Journal of Genetics (2021)100:44
https://doi.org/10.1007/s12041-021-01294-2

RESEARCH ARTICLE

Genomewide alteration of histone H3K4 methylation underlies genetic vulnerability to psychopathology

NICHOLAS NESBIT[1], RACHEL WALLACE[1], SOURABH HARIHAR[1], MILLIE ZHOU[1], JAE-YOON JUNG[2], MICAH SILBERSTEIN[1] and PHIL H. LEE[1,3,4]*

[1]*Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA 02114, USA*
[2]*Department of Pediatrics, Stanford University, San Francisco, CA 94305, USA*
[3]*Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA*
[4]*Department of Psychiatry, Harvard Medical School, Boston, MA 02215, USA*
*For correspondence. E-mail: PLEE0@mgh.harvard.edu.

Received 6 December 2020; revised 14 February 2021; accepted 15 February 2021

Abstract. Dysregulated histone methylation has emerged as a recurring theme in multiple neuropsychiatric disorders. However, it is yet unclear whether the altered histone methylation is associated with aetiologic mechanisms or an outcome of disease manifestation. In this study, we examined the genomewide association studies datasets of three major psychiatric disorders, schizophrenia (SCZ), bipolar disorder (BIP), and major depression disorder (MDD), which represents a total of 231,783 cases and 425,444 controls, to clarify the relationship. Our gene-set enrichment analysis results identified statistically significant association of genes involved in three histone methylation biological processes with the three adult-onset psychiatric disorders, which is mainly driven by the histone H3K4 methylation pathway (GO: 0051568). Further analysis of histone H3K4 methylation pathway genes revealed a widespread role of the genes in brain function and disease; 29 (52%) and 41 genes (73.2%) were associated with at least one brain-related trait or brain disorder, respectively. Spatiotemporal gene expression analysis suggests that these pathway genes play a critical role during the prenatal period and are consistent regulators in the cerebral cortex throughout an individual's life. *AUTS2*, *DNMT1* and *TET2* are genes of particular interest due to their pervasive role in various aspects of brain function. Our findings support a critical aetiologic role of H3K4 methylation genes shared across SCZ, BIP and MDD, providing new direction for the development of epigenetically-focussed drugs targeting common causal factors of these devastating disorders.

Keywords. H3K4 methylation; epigenetics; psychopathology; genomewide association studies; brain-disorder aetiology; cross-disorder analysis.

Introductions

Psychiatric disorders are a leading cause of major worldwide health concerns. In the US alone, more than 47.6 million adults, corresponding to nearly 20% of the adult population, live with mental illness (Substance Abuse and Mental Health Services Administration 2019). No cure exists, and thus many of the patients suffer from lifelong disability, impaired quality of life, and higher rates of medical comorbidity and mortality (Smeland *et al.* 2019). Yet, the quest to improve psychiatric care has been hindered mainly by our limited understanding of the aetiologic mechanisms and difficulties in disease classification (Network and Pathway Analysis Subgroup of Psychiatric Genomics Consortium 2015). Indeed, psychiatry is the only medical field in which no biologic or pathologic links are used in nosology, diagnosis, and treatments. The evolving nature of disease, classification, extensive polygenicity, and pervasive pleiotropy have created major challenges in understanding disease aetiology and developing new treatments.

In this study, we aim to clarify the aetiologic role of histone methylation pathway genes in major psychiatric disorders. Histone methylation acts as a central feature of

Supplementary Information: The online version contains supplementary material available at https://doi.org/10.1007/s12041-021-01294-2.

Published online: 21 June 2021

dynamic gene regulation by enabling rapid and precise adaptation of responses to abruptly changing environments (Cross-Disorder Group of the Psychiatric Genomics Consortium 2019). DNA forms a chromatin structure by wrapping around a histone octamer that contains two copies of the histone variants H2A, H2B, H3, and H4. Here, histones undergo different forms of modifications at various N-terminus sites of histone tails, specifically lysine and arginine, resulting in a covalent modification that can either increase or decrease gene expression (figure 1) (Cedar and Bergman 2009; Jambhekar *et al.* 2019). In recent years, various forms of dysregulated histone methylation have emerged as a recurring theme in multiple psychiatric disorders (Abrahams *et al.* 2013). Changes in histone 3 lysine 4 (H3K4) methylation have been reported at multiple neuronal signalling genes in schizophrenia (SCZ) and major depressive disorder (MDD) (Shen *et al.* 2014; Collins *et al.* 2019) and at synapsin genes in bipolar disorder (BIP) and MDD (Cruceanu *et al.* 2013). O'Dushlaine and colleagues (Network and Pathway Analysis Subgroup of Psychiatric Genomics Consortium 2015) reported the association of the H3K4 methylation pathway genes with major psychiatric disorders, along with neuronal and immune pathways, but the finding has not been replicated. Postmortem schizophrenic brain studies have reported increased levels of histone 3 lysine 27 trimethylation (H3K27me3) (Akbarian and Huang 2009; Akbarian 2010; Peter and Akbarian 2011). H3K27me3, specifically the region of a truncated form of tropomyosin-related receptor kinase type B called *Trk B-T1,* has also been observed in patients with depression and suicide victims (Dalton *et al.* 2014). Despite these intriguing findings, it is unclear whether altered histone methylation is associated with disease onset or is an outcome of disease manifestation itself.

