Ehrlich, S., Bustillo, J. R., Perrone-Bizzozero, N., Walton, E., Calhoun, V. D., & Liu, J. (2017). Cross-Tissue Exploration of Genetic and Epigenetic Effects on Brain Gray Matter in Schizophrenia. *Schizophrenia Bulletin*. https://doi.org/10.1093/schbul/sbx068

Peer reviewed version

Cross tissue exploration of genetic and epigenetic effects on brain gray matter in schizophrenia

Dongdong Lin[1], Jiayu Chen[1], Stefan Ehrlich[6], Juan R. Bustillo[2,4], Nora Perrone-Bizzozero[2,4], Esther Walton[7], Vincent P. Clark[1,5], Yu-Ping Wang[8], Jing Sui[1,9], Yuhui Du[1], Beng C. Ho[10], Charles S. Schulz[11], Vince D. Calhoun[1,2,3,*], Jingyu Liu[1,3,*,†]

[1]The Mind Research Network and Lovelace Biomedical and Environmental Research Institute, Albuquerque, NM, 87106, USA

[2]Dept. of Neurosciences, University of New Mexico, Albuquerque, NM, 87131, USA

[3]Dept. of Electronic and Computer Engineering, University of New Mexico, Albuquerque, NM, 87131, USA

[4]Dept. of Psychiatry, University of New Mexico, Albuquerque, NM, 87131, USA

[5]Dept. of Psychology, University of New Mexico, Albuquerque, NM, 87131, USA

[6] Division of Psychological and Social Medicine and Developmental Neurosciences, Faculty of Medicine, Technische Universität Dresden, Dresden , 01307,Germany

[7]Dept. of Psychology, Georgia State University, Atlanta, GA, 30302, USA

[8]Dept. of Biomedical Engineering, Tulane University, New Orleans, Louisiana, 70118, USA

[9]Brainnetome Center and National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China

[10]Department of Psychiatry, University of Iowa Carver College of Medicine, Iowa City, IA 52242, USA

[11]Department of Psychiatry, University of Minnesota, Minneapolis, MN, 55454, USA

*these authors contributed equally to the work

[†]Corresponding author: Jingyu Liu, 1101 Yale Blvd NE, Albuquerque, NM 87131, jliu@mrn.org

Running title: genetic, epigenetic, and brain interactions

Keywords: Schizophrenia, Genetics, DNA methylation, brain gray matter, meQTL, cross tissue

Abstract words: 243; Article total words: 3998; Figure count: 3; Table count: 2; supplementary count: 1

Abstract

Closely linking genetics and environment factors, epigenetics has been of increasing interest in psychiatric disease studies. In this work, we integrated single nucleotide polymorphisms (SNPs), DNA methylation of blood and saliva, and brain gray matter (GM) measures to explore the role of genetic and epigenetic variation to the brain structure changes in schizophrenia (SZ). By focusing on the reported SZ genetic risk regions, we applied a multi-stage multivariate analysis to a discovery dataset (92 SZ patients and 110 controls, blood) and an independent replication dataset (93 SZ patients and 99 controls, saliva).

Two pairs of SNP-methylation components were significantly correlated (r=0.48 and 0.35) in blood DNA, and replicated (r=0.46 and 0.29) in saliva DNA, reflecting cross-tissue SNP cis-effects. In the discovery data both SNP-related methylation components were also associated with one GM component primarily located in cerebellum, caudate and thalamus. Additionally, another methylation component in *NOSIP* gene with significant SZ patient differences (p=0.009), was associated with eight GM components (seven with patient differences) including superior, middle, and inferior frontal gyri, superior, middle, and inferior temporal gyri, cerebellum, insula, cuneus, and lingual gyrus. Of these, five methylation-GM associations were replicated ($p<.05$). In contrast, no pairwise significant associations were observed between SNP and GM components.

This study strongly supports that compared to genetic variation, epigenetics show broader and more significant associations with brain structure as well as diagnosis, which can be cross tissue, and the potential in explaining the mechanism of genetic risks in SZ.

Introduction

Schizophrenia (SZ) is a highly heritable complex disease with a lifetime prevalence of around 1%[1-3]. Recent landmark work from the psychiatric genomics consortium (PGC) has identified 128 genome-wide significant SZ-risk SNPs[4] from 108 highly linkage genomic regions (R^2>0.6), providing reliable guidance for investigating the etiology and pathophysiology of schizophrenia.

To uncover the biological mechanisms of such SZ-risk variants , a valuable strategy is to explore the genetic influence on brain-based phenotypes through magnetic resonance imaging (MRI) techniques[5]. Imaging-derived structural or functional measures of the brain (e.g., gray matter density) have been shown to be reliable and highly heritable[6-8]. Considerable success in the imaging genetics field has enriched our understanding of the putative neurobiological pathways from genetic mutations to system level and behavioral changes resulting in psychiatric disorders[9-12]. However, a recent study of genetic influences on SZ and subcortical brain volumes found no clear evidence of shared genetic variants[13]. Although this study is limited to the selected subcortical regions, it suggests that more complex biological pathways may involve bridging the gap from DNA variation to brain system changes in SZ.

Epigenetics, beyond genetic sequence variation, regulates gene expression and thus influences biological functions[14,15]. Epigenetics can be modulated by genetics[16], potentially mediating genetic risks[17], and also integrate environmental factors in the etiopathogenesis of SZ[18-20]. DNA methylation, as one of the main epigenetic modifications, has shown to be related to SZ symptoms[21], disease onset[22], cognitive deficit[23], and be significantly changed in schizophrenic brain[24,25] and peripheral tissues[26]. Although methylation tends to be tissue specific in general, recent studies have noted a set of methylation sites significantly correlated between blood and brain[27-29]. This cross-tissue mechanism is hypothesized to be through genetics, called methylation quantitative trait loci (meQTL) effect[30,31]. In parallel, accumulating evidence has shown the relationship of peripheral DNA methylation with brain structure and function[23,32-

[35]. These findings support the potential of peripheral methylation in association with brain structure or function to further understand psychiatric disorders.

