic model that includes tens of thousands of SNPs with disease association  $p < 0.50$ ) utilized by a recent study<sup>24</sup> that reports enrichment for alleles that affect gene expression in brain among Schizophrenia susceptibility alleles, our enrichment results were robustly observed in the significant Bipolar Disorder susceptibility associations at different significance thresholds ( $p < 0.001$  and  $p < 0.0001$ ) for association with disease. Finally, through the use of local mQTL information, we identified a significant association, *rs12618769*, that can stand multiple testing correction for all mQTLs tested in the GWA studies of TGen+GAIN Bipolar Disorder. We found the SNP to regulate ( $p = 0.0009$ ) the methylation of inositol polyphosphate phosphatase 4A (*INPP4A*) in *cis*. The gene *INPP4A* has been shown to protect neurons from excitotoxic cell death and to maintain the functional integrity of the brain<sup>26</sup>. Remarkably, a previous study has shown that increased excitotoxicity may account for disease progression in Bipolar Disorder<sup>27</sup>. The gene *INPP4A* encodes one of the enzymes involved in the phosphatidylinositol signaling pathway, which is a target for the effects of lithium<sup>28</sup>. We obtained replication of the association in the independent WTCCC GWA study. We also found the SNP was a *cis* eQTL for *UNC50* ( $p = 0.009$ ). It is worth noting that considering SNPs that are both *cis* mQTLs and *cis* eQTLs even if the methylation and expression targets are not the same gene may increase power to detect an association in genome-wide studies of Bipolar Disorder by further reducing the multiple testing burden and providing a potentially more robust genetic association finding.

We believe that genome-wide association studies of neuropsychiatric phenotypes merit the kind of investigation made possible by the genome-wide maps of mQTLs and eQTLs. Though much of the functionality of mQTLs that do not correlate with eQTLs remains to be elucidated, the effect of genetic variation on methylation suggests an important new area to explore and may provide mechanistic insights into genome-wide association findings.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

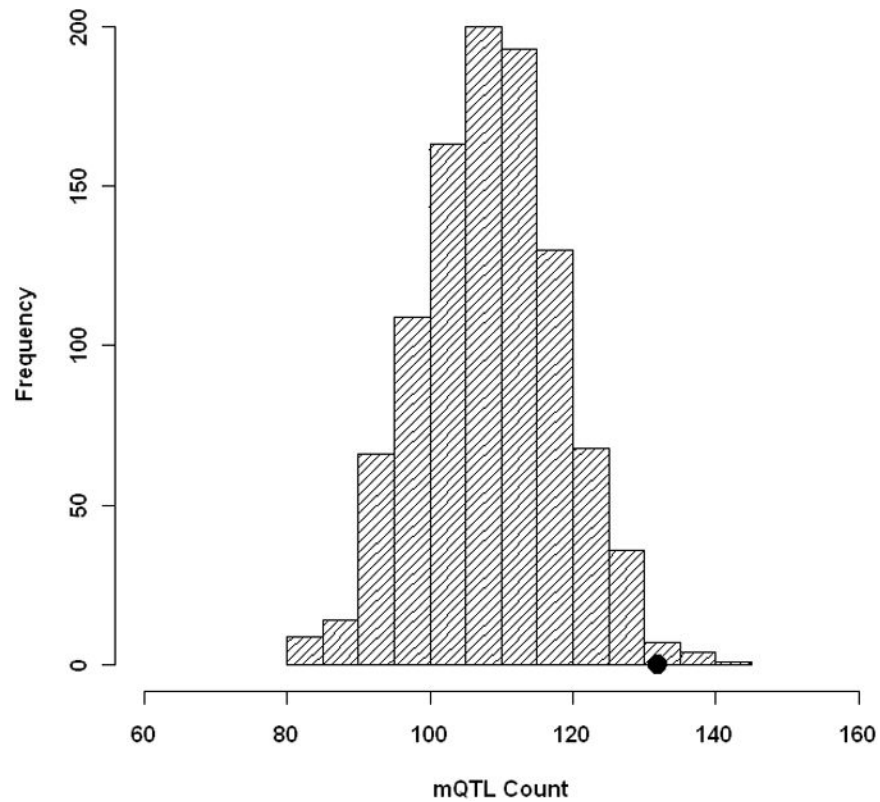
## Acknowledgments

This work was funded through 5R01MH080425 (to CL), PAAR (Pharmacogenetics of Anti-cancer Agents Research; U01 GM61393), ENDGAME (ENhancing Development of Genome-wide Association Methods) initiative (U01 HL084715), and the Genotype-Tissue Expression project (GTEx) (R01 MH090937).