Here we performed GWAS-based gene-set enrichment analysis to examine the association of histone methylation genes in psychiatric disorders. We focussed on three adult-onset psychiatric disorders, SCZ, BIP and MDD: all of which share a group of symptoms, such as mood instability, emotional withdrawal, cognitive impairment (e.g., sustained attention deficits) (American Psychological Association 2013; Zhu *et al.* 2019), known biological processes such as calcium channel activity (American Psychological Association 2013; Heyes *et al.* 2015; Zhu *et al.* 2019), and a significant level of genomewide genetic risk (Cross-Disorder Group of the Psychiatric Genomics Consortium 2019). We specifically examined: (i) whether the risk genes associated with psychiatric disorders are enriched among genes involved in specific histone methylation pathways; (ii) whether this association is specific to psychiatric disorders, or other neuropsychiatric, cognitive, and behavioural traits; and (iii) whether the associated histone methylation genes show heightened gene expressions specific to certain brain regions, cell-types, or developmental stages.

Materials and methods

GWAS datasets

GWAS summary association statistics of three adult-onset psychiatric disorders were downloaded from the Psychiatric Genomics Consortium: SCZ , *n*=40,675 cases and 64,643 controls; BIP, *n* =20,352 cases and 31,358 controls; and

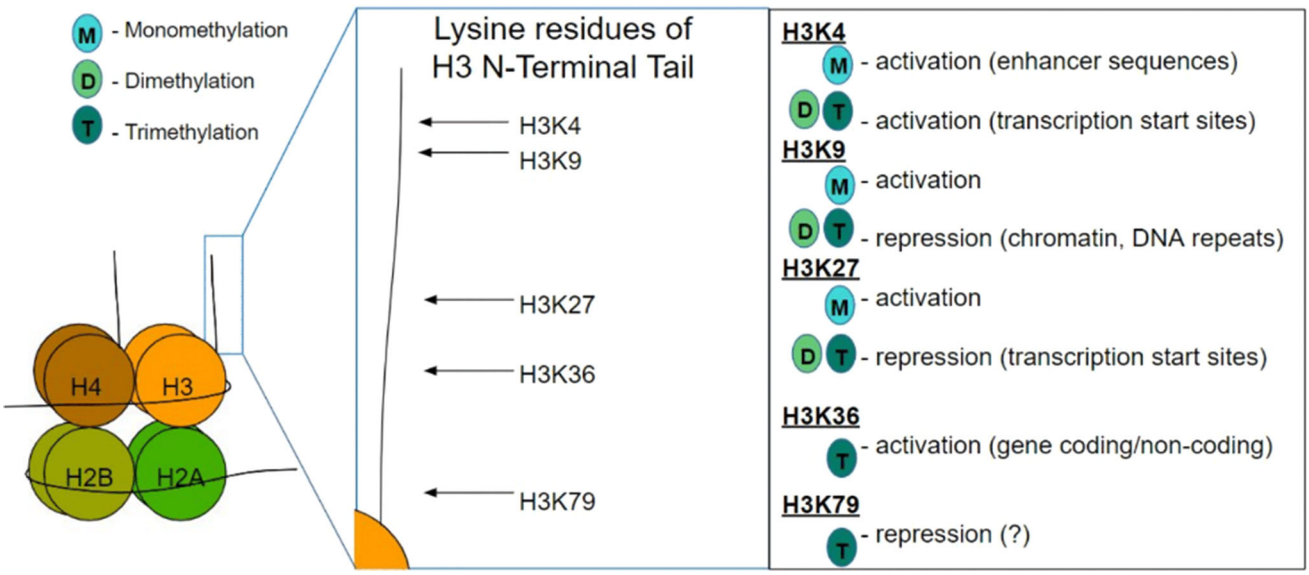


Figure 1. Representation of locations and effects of known H3 N-terminal tail methylation sites. Methylation of n-terminal tails can take place at various lysine sites and occur either as mono-, di-, or trimethylation. Whether histone lysine methylation increases or decreases gene expression varies by gene, lysine and methylation type, as in the case of gene-specific instances of histone 3 lysine 9 (H3K9) where monomethylation results in activation but di- or trimethylation result in repression.

MDD; n =170,756 cases and 329,443 controls. The compiled studies represented a total of 425,444 healthy volunteers and 231,783 patients (table 1 in electronic supplementary material at http://www.ias.ac.in/jgenet/). Chromosome coordinates of all association summary statistics were standardized to the reference genome GRCh37/hg19. We removed rare SNPs (minor allele frequency (MAF) < 1%), nonbiallelic SNPs, those without rs IDs, and those not in 1000 Genomes Project Phase 3. The alleles of SNPs which did not match the 1000 Genomes data on both strands were also excluded. Palindromic SNPs were removed if no strand information was available and the differences of allele frequency between the two alleles were less than 15%.