Two recent studies have demonstrated the significant role of meQTL in brain development and SZ[31,36]. However, the underlying mechanisms of how epigenetic factors interplay with genetic factors, and further influence brain structural variation and psychiatric disorders are far from clear. In this study, we explored the interconnections between the three modalities -single nucleotide polymorphisms(SNPs), DNA methylation from blood and saliva, and whole brain gray matter (GM)- from SZ patients and healthy controls (HCs) in two independent datasets. We focused on SNPs and methylation sites within the 108 SZ-risk regions identified by PGC[4]. Given the dependence of loci in both genotype and methylation, a multivariate analysis based on parallel independent component analysis (ICA)[37], which has been suggested to identify clusters of correlated loci in genetic applications[38,39], was performed to capture the intrinsic genetic-epigenetic relationships. Brain networks were extracted by ICA using GM images from structural MRI[40]. Finally, we tested the relationships among components of three modalities, as well as their associations with clinical diagnosis, symptoms, and cognition.

Methods and Materials

More details of methods and materials were provided in the supplementary information (SI Materials and Methods).

Participants

The discovery dataset consisted of 202 subjects recruited by the Mind Clinical Imaging Consortium study[41]. Patients met criteria for DSM-IV-TR schizophrenia, schizoaffective, or schizophreniform disorders. Healthy controls had no current or past history of psychiatric illness including substance abuse or dependence. Following a similar recruitment protocol, the replication data were collected from the Center for Biomedical Research Excellence study[42] and Glutamate and Outcome in Schizophrenia study[43]. All participants of these studies as shown in Table 1 provided written consent forms.

In the discovery data symptoms were assessed with the Scale of the Assessment of Positive Symptoms[44] and the Scale of the Assessment of Negative Symptoms[45]. Alcohol abuse/dependence status was assessed by DSM-IV and 101 participants provided reliably nicotine use information measured by pack per year[46]. In the replication data, participants had no current diagnosis of alcohol abuse or dependence, and nicotine use was assessed by Fagerstrom Test for Nicotine Dependence[47]. For the cognitive assessment, we focused on two comparable tests between studies: 1) letter-number sequencing test of the WAIS-III and MATRICS for measuring working memory[48,49], and 2) computerized version of the Tower of London(TOL) test[50] and the NAB maze test[51] for measuring planning and problem solving.

Data preprocessing

DNA was extracted from whole blood in the discovery data. Genotyping was performed using the Illumina Infinium HumanOmni1-Quad assay. Data was quality controlled including checking for missing data, sex mismatch, relatedness, Hardy-Weinberg Equilibrium($<1\times10^{-6}$), minor allele frequency(>0.01) and population structure correction. DNA methylation was assayed by the Illumina Infinium HumanMethylation27 assay. Quality control included removal of bad samples and probes, batch correction, and removal of non-variant sites. Age and gender were further adjusted by linear regression model. Finally, we narrowed down our analyses to the 108 SZ-susceptible regions[4], resulting in 4,475 SNPs and 102 cytosine-phosphate-guanine(CpG) sites.

Saliva samples were used for DNA extraction in the replication data. Genotyping was performed using both Illumina Infinium HumanOmni1-Quad and Omni5 assays. The SNP data was quality controlled by following the same procedure as that in MCIC data. Missing genotypes were imputed with the 1000 genomes reference panel. DNA methylation was assayed by the Infinium MethylationEPIC beadchip. Quality control was performed using package 'minfi'[52]. Batch, age and gender were adjusted on

each site prior to analysis. Cell type proportions in blood and saliva samples were estimated using the algorithm by Houseman et.al[53].

T1-weighted MRI images in both datasets were preprocessed using SPM12 (http://www.fil.ion.ucl.ac.uk/spm/software/spm12/). Unmodulated gray matter images were smoothed by a full width half maximum Gaussian kernel of 8 mm, followed by the correction for scanning sites (for the discovery data), age and gender on each voxel as suggested in previous studies[54,55]. No significant differences on GM networks were observed when using modulated images (data not shown).

Primary analysis

A multi-level association analysis was applied as shown in Fig.1. Parallel ICA was first applied to SNP and DNA methylation data to derive independent genetic and methylation components, while maximizing inter-modality correlations of the components' loadings. Independent components of each modality were linear combinations of individual variants, capturing interactions among SNPs or CpG sites. A loading vector presented the expression of each component across subjects. Ten-fold subsample and permutation tests were conducted to verify the stability and significance of the SNP-methylation component links. For the linked SNP-methylation components, meQTL analysis was applied to the contributing SNPs and CpG sites to verify SNP effect on methylation at individual site level. In parallel, independent GM components were derived by a regular ICA on whole brain GM images. Then, the SNP and methylation components were tested for associations with GM components using a linear regression model. For any three-modality links, a mediation model was further tested for the methylation component mediating the genetic component's effect on the GM component. In addition, all SNP, methylation and GM components were tested for the associations with diagnosis, symptoms, and cognition to help elucidate their relation to schizophrenia disease. Medication, cell type proportion, alcohol abuse and smoking were also assessed with their influences on the results as additional covariates in regression model. All the tests were

corrected for multiple comparison using the false discovery rate (FDR<=.05)[56]. For replication, we extracted the same SNP, methylation and GM components identified in the discovery data using a projection method and assessed their associations. The significance $p<0.05$ and the same direction of correlation were set for replication.