## References

1. Egger G, Liang G, Aparicio A, Jones PA. Epigenetics in human disease and prospects for epigenetic therapy. *Nature*. 2004; 429(6990):457–463. [PubMed: 15164071]
2. Park PJ. Epigenetics meets next-generation sequencing. *Epigenetics*. 2008; 3(6):318–321. [PubMed: 19098449]
3. Schumacher A, Kapranov P, Kaminsky Z, Flanagan J, Assadzadeh A, Yau P, et al. Microarray-based DNA methylation profiling: technology and applications. *Nucleic Acids Res*. 2006; 34(2): 528–542. [PubMed: 16428248]
4. Feinberg AP. Phenotypic plasticity and the epigenetics of human disease. *Nature*. 2007; 447(7143): 433–440. [PubMed: 17522677]
5. Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet*. 1999; 23(2): 185–188. [PubMed: 10508514]
6. Gamazon ER, Zhang W, Konkashbaev A, Duan S, Kistner EO, Nicolae DL, et al. SCAN: SNP and copy number annotation. *Bioinformatics*. 2010; 26(2):259–262. [PubMed: 19933162]
7. Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, Cox NJ. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. *PLoS Genet*. 2010; 6(4):e1000888. [PubMed: 20369019]
8. Nica AC, Montgomery SB, Dimas AS, Stranger BE, Beazley C, Barroso I, et al. Candidate causal regulatory effects by integration of expression QTLs with complex trait genetic associations. *PLoS Genet*. 2010; 6(4):e1000895. [PubMed: 20369022]
9. Zhang D, Cheng L, Badner JA, Chen C, Chen Q, Luo W, et al. Genetic control of individual differences in gene-specific methylation in human brain. *Am J Hum Genet*. 2010; 86(3):411–419. [PubMed: 20215007]
10. Liu C, Cheng L, Badner JA, Zhang D, Craig DW, Redman M, et al. Whole-genome association mapping of gene expression in the human prefrontal cortex. *Mol Psychiatry*. 2010; 15(8):779–784. [PubMed: 20351726]
11. Smith EN, Bloss CS, Badner JA, Barrett T, Belmonte PL, Berrettini W, et al. Genome-wide association study of bipolar disorder in European American and African American individuals. *Mol Psychiatry*. 2009; 14(8):755–763. [PubMed: 19488044]
12. Smith EN, Koller DL, Panganiban C, Szelinger S, Zhang P, Badner JA, et al. Genome-Wide Association of Bipolar Disorder Suggests an Enrichment of Replicable Associations in Regions near Genes. *PLoS Genet*. 2011; 7(6):e1002134. [PubMed: 21738484]
13. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007; 447(7145):661–678. [PubMed: 17554300]
14. Kato T, Iwamoto K, Kakiuchi C, Kuratomi G, Okazaki Y. Genetic or epigenetic difference causing discordance between monozygotic twins as a clue to molecular basis of mental disorders. *Mol Psychiatry*. 2005; 10(7):622–630. [PubMed: 15838537]
15. Petronis A. Epigenetics and bipolar disorder: new opportunities and challenges. *Am J Med Genet C Semin Med Genet*. 2003; 123C(1):65–75. [PubMed: 14601038]
16. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol*. 2010; 34(8):816–834. [PubMed: 21058334]
17. Li Y, Willer C, Sanna S, Abecasis G. Genotype imputation. *Annu Rev Genomics Hum Genet*. 2009; 10:387–406. [PubMed: 19715440]
18. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*. 2007; 8(1):118–127. [PubMed: 16632515]
19. Leek JT, Storey JD. Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS Genet*. 2007; 3(9):1724–1735. [PubMed: 17907809]
20. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007; 81(3): 559–575. [PubMed: 17701901]