Gene-set enrichment analysis of histone methylation pathway genes

We downloaded gene ontology (GO) gene sets from the MSigDB database (v.7.1) and found 47 GO gene sets related to histone methylation (table 2 in electronic supplementary material). Gene information including the genomic locus, strand, gene type, description, and biological function was obtained from the NCBI Entrez (table 3 in electronic supplementary material). We limited genes to chromosomes 1–22, X, and Y, with Entrez IDs. Next, we estimated the number of causal risk variants underlying each psychiatric disorder using MiXeR (Frei *et al.* 2019). While incorporating the effects of linkage disequilibrium (LD) structure, MAFs, sample sizes and cryptic relationships, the method estimated the number of disease-risk variants as 8.3k, 6.4k, and 14.9k for SCZ, BIP, and MD, respectively. Disease-associated risk loci for each disorder were then generated using PLINK (Chang *et al.* 2015) LD clumping (r^2=0.4, kb=500). We conducted gene set enrichment analyses of disease-associated risk loci for each psychiatric disorder using INRICH (Lee *et al.* 2012), a GWAS-specific analysis method that adjusts for various genomic confounding factors including SNP LD, gene density, and sizes. For specificity, we restricted gene-set enrichment analysis to 61 gene sets of which the number of genes ranged from 5 to 1000. Meta-analysis of gene set enrichment analysis results across disorders was conducted using the Fisher's method implemented in the R *metap* library.

Investigation of cognitive, behavioural and brain disorders

For methylation gene sets with significant association with three psychiatric disorders, we examined whether any brain-related traits have been associated with the annotated genes using the GWAS catalog data. First, we downloaded the full GWAS catalog data (v1.0.2; on Jan 2020). We searched 'mapped genes' or 'reported genes' columns, which list gene symbols that are mapped to or reported for the strongest SNP of associated risk loci. The GWAS catalog data included genotype–phenotype association data for 3643 traits. Following the GWAS catalog, we defined the seven domains of brain-related traits using the European Bioinformatics Institute Experimental Factor Ontology (EFO) data. Domains of brain-related traits were defined as: (i) mental process (e.g., intelligence); (ii) emotion (e.g., irritability); (iii) behaviour (e.g., risk-taking behaviour); (iv) psychiatric disorder (e.g., anxiety disorder); (v) brain diseases (e.g., Alzheimer's disease); (vi) brain-related measurement (e.g., brain volume measurement); and (vii) multi-domain studies that analyse traits in different domains together (e.g., schizophrenia and intelligence). There were 94 brain-related traits in the seven domains. For each of the trait-associated genes, we examined as how many of the genes overlapped with our methylation gene sets. We calculated the statistical significance of the overlap by randomly selecting the same number of genes 10,000 times and by counting how many of the random overlap is at least the same as the original overlap observed for the methylation gene set. False discovery rates of 10% was used to identify statistically significant overlaps.

We also investigated various bioinformatics databases that provide rare-variant-based evidence for gene–phenotype associations. The resources we used include: online Mendelian inheritance in man (OMIM) (Amberger *et al.* 2019), ClinVar (Landrum *et al.* 2018), GeneCards (Stelzer *et al.* 2016), and the genetic testing registry (GTR) (Rubinstein *et al.* 2013). We also examined disease-specific databases, SFARI (Abrahams *et al.* 2013), VariCarta (Belmadani *et al.* 2019), and SZBD (Wu *et al.* 2017), and literatures (Rhoades *et al.* 2019) that reported sequencing-based evidence for histone H3K4 methylation pathway-associated genes. SFARI and VariCarta gene databases are dedicated to integrating information about autism spectrum disorder (ASD), which showed significant genomewide genetic correlations with SCZ, BIP, and MD (Cross-Disorder Group of the Psychiatric Genomics Consortium 2019). Both databases utilize peer-reviewed publications to build a library of in-depth knowledge about genes associated with ASD. SZBD provides a collection of 7377 schizophrenia-associated genes identified through integration of various sources, including CNVs, exons, GWAS, and literature.

Spatiotemporal brain gene expression analysis

We also performed temporal analysis of brain-expressed genes using the Human Brain Transcriptome (https://hbatlas.org/pages/hbtd) (Kang *et al.* 2011). This database provides spatial and temporal gene expression analysis of over 1340 specimens gathered from 16 areas of the postmortem human brain: cerebellar cortex, mediodorsal nucleus of the thalamus, striatum, amygdala, hippocampus, and 11 areas of the neocortex. Study samples span individuals of both genders and all ethnicities from four post-conceptual weeks of age

through adulthood and are divided into 15 period groups. Each of our genes of interest was entered into the database, and spatiotemporal output was downloaded and analysed for expression patterns. For gene expression analysis, we observed the trends of six different represented brain areas and defined high prenatal as a decrease of ≥1 in signal intensity in four or more brain areas after birth, while high postnatal was defined as an increase of ≥1 in signal intensity in four or more brain areas after birth. Expression patterns without changes significant enough to meet high or low expression thresholds were defined as uniform throughout. We further analysed spatiotemporal data using the brain expression spatiotemporal pattern (BEST) (http://best.psych.ac.cn/) (Guo *et al.* 2019). This database provides spatial and temporal coexpression data for brain-related genes based on eight postmortem brain datasets, including the Allen Brain Atlas (Hawrylycz *et al.* 2012) and BrainSpan (Miller *et al.* 2014). The average expression level of each input gene is calculated with either a 16x10 or 25x9 spatiotemporal matrixes, with the first number representing brain regions and the second number representing time periods (Johnson *et al.* 2016). Coexpression networks and heatmaps are based upon 525 RNA-seq data samples from BrainSpan and illustrate the top 20 input genes by network connectivity in the top five enriched clusters by adjusted *P* value (Langfelder and Horvath 2008).