Results

SNP, DNA methylation components and inter-modality relationship

In the discovery data, 15 SNP components and 10 methylation components were extracted. Among them, three methylation components, but no SNP component showed significant disease association. After correction for multiple comparisons, two pairs of SNP-methylation components were significantly correlated with r=0.48 (p=5.46×10^{-11}) and 0.35 (p=3.28×10^{-5}). The correlations were consistent in ten-fold stability tests with the average correlations of 0.50±0.04 and 0.25±0.11, respectively. No significant effects of race and alcohol abuse were observed on the SNP-methylation associations. Permutation tests showed the significance of the two links with empirical p-values of 1×10^{-3} and 0.01.

Fig.2(B,E) shows the loadings of the two pairs of SNP and methylation components across subjects. Both pairs indicated increased SNP loading along with the higher methylation loading. A higher loading indicates more presentation of positively weighted SNPs, and hyper-methylation of positively weighted CpG sites. The HC and SZ groups exhibited comparable correlations for both pairs (1st pair: 0.49 and 0.46; 2nd pair: 0.39 and 0.23), respectively. Excluding the schizoaffective and schizophreniform patients did not change the correlation by much. Marginal group differences were found in components of the second pair (p_{snp}=0.09, $p_{methylation}$=0.04, uncorrected). After correcting for clinical diagnosis, the SNP-methylation associations remained significant for both pairs (p=4.41×10^{-10}, and 2.36×10^{-4}). By thresholding the contribution weights of loci in the components with |z-score|>2.5, there were 94 significant SNPs from 20 genes or their nearby regions in the first linked SNP component (Table S1), and five CpG sites in the genes *NT5C2*, *CSMD1*, *FES*, and *ARL3* in the corresponding methylation

component. In the second pair, 60 highly weighted SNPs from 19 genes and one CpG site in the gene *ITIH4* were involved in the correlation (Table S2).

meQTL analysis on the contributing SNPs and CpG sites found 69 meQTLs in the first pair and 59 meQTLs in the second pair after FDR correction, mostly showing *cis* effects (SNP effect of CpG site within 500kbp distance[31]). In the first pair *cis*-acting meQTLs were located on chromosome 10 around three CpG sites in *NT5C2* and *ARL3* (Fig.2A), where the strongest cis-effect resides in SNP rs1926030 and CpG cg00035347 with partial η^2 (% of variance explained) =50.3%, $p=3.6\times10^{-32}$ (Fig.2C), indicating that an increasing minor allele *A* count in rs1926030 was associated with hypomethylation at cg00035347 located in the promoter of *NT5C2*. The *cis*-effecting meQTLs in the second pair was found in chromosome 3 around *ITHI4* (Fig.2D). The strongest meQTL effect was between SNP rs2071041 and CpG cg17890764 with partial η^2 = 13.9%, $p=4.5\times10^{-8}$ (Fig.2F). Other meQTL relations were listed in Tables S3 and S4. Overall, the cis-effecting meQTLs strongly affected the methylation change (partial η^2 median value: 10.88%, interquartile change IQR: 6.95%-21.24%).

We further compared our meQTLs with two recent meQTL studies in fetal[31] and adult brain tissues[36]. Among 69 meQTLs identified in the first SNP-methylation pair, 14 were also separately identified in both brain tissues with 12 in common. For the 59 meQTLs in the second pair, 13 were found in the fetal brain tissue and 11 were in adult brain tissue with 7 in common.

In the replication data 72.21% of SNPs and 52% of methylation sites used in the discovery data were present, including over 96% of contributing loci and CpG sites ($|z|>2.5$) in the identified significant SNP-methylation pairs. The replicated correlations for the two linked SNP-methylation pairs were r=0.46, $p=1.59\times10^{-11}$ and r=0.29, $p=5.9\times10^{-5}$, respectively, and the correlations (r=0.53, $p=4.93\times10^{-8}$; r=0.26, $p=1.08\times10^{-2}$) were consistent in SZ group only (no significant difference after excluding schizoaffective patients). In the first SNP-methylation pair, 72 meQTLs were identified including 67 (97.1%) found in

the discovery set. Similarly, for the second SNP-methylation pair, 57 meQTLs were identified and replicated 98.28% meQTLs in the discovery data.

GM components and their associations with DNA methylation

GM data were decomposed into 15 independent components by ICA, of which 10 components showed significant associations with diagnosis (Fig.S2). Nine components (Fig.3) were significantly associated with three methylation components, as listed in Table 2. Interestingly, seven out of nine GM components showed significant group differences, and remained significant after controlling for alcohol abuse except component 7 with p=0.06. It was worth noting that in the GM components 1 and 2 with the largest group differences, SZ patients showed substantial GM loss in the areas of superior/middle/medial/inferior frontal gyri, and superior/middle/inferior temporal gyri.

Among the three methylation components, methylation component 1, mainly involving the *NOSIP* gene, showed a significant group difference (Cohen's d = -0.348; p =0.009). Consistently, the CpG site (cg05696092) located in the gene promoter region was hypomethylated in SZ ($p= 6.4\times10^{-3}$,uncorrected), and showed 5.17% mean methylation beta difference between two groups. This component was positively associated with GM components 1,2,3,5 and 7, and negatively associated with GM components 4,6 and 8 after multiple comparisons correction, as listed in Table 2. These GM-methylation associations remained significant after controlling for diagnosis or in SZ group only (Table S8), suggesting that the GM-methylation relationship is more general rather than driven by the disease. The interaction between methylation component 1 and diagnosis was significant in four pairs of GM-methylation components with stronger correlation in the patient group than in the HC group (Fig.S3 and Table S5), likely reflecting larger variations (e.g., much lower loadings in the GM and methylation components) only observed in patients.