21. Gamazon ER, Nicolae DL, Cox NJ. A study of CNVs as trait-associated polymorphisms and as expression quantitative trait loci. *PLoS Genet.* 2011; 7(2):e1001292. [PubMed: 21304891]
22. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A.* 2009; 106(23):9362–9367. [PubMed: 19474294]
23. Bell JT, Pai AA, Pickrell JK, Gaffney DJ, Pique-Regi R, Degner JF, et al. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. *Genome Biol.* 2011; 12(1):R10. [PubMed: 21251332]
24. Richards AL, Jones L, Moskvina V, Kirov G, Gejman PV, Levinson DF, et al. Schizophrenia susceptibility alleles are enriched for alleles that affect gene expression in adult human brain. *Mol Psychiatry.* 2011
25. Gibbs JR, van der Brug MP, Hernandez DG, Traynor BJ, Nalls MA, Lai SL, et al. Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet.* 2010; 6(5):e1000952. [PubMed: 20485568]
26. Sasaki J, Kofuji S, Itoh R, Momiyama T, Takayama K, Murakami H, et al. The PtdIns(3,4)P(2) phosphatase INPP4A is a suppressor of excitotoxic neuronal death. *Nature.* 2010; 465(7297):497–501. [PubMed: 20463662]
27. Rao JS, Harry GJ, Rapoport SI, Kim HW. Increased excitotoxicity and neuroinflammatory markers in postmortem frontal cortex from bipolar disorder patients. *Mol Psychiatry.* 2010; 15(4):384–392. [PubMed: 19488045]
28. Berridge MJ. The Albert Lasker Medical Awards. Inositol trisphosphate, calcium, lithium, and cell signaling. *JAMA.* 1989; 262(13):1834–1841. [PubMed: 2674489]



**Figure 1.**

Top susceptibility loci from the Wellcome Trust genome-wide association study of Bipolar Disorder are significantly enriched for *cis* mQTLs. The distribution of the number of *cis* mQTLs in 1,000 draws (each of same count as the number of susceptibility loci tested) of frequency-matched variants is shown in the bar graphs, with the actual number of *cis* mQTLs observed in Wellcome Trust study shown as a solid circle. The expected value for the *cis* mQTL count was 108.9 (with standard deviation of 9.7) while the observed count was 132, which yields an enrichment p value of 0.01.

**Table 1**  
**A list of mQTLs associated with Bipolar Disorder in the NHGRI catalog of published genome-wide association studies**

The Gene column is the target gene for the mQTL. P-value is from the association between the corresponding SNP and methylation level. PMID is the Pubmed ID of the study of origin. All SNPs included here satisfy  $p < 10^{-5}$  with Bipolar Disorder in the original study.

Gene	Chromosome	Position	SNP	mQTL P-value	PMID
<i>BRN1</i>	12	2215556	rs1006737	5.40E-05	20351715
<i>C12orf4</i>	12	2215556	rs1006737	0.002529	20351715
<i>TEKT1</i>	12	2215556	rs1006737	4.65E-05	20351715
<i>CCL3L3</i>	13	41551437	rs1012053	2.37E-07	17486107
<i>PDK3</i>	14	103578829	rs11622475	8.35E-05	17554300
<i>FLJ35695</i>	15	36782783	rs12899449	1.63E-10	18711365
<i>NECAP1</i>	3	3633840	rs1601875	2.84E-05	18711365
<i>TAS2R48</i>	9	33789370	rs216345	5.15E-05	18711365
<i>C9orf82</i>	2	241164269	rs2953145	9.67E-05	17554300
<i>FLJ38379</i>	2	241164269	rs2953145	0.002306	17554300
<i>GUCY1A3</i>	2	241164269	rs2953145	4.79E-05	17554300
<i>MYOZ3</i>	9	129147785	rs4130590	5.64E-05	18711365
<i>GCKR</i>	16	23541527	rs420259	8.92E-05	17554300
<i>HLA-C</i>	16	23541527	rs420259	6.88E-05	17554300
<i>ZNF300</i>	9	112237084	rs7042161	2.73E-05	18711365