Brain cell-type-specific data analysis

We examined whether methylation-related genes are expressed in the brain using the GeneCards database. For brain-expressed genes, we investigated whether they show cell-type specific expression activities in a brain using a web-based data resource (http://celltypes.org/brain/). This database provides RNA expression data (MacArthur *et al.* 2017) for six brain cell-types: astrocyte, endothelial, microglia, neuron, oligodendrocyte and OPC. We used human cortical sample data generated by Darmanis *et al.* (2015). Exon-level analysis was conducted by entering the number of exons for each gene, which was obtained from the Ensembl database (Hunt *et al.* 2018). To measure the statistical significance of cell-type-specific gene expression, we ran a permutation test as follows. In each simulation run, gene expression data were randomly permuted across cell-types for each gene. This reserves the number of single cell-type gene expression for each gene, while randomizing the cell-type membership. Then for each cell-type, we counted the number of genes that were expressed in the cell-type. This was repeated 1000 times. For each of the six cell-types, the statistical significance of enrichment was measured by comparing the number of originally observed gene expression to those counted in permutated datasets.

Results

Enrichment of H3K4 methylation pathway genes in three-adult-onset psychiatric disorders

We assembled GWAS datasets for three adult-onset psychiatric disorders, SCZ, BIP and MDD, which represent a total of 231,783 cases and 425,444 controls (table 1 in electronic supplementary material). Among 47 GO terms related to histone methylation, we found statistically significant association of three GO biological process (BP) and one cellular component (CC) gene sets across the three psychiatric disorders (FDR<1%). The top 15 findings are summarized in table 1 (full results is provided in table 2 in electronic supplementary material). Three significant BP gene sets included: (i) 'histone H3K4 methylation' (GO:0051568, gene-set association meta-analysis *P* value = 2.51×10^{-5}, FDR *q* = 0.0012); (ii) 'histone methylation' (GO:0016571, *P* = 2.82×10^{-5}, *q* = 0.0013); and (iii) 'regulation of histone

Table 1. Pathway analysis results of histone methylation genes across SCZ, BIP and MDD.

Category	Gene set	metaP	FDR
bp	**HISTONE H3 K4 METHYLATION**	**2.51E-05**	**1.18E-03**
	Histone methylation	**2.82E-05**	**1.33E-03**
	Regulation of histone methylation	**8.77E-05**	**4.12E-03**
	Regulation of histone H3 K4 Methylation	2.80E-03	1.32E-01
	Negative regulation of histone H3 K9 methylation	3.41E-03	1.60E-01
	Negative regulation of histone methylation	3.50E-03	1.65E-01
	Histone H3 K9 methylation	9.07E-03	4.26E-01
	Positive regulation of histone methylation	9.66E-03	4.54E-01
	Regulation of histone H3 K9 methylation	3.48E-02	1.00E+00
	Histone H3 K9 demethylation	4.03E-02	1.00E+00
cc	**Histone methyltransferase complex**	**4.68E-04**	**2.20E-02**
mf	Histone methyltransferase activity	1.42E-03	6.68E-02
	Methylated histone binding	2.35E-03	1.10E-01
	Histone lysine n methyltransferase activity	6.98E-03	3.28E-01
	Histone demethylase activity	3.74E-02	1.00E+00

Bold text indicates gene sets of which meta-analysis association is significant at false discovery rate (FDR) < 0.05.

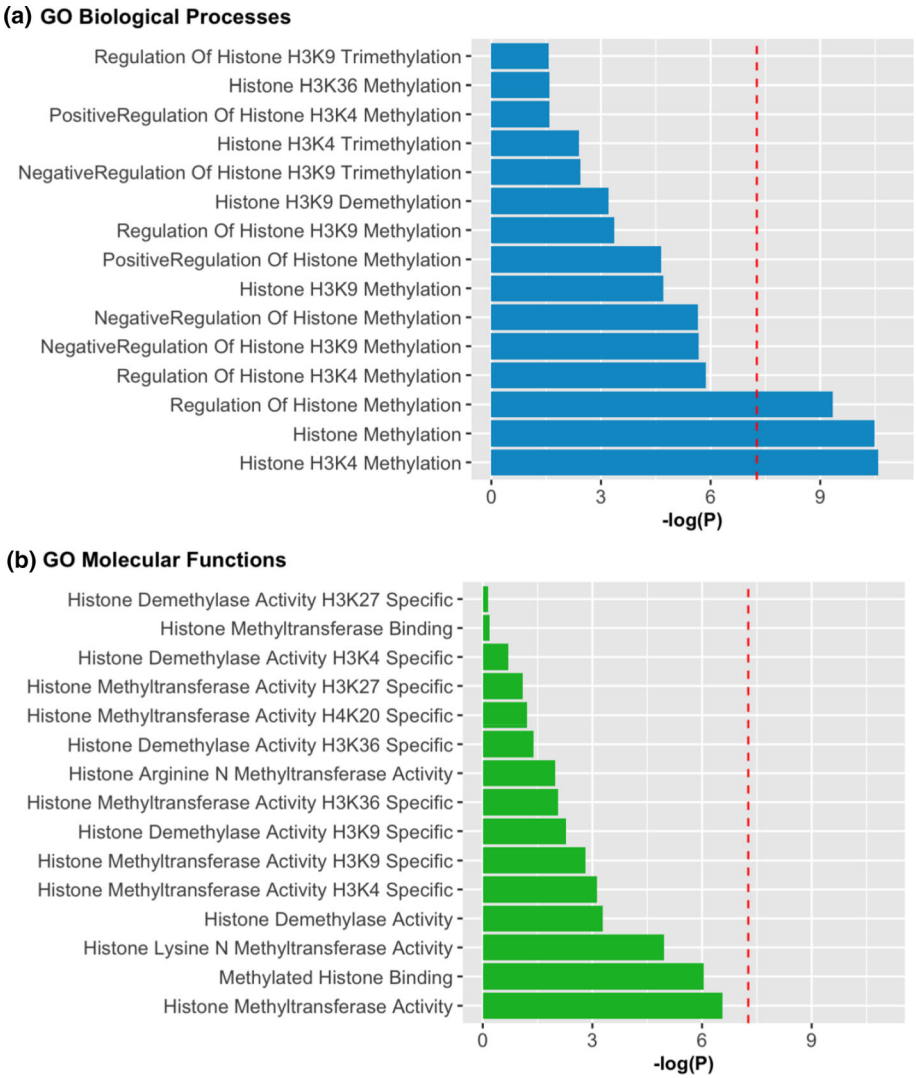
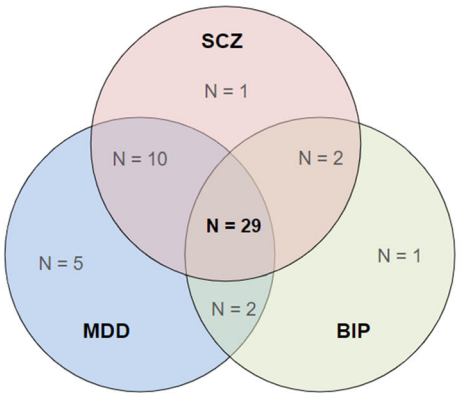