Methylation components 2 and 3 were the ones linked to the SNP components aforementioned, and also significantly associated with GM component 9, mainly located at cerebellum, thalamus and

caudate (R^2=0.038, and 0.056; p=0.05 and 8.7×10^{-3}, respectively). Although connected to the same GM component, increasing GM in the component 9 regions was associated with decreased loadings of methylation component 2, and increased loadings of methylation component 3.

With respect to medication and symptoms, no significant effects were observed on the three methylation and nine GM components using the threshold FDR≤0.05 in SZ patients. However, significant associations were found between methylation component 1 with the working memory score (R^2=0.053, p $=1.27\times10^{-3}$) and the TOL score (R^2=0.023, p=0.04). Seven out of eight GM components related to the methylation component 1 were also significantly associated with the working memory score while four of them were related to the TOL score after correcting for multiple testing (Table S7).

In the replication data, we replicated significant correlations between methylation component 1 and five GM components 1,2,3,5,7 (Table S6).The correlations remained significant after controlling for diagnosis($p<0.05$) and in SZ group only (Table S8). We also replicated the associations of GM components 1,2,4,7 with the working memory score and GM components 1,2,5,7 with the TOL score ($p<0.06$; Table S7). GM components 1,2,3,7 showed significant group differences with the same direction of GM changes as in the discovery data, indicating reduced GM in SZ patients on the corresponding brain regions. The correlations between methylation components 2 and 3 with the cerebellum GM component were not significant in the replication data (p=0.15), but showed similar effects in HC group (R^2 =0.05, p = 0.027).

Three-way link of SNP, DNA methylation and GM

There was no significant association between SNP components and GM components in both datasets. However, the SNP component linked to methylation component 3 (the second SNP-methylation pair) was also marginally associated with GM component 9 ($p=3.38\times10^{-4}$, uncorrected). A mediation test showed that methylation component 3 significantly mediated the SNP component's effect on GM component 9 (average mediation effect = 0.0039, the proportion of SNP effect mediated: 23.63%, p=0.02).

Discussion and conclusions

In this work, we observed, in blood and saliva samples, two pairs of significantly related SNP-methylation components by multivariate analysis. The correlations were consistent in both SZ and HC groups and mainly contributed by cis-meQTLs (Fig.S1), suggesting that such intrinsic SNP-methylation relationship was driven by the effects of the SNPs on DNA methylation in close proximity, consistent with previous findings[4,18,31,36]. The benefit of identifying component based SNP-methylation correlations, instead of individual meQTLs, lies in the fact that multiple SNP or methylation loci can be functionally related and be grouped into one factor[38]. Component analysis not only reduced the number of statistical test but also derived a region of interest. The main regions covered by both SNP-methylation component pairs (Chr.10 around *NT5C2*, *ARL3*, and Chr.3 around *ITIH4* respectively) were also identified as prominent meQTL regions in previous studies using brain tissues[31,36].

More than 97% of meQTLs were replicated in saliva, and about 20% were also reported in studies of brain tissues passing genome-wide significance[31,36]. We speculated more meQTLs might be replicated in brain if less strict significance thresholds were used. Given meQTL has been proposed to contribute to cross-tissue similarity[30,57], we used data of blood and brain samples from the same 12 epilepsy subjects[29] to test cross-tissue similarity of the identified CpG sites. Two CpG sites(cg00035347 on *NT5C2* and cg10872209 on *ARL3*) involving the majority of significant cis-effect, showed strong blood-brain correspondence with correlation r=0.73 and 0.76, respectively. These findings support that the meQTL effects are relatively consistent across tissue types[31,36], and likely mediate cross-tissue methylation similarity. Hence it is plausible to study effects of this specific set of methylation sites, assessed from peripheral tissues, on brain structure in psychiatric disorders[17].

SNP-affected methylation components 2 and 3 in the blood sample were significantly associated with one cerebellum GM component, and showed an interesting mediation effect, even though these correlations were replicated significantly only in HC group. Although SNP-effect on methylation was consistent cross tissues, different types of peripheral tissues may have different surrogate ability for brain.

Given the contributing genes of these two components such as *ITIH4*, *NT5C2* have shown expressional changes in brain tissues in SZ patients (p=0.0025 and 0.01 respectively)[58], we speculate these methylation components are somewhat relevant to GM in brain. This hypothesis warrants further investigation using a larger sample.

Methylation component 1 from the gene coding for nitric oxide synthase interacting protein (*NOSIP*) was significantly associated with five GM components in both datasets. Further examination found a significant negative blood-brain correlation of the CpG site cg05696092 (r=-0.63)[29]. Considering that there was no meQTL identified on this site, a negative cross-tissue correlation suggests another mechanism, more likely environment-related, which may underlie the cross-tissue similarity as negative correlations were also observed in another study[59]. The site cg05696092 was located in the promoter region of *NOSIP* and was hypomethylated in blood and saliva samples of SZ patients, suggesting hypermethylation in brain of SZ. Testing for the change of *NOSIP* gene expression in brain showed significant decreased expression in SZ patients[58], suggesting the hypermethylation may repress *NOSIP* expression. The protein encoded by the gene *NOSIP* mainly sequesters and inhibits the activity of nitric oxide synthase (*NOS*) and reduces nitric oxide (*NO*) production[60,61]. Several studies have found higher levels of *NOS* expression or *NO* in subjects with psychiatric disorders cross brain including the frontal cortex[62,63], caudate[64], hippocampus[65] and cerebellum[66]. Additionally, *NO* or *NOS* have been shown to relate to neural cell apoptosis, synaptic plastic and remodeling, neurodegeneration, and the development of neuronal structures[67-69]. Taken together, we speculate that hypermethylation in the promoter of the *NOSIP* gene induces a decrease of *NOSIP* expression and an increase of *NO* production in the brain of patients, and further affects neuron activity which may be partially reflected by GM signal.