Figure 2. Analysis of GO biological processes related to histone methylation. Threshold for significance is indicated with a dashed red line at -log(0.0007). (a) Top 10 BF gene sets ordered by extent of association based on -log(p). The most significant gene-set is *histone H3K4 methylation*, followed by *histone methylation*, and *regulation of histone methylation*. (b) Top 10 MF gene sets ordered by extent of association based on -log(p). No MF gene sets meet the criteria for significance.

methylation' (GO:0031060, $P = 8.77\text{x}10^{-5}$, $q = 0.0041$) (figure 2a). We found no significant enrichment for gene sets under GO molecular functions (figure 2b). Association of the top three histone methylation pathway gene sets was more notable for two mood disorders, BIP and MDD, compared to SCZ (table 3 in electronic supplementary material). Comparison of the top three pathway genes indicated that association of the latter two gene sets, 'histone methylation' and 'regulation of histone methylation' was largely driven by genes involved in 'histone H3K4 methylation' (figure 1 in electronic supplementary material). Cross-disorder comparison showed that 50 of 56 genes shared across SCZ, BIP and MDD, showed significant levels of enrichment. Forty three of those 50 genes were implicated in at least two disorders, and 29 of 50 genes were implicated in all three disorders. Eighteen of these 29 overlapping genes play a role in the epigenetic processes of chromatin compaction ($n = 3$) or the methylation of DNA ($n = 3$), histone ($n = 5$), or histone lysine ($n = 7$) (figure 3). Thus, we further focussed analyses on 'histone H3K4 methylation'. Excluding the MHC regions (chr6: 25–35 Mb) resulted in consistent findings (table 4 in electronic supplementary material).

Implication of H3K4 methylation genes in brain-related traits

We hypothesized that pleiotropic association of histone H3K4 methylation pathway (GO: 0051568) genes across three adult-onset psychiatric disorders may indicate their broad impact on brain function and structure. To test the hypothesis, we examined whether common single-nucleotide polymorphisms (SNPs) in histone H3K4 methylation pathway genes had shown association with brain-related traits using the GWAS catalog data (MacArthur *et al.* 2017).



Functional Mechanism of Genes Shared by SCZ, BIP, and MDD (N = 29)

Cellular Regulation
CTNNB1

Chromatin Compaction
HIST1H1C
HIST1H1D
HIST1H1E

DNA Methylation
DNMT3B
TET2
TET3

Histone Methylation
DPY30 DYDC2
PAGR1 SETD1A
SETD1B

Histone Lysine Methylation
ASH1L KMT2B
KMT2E RBBP5
WDR82 WDR5
WDR5B

Pathway Signal Transduction
PYGO2

RNA/rRNA Transcription
AUTS2 CTCFL
CXXC1 RTF1
GFI1 NELFE
PIH1D1 WDR61
ZNF335

Figure 3. Cross-disease comparison of genes associated with three significant GO biological processes shows a strong correlation between SCZ, BIP and MDD. Each circle in the diagram represents gene sets for one of the three diseases. N represent number of genes, with n values in areas where circles overlap representing genes common to two or all three diseases. Twenty-nine genes are associated with all three, and these genes are categorized by functional mechanism in the text to the right of the diagram.

We used European Bioinformatics Institute Experimental Factor Ontology (EFO) to define seven domains of brain-related traits: (i) mental processes (e.g., intelligence); (ii) emotion-related phenotypes (e.g., irritability); (iii) behaviours (e.g., risk-taking behaviour); (iv) psychiatric disorders (e.g., anxiety disorder); (v) brain diseases not including psychiatric disorders (e.g., Alzheimer's disease); (vi) brain-related measurement (e.g., brain volume measurement); and (vii) multi-domain phenotypes that combine traits in distinct domains together (e.g., schizophrenia and intelligence).