No significant SNP-GM associations were detected, and no SNP component showed group difference, perhaps due to the small sample size. Even though we focused on 108 SZ-risk regions, genetic effects on the brain may not be as large as expected[36]. In contrast, methylation component 1 showed altered methylation level in SZ in both datasets, and its associated four GM components also had

replicated group differences. The most significance GM density reduction included frontal, temporal, and cerebellum regions, and an increase of GM from small regions of pre/post-central gyri, consistent with SZ brain structural abnormalities reported in a meta-analysis[54]. Additionally, cognitive ability presented replicated, strong relation with GM components in various regions. Methylation in *NOSIP* also showed an interesting relationship with GM components, though it was not replicated. Thus, regarding the relationship between SZ and cognition, our data suggest that GM measures show strong association, followed by methylation, and no SNP association.

The findings of this study should be interpreted with consideration of several limitations**.** First, methylation assay in the discovery data only covers promoter regions, which significantly limits our discovery power. Although the replication data have much higher coverage, we chose to use it to replicate our findings. Secondly, we were unable to confirm the methylation measures via a different platform. However, some publications[70,71] have reported high reproducibility of Illumina array for DNA methylation assay, in line with our own test-retest stability (Spearman rank correlation: 0.96-0.99). Furthermore, we used an independent cohort from a different tissue to replicate the identified associations, strengthening the credibility of our findings. Thirdly, DNA from whole blood reflects a mixture of cell types. The estimated cell type proportion from blood[53] was correlated with methylation components 2 and 3 but did not alter the significance of the association findings. Fourthly, except alcohol abuse or dependence, other substance uses or comorbidities in SZ may have influence on epigenetics or GM. We tested the effect of nicotine use and further top 8 PCs of methylation data as surrogate of other potential confounding factors, and no significant change on the findings was observed. Finally, the findings in this study were mainly built on correlation analyses. We specifically intended to identify SNP-methylation associations, and then explore their effects on GM which could be further linked to the disease. While these associations suggest underlying biological pathways for SZ, we are unable to assess a causal relationship among the features. Other ICA-based methods (e.g., three-way parallel ICA[72]) can

be applied to the analysis of more than two modalities, but these methods may not optimally estimate the hierarchical relationship.

In summary, we first identified cross tissue SNP-methylation relationships within the known genetic SZ-risk regions, which were largely driven by cis-meQTL effects. We also found significant associations between methylation and GM components as well as SZ diagnosis in both blood and saliva tissues. These findings support a plausible mechanism of genetic risk to influence brain GM variation through epigenetic mediation. Furthermore, we identified cross-tissue similarities of methylation patterns unlikely to be regulated by SNP variants. Methylation-brain associations either regulated by SNPs or other factors, suggest the integral role epigenetics plays on brain structural variation and in the pathogenesis of SZ conveyed by both genetic and environmental influences.

Acknowledgements

This study was funded by the National Institutes of Health, grant number: P20GM103472 and R01EB005846, and National Science Foundation, grant number: 1539067.

The authors declare that they have no competing interests.

References

1. Mathers C, Fat DM, Boerma JT. *The global burden of disease: 2004 update.* World Health Organization; 2008.

2. Gottesman II. *Schizophrenia genesis: The origins of madness.* WH Freeman/Times Books/Henry Holt & Co; 1991.

3. Sullivan PF, Kendler KS, Neale MC. Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. *Archives of general psychiatry.* 2003;60(12):1187-1192.

4. Ripke S, Neale BM, Corvin A, et al. Biological insights from 108 schizophrenia-associated genetic loci. *Nature.* 2014;511(7510):421.

5. Meyer-Lindenberg A, Weinberger DR. Intermediate phenotypes and genetic mechanisms of psychiatric disorders. *Nature Reviews Neuroscience.* 2006;7(10):818-827.

6. Baaré WF, Pol HEH, Boomsma DI, et al. Quantitative genetic modeling of variation in human brain morphology. *Cerebral Cortex.* 2001;11(9):816-824.

7. Winkler AM, Kochunov P, Blangero J, et al. Cortical thickness or grey matter volume? The importance of selecting the phenotype for imaging genetics studies. *Neuroimage.* 2010;53(3):1135-1146.

8. Rose EJ, Donohoe G. Brain vs behavior: an effect size comparison of neuroimaging and cognitive studies of genetic risk for schizophrenia. *Schizophr Bull.* May 2013;39(3):518-526.

9. Lin D, Cao H, Calhoun VD, Wang Y-P. Sparse models for correlative and integrative analysis of imaging and genetic data. *Journal of neuroscience methods.* 2014;237:69-78.

10. Liu J, Calhoun VD. A review of multivariate analyses in imaging genetics. *Frontiers in neuroinformatics.* 2014;8:29.

11. Hibar DP, Stein JL, Renteria ME, et al. Common genetic variants influence human subcortical brain structures. *Nature.* Apr 9 2015;520(7546):224-229.

12. Stein JL, Medland SE, Vasquez AA, et al. Identification of common variants associated with human hippocampal and intracranial volumes. *Nat. Genet.* May 2012;44(5):552-561.

13. Franke B, Stein JL, Ripke S, et al. Genetic influences on schizophrenia and subcortical brain volumes: large-scale proof of concept. *Nature neuroscience.* 2016.

14. Nikolova YS, Hariri AR. Can we observe epigenetic effects on human brain function? *Trends in cognitive sciences.* 2015;19(7):366-373.

15. Jansen RC, Nap J-P. Genetical genomics: the added value from segregation. *TRENDS in Genetics.* 2001;17(7):388-391.

16. Gertz J, Varley KE, Reddy TE, et al. Analysis of DNA methylation in a three-generation family reveals widespread genetic influence on epigenetic regulation. *PLoS Genet.* 2011;7(8):e1002228.

17. Liu Y, Aryee MJ, Padyukov L, et al. Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. *Nat. Biotechnol.* Feb 2013;31(2):142-147.

18. Maric N, Svrakic D. Why schizophrenia genetics needs epigenetics: a review. *Psychiatria Danubina.* 2012;24(1.):2-18.

19. Tsuang M. Schizophrenia: genes and environment. *Biological psychiatry.* 2000;47(3):210-220.

20. Thibaut F. Why schizophrenia genetics needs epigenetics: a review. *Psychiatria Danubina.* Mar 2012;24(1):25-27.

21. Liu J, Chen J, Ehrlich S, et al. Methylation patterns in whole blood correlate with symptoms in schizophrenia patients. *Schizophrenia bulletin.* 2014;40(4):769-776.

22. Melas PA, Rogdaki M, Osby U, Schalling M, Lavebratt C, Ekstrom TJ. Epigenetic aberrations in leukocytes of patients with schizophrenia: association of global DNA methylation with antipsychotic drug treatment and disease onset. *FASEB J.* Jun 2012;26(6):2712-2718.
23. Walton E, Liu J, Hass J, et al. MB-COMT promoter DNA methylation is associated with working-memory processing in schizophrenia patients and healthy controls. *Epigenetics.* 2014;9(8):1101-1107.
24. Huang H-S, Akbarian S. GAD1 mRNA expression and DNA methylation in prefrontal cortex of subjects with schizophrenia. *PloS one.* 2007;2(8):e809.
25. Aston C, Jiang L, Sokolov BP. Microarray analysis of postmortem temporal cortex from patients with schizophrenia. *Journal of neuroscience research.* Sep 15 2004;77(6):858-866.
26. Aberg KA, McClay JL, Nerella S, et al. Methylome-wide association study of schizophrenia: identifying blood biomarker signatures of environmental insults. *JAMA psychiatry.* Mar 2014;71(3):255-264.
27. Davies MN, Volta M, Pidsley R, et al. Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. *Genome biology.* 2012;13(6):1.
28. Horvath S, Zhang Y, Langfelder P, et al. Aging effects on DNA methylation modules in human brain and blood tissue. *Genome biology.* 2012;13(10):1.
29. Walton E, Hass J, Liu J, et al. Correspondence of DNA methylation between blood and brain tissue and its application to schizophrenia research. *Schizophrenia bulletin.* 2015:sbv074.
30. Smith AK, Kilaru V, Kocak M, et al. Methylation quantitative trait loci (meQTLs) are consistently detected across ancestry, developmental stage, and tissue type. *BMC genomics.* 2014;15(1):1.
31. Hannon E, Spiers H, Viana J, et al. Methylation QTLs in the developing brain and their enrichment in schizophrenia risk loci. *Nature neuroscience.* 2016;19(1):48-54.
32. Booij L, Szyf M, Carballedo A, et al. DNA methylation of the serotonin transporter gene in peripheral cells and stress-related changes in hippocampal volume: a study in depressed patients and healthy controls. *PloS one.* 2015;10(3):e0119061.
33. Frodl T, Szyf M, Carballedo A, et al. DNA methylation of the serotonin transporter gene (SLC6A4) is associated with brain function involved in processing emotional stimuli. *Journal of psychiatry & neuroscience: JPN.* 2015;40(5):296.
34. Hass J, Walton E, Wright C, et al. Associations between DNA methylation and schizophrenia-related intermediate phenotypes—A gene set enrichment analysis. *Progress in Neuro-Psychopharmacology and Biological Psychiatry.* 2015;59:31-39.
35. Liu J, Julnes PS, Chen J, Ehrlich S, Walton E, Calhoun VD. The association of DNA methylation and brain volume in healthy individuals and schizophrenia patients. *Schizophrenia research.* 2015;169(1):447-452.
36. Jaffe AE, Gao Y, Deep-Soboslay A, et al. Mapping DNA methylation across development, genotype and schizophrenia in the human frontal cortex. *Nat. Neurosci.* Jan 2016;19(1):40-47.
37. Liu J, Demirci O, Calhoun VD. A Parallel Independent Component Analysis Approach to Investigate Genomic Influence on Brain Function. *IEEE Signal Processing Letters.* 2008;15:413.
38. Chen J, Calhoun VD, Pearlson GD, et al. Independent component analysis of SNPs reflects polygenic risk scores for schizophrenia. *Schizophrenia Research.* 2016.
39. Pearlson GD, Liu J, Calhoun VD. An Introductory review of parallel independent component analysis (p-ICA) and a guide to applying p-ICA to genetic data and imaging phenotypes to identify disease-associated biological pathways and systems in common complex disorders. *Frontiers in genetics.* 2015;6.
40. Xu L, Groth KM, Pearlson G, Schretlen DJ, Calhoun VD. Source-based morphometry: the use of independent component analysis to identify gray matter differences with application to schizophrenia. *Hum. Brain Mapp.* Mar 2009;30(3):711-724.