We found implications of histone H3K4 methylation genes in all of the seven domains, most notably, cognition, behaviour, emotion, brain morphology, and mental health (table 5 in electronic supplementary material at http://www.ias.ac.in/jgenet/). We summarize brain-related traits associated with more than one H3K4 methylation gene in table 2. Notable association was found for 'intelligence' in which 10 of 56 histone H3K4 methylation pathway genes had been implicated (17.6%). Simulation studies indicated this overlap is highly unlikely to occur by a random chance ($P = 1.19 \times 10^{-3}$). Significant associations were also found for risk-taking behaviours (overlapping gene $n = 6$, $P = 3.99 \times 10^{-4}$), neuroticism ($n = 4$, $P = 2.99 \times 10^{-4}$), and reaction time measurement ($n = 3$, $P = 2.99 \times 10^{-4}$). As expected, multiple histone H3K4 methylation pathway genes had shown genomewide significant association in the GWAS catalog with adult-onset psychiatric disorders, including BIP ($n = 4$, $P = 1.99 \times 10^{-4}$), SCZ ($n = 3$, $P = 9.99 \times 10^{-5}$), and insomnia ($n = 3$, $P = 1.99 \times 10^{-4}$).

We also examined various bioinformatics resources and published literature to check whether rare inherited or *de-novo* exonic mutations in histone H3K4 methylation genes had contributed to the risk of mental health conditions. The results are summarized in table 6 in electronic supplementary material. We found 41 out of 56 histone H3K4 methylation pathway genes have been implicated in at least one brain disorder (73.2%), with DNA methyltransferase 1 (*DNMT1*) having the highest number of associated disorders (n=17). *DNMT1* showed association with narcolepsy, hereditary sensory neuropathy, neuropathy, dementia, bipolar disorder, autosomal dominant cerebellar ataxia, hereditary sensory and autonomic neuropathy type 1, mood disorder, epilepsy, foetal alcohol spectrum disorder, Rett syndrome, autism, pervasive developmental disorder, Charcot–Marie tooth disease/*DNMT1*-related disorder1, and SCZ (Rubinstein *et al.* 2013; Stelzer *et al.* 2016; Wu *et al.* 2017). A cross comparison of the disorders associated with the 56 histone H3K4 pathway genes showed some striking pathological patterns. Of the 56 genes, nine were associated with a form of mental retardation (16.1%), 10 were associated with Kleefstra syndrome (17.6%), 11 were associated with microcephaly (19.6%), 14 were associated with schizophrenia (25%), and 16 were associated with Kabuki syndrome (28.6%).

Spatiotemporal and cell-type-specific brain expression activity of histone H3K4 methylation genes

We examined how the expression of 56 H3K4 methylation genes were regulated in different brain areas throughout the lifespan using BrainSpan data (Kang *et al.* 2011). The postmortem brain data were gathered from the cerebellar cortex, mediodorsal nucleus of the thalamus, striatum, amygdala, hippocampus, and 11 areas of the neocortex. Overall, among the 52 H3K4 methylation genes for which gene expression data was available, 35 showed higher expression during the prenatal period compared to postnatal (67.3%), while one gene showed an opposite trend (1.9%). The remaining 16 genes showed uniform expression

Table 2. Brain-related traits related to histone H3-K4 methylation.

Domain	Brain-related traits	Number of genes	Genes
behaviour	Alcohol consumption measurement	3	*ARID4A*, *AUTS2*, *TET3*
	Coffee consumption measurement	2	*AUTS2*, *TET2*
	Neuroticism measurement	4	*HIST1H1C*, *HIST1H1E*, *KMT2A*, *SETD1A*
	Risk-taking behaviour	6	*AUTS2*, *GATA3*, *HISTH1E*, *KMT2E*, *TET2*, *TET3*
	Smoking initiation	2	*AUTS2*, *DNMT3B*
	Smoking status measurement	5	*AUTS2*, *DNMT3B*, *HISTH1E*, *SETD3*, *TET3*
	Social interaction measurement	2	*GCG*, *TET2*
Brain disease	Parkinson's disease	2	*SETD1A*, *SMAD4*
Brain-related measurement	Brain volume measurement	2	*AUTS2*, *HIST1H1E*
	Chronotype measurement	2	*AUTS2*, *KMT2D*
Mental process	Intelligence	10	*AUTS2*, *HIST1H1C*, *HIST1H1E*, *KMT2D*, *KMT2E*, *PRMT6*, *RTF1*, *TET2*, *WDR82*, *ZNF335*
	Reaction time measurement	3	*AUTS2*, *KMT2D*, *TET2*
	Self-reported educational attainment	2	*AUTS2*, *TET2*
Multi-domain	Well-being measurement	2	*AUTS2*, *KMT2A*
Psychiatric disorder	Bipolar disorder	4	*AUTS2*, *KMT2D*, *PYGO2*, *WDR82*
	Schizophrenia	3	*ASH2L*, *KMT2E*, *WDR82*
	Schizophrenia/autism spectrum disorder	6	*HIST1H1C*, *HIST1H1D*, *HIST1H1E*, *KMT2E*, *NELFE*, *WDR82*

throughout the lifespan (30.8%) (table 7 in electronic supplementary material). The cerebral cortex exhibited higher general uniformity throughout the lifespan across all examined genes when compared to other brain areas. Activator of transcription and developmental regulator AUTS2 (*AUTS2*), tet methylcytosine dioxygenase 2 (*TET2*), and *DNMT1*, which were associated with the highest levels of brain-related traits and psychiatric disorders, all exhibited higher prenatal expression compared to postnatal. Coexpression analysis derived from BEST correlated to these spatial and temporal findings for all genes for which data was available (figures 2 and 3 in electronic supplementary material) (Guo *et al.* 2019).

We examined the levels of expression for each of the 56 H3K4 methylation genes in six specific brain cell-types: astrocytes, endothelial cells, microglia, neurons, oligodendrocytes, and oligodendrocyte progenitor cells. We found no enrichment of single cell-type-specific expression for H3K4 methylation genes ($P>0.05$). Of the 47 genes for which data was available, close to half of the genes were expressed in at least five cell-types, including *AUTS2* and *TET2* (47.8%) (table 8 in electronic supplementary material).

Discussion

Genetic and epigenetic regulatory mechanisms, particularly histone methylation, play a critical role in brain development and function. Previous studies have shown abnormal H3K4 methylation disrupts normal skeletal and neurological development, which can lead to conditions with cognitive defects such as Kabuki syndrome and microcephaly (Jambhekar *et al.* 2019). Using GWAS data representing more than 657,000 individuals, we showed that genes involved in histone H3K4 methylation (GO: 0051568) are shared between three major adult-onset psychiatric disorders: MDD, BIP, and SCZ. Brain gene expression data analysis indicated that these genes are specifically highly expressed during prenatal encephalic development, suggesting genetically-driven variations in histone H3K4 methylation may increase MDD, BIP, and SCZ susceptibility through disruption of important neurodevelopmental mechanisms. We did not find statistically significant association of other types of histone methylation mechanisms, including H3K27, H3K9, and H3K36, despite prior studies suggesting their role in psychiatric disorders (Akbarian and Huang 2009; Akbarian 2010; Peter and Akbarian 2011; Hyun *et al.* 2017; Jambhekar *et al.* 2019).

We found robust evidence supporting the implication of histone H3K4 methylation pathway genes in various brain disorders and traits. Overall, 29 of 56 histone H3K4 methylation pathway genes had at least one publication reporting genotype–phenotype association with brain-related traits (52%). Two genes, *AUTS2* and *TET2,* were found to be of particular interest. *AUTS2* was associated with 18 brain-traits, which encompassed all seven domains we studied, while *TET2* was associated with 11 brain-traits in six domains, including behaviour, brain-related measurement, emotion, mental process, multi-domain, and psychiatric disorder. Both genes were disproportionately associated with behavioural traits (seven of 18 for *AUTS2* and five of 11 for *TET2*) related to sleep, smoking, coffee consumption, and risk-taking behaviours. Other traits with which both *AUTS2* and *TET2* showed association are related to mental processes, including intelligence, reaction time, and self-reported educational attainment. Of the 41 genes that

implicated in at least one brain disorder, *DNMT1* ranked highest with association to 17 individual disorders. *AUTS2* was associated with 14 disorders (tables 6 and 9 in electronic supplementary material).

Three H3K4 methylation genes, *AUTS2*, *DNMT1* and *TET2,* which are highlighted based on their association with broad brain-related traits play vital functional roles in early brain development. *AUTS2* consists of the polycomb group (PcG) multiprotein PRC1-like complex which modifies histones and chromatin structures during prenatal encephalic development. It shows significant expression in the frontal, parietal, and temporal regions of the foetal brain and several neuronal subtypes to include glutamatergic, GABAergic, and tyrosine hydroxylase (TH)-positive dopaminergic neurons (Oksenberg and Ahituv 2013; Stelzer *et al.* 2016). Spatiotemporal data and association of *AUTS2* copy number variation with multiple neurodevelopmental disorders suggest a role in neuronal maturation or neocortical differentiation (Oksenberg and Ahituv 2013), which is consistent with single cell data analyses. In our brain single cell-data analysis, expression of *AUTS2* was evident in various brain cell types including endothelial cells, neurons, oligodendrocytes, and oligodendrocyte progenitor cells, but markedly high expression was most notable in astrocytes (Darmanis *et al.* 2015) (table 8 in electronic supplementary material). There is strong evidence for links between abnormalities in astrocyte function and SCZ, with cell-specific hyperresponsiveness and morphological abnormality, brain-area specific changes in astrocyte density, and reduction in glutamate transporters in astroglia as just some aspects associated with the disorder (McCullumsmith *et al.* 2016; Tarasov *et al.* 2019). A strong body of research has emerged over the last 15 years showing that astrocytes are capable of producing and reacting to neuroactive signalling with neurons, and that, they heavily influence critical formation and function of synapses and neural circuitry in early life. In correlation with our spatiotemporal data on *AUTS2*, astrocyte activity peaks during prenatal periods and continues to be high during postnatal synaptogenesis and synaptic pruning (Petrelli *et al.* 2016). Studies have shown that forebrain regulation of protein–protein interaction (PPI) and sensorimotor gating is dependent on functional neural circuitry, and disruption of such circuitry is a feature of some SCZ phenotypes (Swerdlow and Light 2018). Disruption in neurocircuitry has also been evidenced by fMRI in MDD, BIP, and other mood disorders (Wessa and Lois 2015; Horowitz-Kraus *et al.* 2018), suggesting *AUTS2* disruption in astrocytes could contribute to brain-disorder aetiology through dysregulation of developmental synapse formation.