41. Gollub RL, Shoemaker JM, King MD, et al. The MCIC collection: a shared repository of multi-modal, multi-site brain image data from a clinical investigation of schizophrenia. *Neuroinformatics.* 2013;11(3):367-388.
42. Liu J, Ulloa A, Perrone-Bizzozero N, Yeo R, Chen J, Calhoun VD. A pilot study on collective effects of 22q13.31 deletions on gray matter concentration in schizophrenia. *PLoS One.* 2012;7(12):e52865.
43. Bustillo JR, Jones T, Chen H, et al. Glutamatergic and Neuronal Dysfunction in Gray and White Matter: A Spectroscopic Imaging Study in a Large Schizophrenia Sample. *Schizophr Bull.* Aug 22 2016.
44. Andreasen NC. *Scale for the Assessment of Positive Symptoms (SAPS).* Iowa City: University of Iowa;1984.
45. Andreasen NC. *Scale for the Assessment of Negative Symptoms (SANS).* Iowa City: University of Iowa;1984.
46. Cullen KR, Wallace S, Magnotta VA, et al. Cigarette smoking and white matter microstructure in schizophrenia. *Psychiatry Res.* Feb 28 2012;201(2):152-158.
47. Heatherton TF, Kozlowski LT, Frecker RC, Fagerstrom KO. The Fagerstrom Test for Nicotine Dependence: a revision of the Fagerstrom Tolerance Questionnaire. *Br. J. Addict.* Sep 1991;86(9):1119-1127.
48. Wechsler D. *WAIS-III: Wechsler adult intelligence scale.* Psychological Corporation San Antonio, TX; 1997.
49. Nuechterlein KH, Green MF, Kern RS, et al. The MATRICS Consensus Cognitive Battery, part 1: test selection, reliability, and validity. *American Journal of Psychiatry.* 2008;165(2):203-213.
50. Morris RG, Ahmed S, Syed G, Toone B. Neural correlates of planning ability: frontal lobe activation during the Tower of London test. *Neuropsychologia.* 1993;31(12):1367-1378.
51. Brown LB, Stern RA, Cahn-Weiner DA, et al. Driving scenes test of the Neuropsychological Assessment Battery (NAB) and on-road driving performance in aging and very mild dementia. *Archives of Clinical Neuropsychology.* 2005;20(2):209-215.
52. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics.* May 15 2014;30(10):1363-1369.
53. Houseman EA, Accomando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC bioinformatics.* 2012;13(1):1.
54. Gupta CN, Calhoun VD, Rachakonda S, et al. Patterns of gray matter abnormalities in schizophrenia based on an international mega-analysis. *Schizophrenia bulletin.* 2015;41(5):1133-1142.
55. Chen J, Liu J, Calhoun VD, et al. Exploration of scanning effects in multi-site structural MRI studies. *J Neurosci Methods.* Jun 15 2014;230:37-50.
56. Storey JD. A direct approach to false discovery rates. *Journal of the Royal Statistical Society: Series B (Statistical Methodology).* 2002;64(3):479-498.
57. Gibbs JR, van der Brug MP, Hernandez DG, et al. Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet.* 2010;6(5):e1000952.
58. Chen C, Cheng L, Grennan K, et al. Two gene co-expression modules differentiate psychotics and controls. *Molecular psychiatry.* 2013;18(12):1308-1314.
59. Hannon E, Lunnon K, Schalkwyk L, Mill J. Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes. *Epigenetics.* 2015;10(11):1024-1032.
60. Dreyer J, Schleicher M, Tappe A, et al. Nitric oxide synthase (NOS)-interacting protein interacts with neuronal NOS and regulates its distribution and activity. *The Journal of neuroscience : the official journal of the Society for Neuroscience.* Nov 17 2004;24(46):10454-10465.

61. Dedio J, Konig P, Wohlfart P, Schroeder C, Kummer W, Muller-Esterl W. NOSIP, a novel modulator of endothelial nitric oxide synthase activity. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology.* Jan 2001;15(1):79-89.
62. Reif A, Herterich S, Strobel A, et al. A neuronal nitric oxide synthase (NOS-I) haplotype associated with schizophrenia modifies prefrontal cortex function. *Mol Psychiatry.* Mar 2006;11(3):286-300.
63. Gerlach M, Blum-Degen D, Ransmayr G, Leblhuber F, Pedersen V, Riederer P. Expression, but not activity, of neuronal nitric oxide synthase is regionally increased in the alcoholic brain. *Alcohol and alcoholism.* 2001;36(1):65-69.
64. Yao JK, Leonard S, Reddy RD. Increased nitric oxide radicals in postmortem brain from patients with schizophrenia. *Schizophr Bull.* 2004;30(4):923-934.
65. Oliveira RM, Guimaraes FS, Deakin JF. Expression of neuronal nitric oxide synthase in the hippocampal formation in affective disorders. *Brazilian journal of medical and biological research = Revista brasileira de pesquisas medicas e biologicas / Sociedade Brasileira de Biofisica ... [et al.].* Apr 2008;41(4):333-341.
66. Karson CN, Griffin WS, Mrak RE, et al. Nitric oxide synthase (NOS) in schizophrenia: increases in cerebellar vermis. *Molecular and chemical neuropathology / sponsored by the International Society for Neurochemistry and the World Federation of Neurology and research groups on neurochemistry and cerebrospinal fluid.* Apr 1996;27(3):275-284.
67. Chen J, Tu Y, Moon C, Matarazzo V, Palmer AM, Ronnett GV. The localization of neuronal nitric oxide synthase may influence its role in neuronal precursor proliferation and synaptic maintenance. *Developmental biology.* May 01 2004;269(1):165-182.
68. Blaise GA, Gauvin D, Gangal M, Authier S. Nitric oxide, cell signaling and cell death. *Toxicology.* Mar 15 2005;208(2):177-192.
69. Chung KK, David KK. Emerging roles of nitric oxide in neurodegeneration. *Nitric Oxide.* 2010;22(4):290-295.
70. Pidsley R, Zotenko E, Peters TJ, et al. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biol.* Oct 07 2016;17(1):208.
71. Breitling LP, Salzmann K, Rothenbacher D, Burwinkel B, Brenner H. Smoking, F2RL3 methylation, and prognosis in stable coronary heart disease. *European heart journal.* Nov 2012;33(22):2841-2848.
72. Vergara VM, Ulloa A, Calhoun VD, Boutte D, Chen J, Liu J. A three-way parallel ICA approach to analyze links among genetics, brain structure and brain function. *Neuroimage.* 2014;98:386-394.

Figure legends

Figure 1. The analysis flow of the study.

Figure 2. (A) and (C) show meQTL mapping of the contributing SNPs and methylation sites (|z|>=2.5) in both pairs of significantly linked SNP and DNA methylation components. (B) and (E) plot their loadings with correlations r=0.48 and 0.35, shown separately by HC and SZ groups. (C) and (F) give examples of the strongest meQTL in each pair (located on chr.10 around gene *NT5C2* and chr.3 around gene *ITIH4*, respectively).

Figure 3. Brain mapping of nine GM components significantly associated with three methylation components. (A) Seven GM components have significant group differences while the other two in (B) show no significant differences.

Table 1. Demographic of participants in the both discovery and replication data

Variables	**Discovery data (n=202)**		**Replication data (n=192)**	
	SZ (Schizoaffective/Schizophreniform)	HC	SZ (Schizoaffective/Schizophreniform)	HC
Diagnosis	92 (1/4)	110	93 (11/0)	99
Gender* (women, %)	26.09	37.27	14.13	28.28
Age, y (SD)	34.14 (10.23)	32.66 (10.91)	37.96 (14.33)	37.09(12.16)
Race* (White/Black/ Asian/Other)	70/15/3/4	99/3/4/4	42/7/0/44	51/6/0/42
Alcohol depend/abuse	27	0	0	0

* indicates significant differences between SZ and HC. The differences were controlled or corrected in the analyses.

Table 2. The relationship between GM and DNA methylation components in the discovery data

Methylation components		**Associated gray matter imaging components**				
No.	Methylation-Diagnosis association Cohen's d (p, direction); *Significant genes*:	No.	Methylation-GM association R^2 (T-score, p)	GM-Diagnosis association Cohen's d (p, direction)	Brain regions (|z|>2.5)	L/R volume (mm^3)
1	-0.348 (0.009, HC>SZ); *NOSIP*	1	0.1 (4.77, 9.1×10^{-5})	-0.863 (7.7×10^{-8}, HC>SZ)	**Positive**: Superior/middle/medial/inferior Frontal Gyri, **Negative**: Precentral/Postcentral Gyri	8.6/11.8 3.5/0.8
		2	0.11 (5.02 , 3.8×10^{-5})	-0.79 (5.6×10^{-7}, HC>SZ)	**Positive**: Superior/middle/inferior Temporal Gyri,	7.7/7
		3	0.078 (4.11 , 9.7×10^{-4})	-0.551 (4.9×10^{-4}, HC>SZ)	**Positive**: Inferior Semi-Lunar Lobule, Uvula, Pyramis,	3.2/2.6, 1.0/1.0, 1.3/1.4
		4	0.071 (-3.89 , 2.0×10^{-3})	0.532 (5.7×10^{-4}, HC<SZ)	**Positive**: Lingual Gyrus, **Negative**: Inferior/middle Frontal Gyri, Lingual Gyrus, Insula	2.6/1.5, 1/1.6, 1.0/0.2, 0.3/0.6
		5	0.13 (5.45 , 1.5×10^{-5})	-0.418 (7.5×10^{-3}, HC>SZ)	**Positive**: Lingual Gyrus, Cuneus, Superior/Middle/inferior Occipital Gyri, Precuneus	1.7/2.0, 3.9/4.6, 4.9/4.6, 0.9/1.8
		6	0.098 (-4.65 , 1.2×10^{-4})	0.401 (9.4×10^{-3}, HC<SZ)	**Positive**: Insula, Precentral/Postcentral Gyrus, **Negative**: Inferior Parietal Lobule, Cingulate Gyrus, Superior/Middle temporal gyri	2.3/2.1, 1.2/1.4, 0.4/0.4, 1.7/2.0, 0.4/0.7
		7	0.11 (5.07 , 4.5×10^{-5})	-0.353 (0.02, HC>SZ)	**Positive**: Superior/Middle/Inferior Temporal Gyri, Middle/Inferior Frontal Gyri	5.5/4.7, 0.6/0.4
		8	0.05 (-3.25 , 0.015)	0.024 (non-significance)	**Positive**: Superior/Middle Temporal gyri, Supramarginal Gyrus, Superior/Inferior Parietal Lobule, Precuneus	3.6/2.7, 0.6/0.8, 1.4/2.1, 0.5/0.3
2 **3**	-0.211 (0.12, HC>SZ); *NT5C2*, *CSMD1*, *FES*, *ARL3* -0.296 (0.06, HC>SZ); *ITIH4*	9	0.038 (-2.82 , 0.05) 0.056 (3.44 , 8.7×10^{-3})	-0.091 (non-significance)	**Positive**: Declive, Pyramis, Vermis, Culmen, Thalamus, Caudate	1.4/1.5, 0.6/0.7, 0.5/0.5, 1.7/1.6, 0.3/0.6, 1.1/1.3,