The protein encoded by *TET2* is involved in histone H2B GlcNAcylation through recruitment of OGT to transcription start sites of active genes containing CpG sites. *TET2* also plays an active role in DNA cytosine demethylation through the multi-step conversion of methyl-cytosine to 5-hydroxymethylcytosine (5hmC), 5hmC to 5-formylcytosine (5fC), and 5fC to 5-carboxylcytosine (5caC). Contribution of *TET2* to the demethylation of cytosine has broad developmental implications, as methylation of the C5 position of cytosine bases are heavily involved in transcriptional regulation (Stelzer *et al.* 2016). Unlike *AUTS2*, we found specifically higher expression of *TET2* in microglia compared to other cell types. Microglia primarily serves to counteract brain-tissue damage and pathogens through macrophage, and microglial phagocytic activity is highly involved in the postnatal synaptic pruning of presynaptic and postsynaptic elements (Petrelli *et al.* 2016). Thus, the same SCZ, BIP and MDD disruption of neural circuitry associated with dysregulated astrocyte function could be impacted by changes in microglia function. This correlates to our data showing consistently high *TET2* expression in both the prenatal period and the microglial cell type, evidencing a link between gene disruption and brain-disease aetiology.

DNMT1 produces an enzyme that transfers methyl groups to genomic DNA cytosine and is highly involved in maintaining methylation patterns following DNA replication. *DNMT1* genetic mutation leads to protein variation, resulting in dysregulated methylation of other genes and subsequent development of *DNMT1*-associated neurological disorders. Recent research suggests such dysregulation can be corrected by a *TET2-DNMT1* complex which downgrades abnormally elevated levels of CpG site methylation during DNA repair following oxidative stress (Baets *et al.* 2015; Stelzer *et al.* 2016; Zhang *et al.* 2017). Prior research has identified altered methylation patterns of specific genes in cases of SCZ, MDD and BIP (Ruzicka *et al.* 2015; Hannon *et al.* 2016; Chen *et al.* 2017), evidencing the shared aetiological role of prenatal *DNMT1* expression in these disorders.

Our study highlights the shared aetiology of H3K4 methylation during early brain development across SCZ, MDD and BIP. Yet much progress remains to be made, as genetic contributors to dysregulation of histone methylation are known to be compounded by environmental perturbations including maternal immune activation and exposure to neuroactive drugs (Dong *et al.* 2018; Nestler *et al.* 2016). At this time, treatment of SCZ, BIP and MDD is largely limited to behavioural interventions and symptom management (Collins *et al.* 2019; Annette *et al.* 2020; Miklowitz *et al.* 2020). However, as contemporary advances are made into highly-selective pharmacological interventions that can cross the blood–brain barrier, direct enzymatic manipulation has become an increasingly possible treatment option for dysregulated H3K4 methylation in psychiatric disorders (Ricq *et al.* 2016).

References

Abrahams B. S., Arking D. E., Campbell D. B., Mefford H. C., Morrow E. M., Weiss L. A. *et al.* 2013 SFARI Gene 2.0: a community-driven knowledgebase for the autism spectrum disorders (ASDs). *Mol. Autism* **4**, 36.

Akbarian S. 2010 The molecular pathology of schizophrenia–focus on histone and DNA modifications. *Brain Res. Bull.* **83**, 103–107.
Akbarian S. and Huang H. S. 2009 Epigenetic regulation in human brain-focus on histone lysine methylation. *Biol. Psychiatry* **65**, 198–203.
Amberger J. S., Bocchini C. A., Scott A. F. and Hamosh A. 2019 OMIM.org: leveraging knowledge across phenotype-gene relationships. *Nucleic Acids Res.* **47**, D1038–D1043.
American Psychological Association 2013 Diagnostic and Statistical Manual of Mental Disorders. Diagnostic and Statistical Manual of Mental Disorders, 5th edition. American Psychological Association, Arlington.
Annette S., Stephan G., Mueser K. T., Martin H., Elisabeth R., Ulrich G. *et al*. 2020 A 2-year longitudinal study of neuropsychological functioning, psychosocial adjustment and rehospitalisation in schizophrenia and major depression. *Eur. Arch. Psychiatry Clin. Neurosci.* **270**, 699–708.
Belmadani M., Jacobson M., Holmes N., Phan M., Nguyen T., Pavlidis P. and Rogic S. 2019 VariCarta: A comprehensive database of harmonized genomic variants found in autism spectrum disorder sequencing studies *Autism Res.* **12**, 1728-1736.
Cedar H. and Bergman Y. 2009 Linking DNA methylation and histone modification: patterns and paradigms. *Nat Rev. Genet.* **10**, 295–304.
Chang C. C., Chow C. C., Tellier L. C., Vattikuti S., Purcell S. M. and Lee J. J. 2015 Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7.
Chen D., Meng L., Pei F., Zheng Y. and Leng J. 2017 A review of DNA methylation in depression. *J. Clin. Neurosci.* **43**, 39–46.
Collins B. E., Greer C. B., Coleman B. C. and Sweatt J. D. 2019 Histone H3 lysine K4 methylation and its role in learning and memory. *Epigenet. Chromatin* **12**, 7.
Cross-Disorder Group of the Psychiatric Genomics Consortium 2019 Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. *Cell* **179**, 1469–1482.
Cruceanu C., Alda M., Nagy C., Freemantle E., Rouleau G. A. and Turecki G. 2013 H3K4 tri-methylation in synapsin genes leads to different expression patterns in bipolar disorder and major depression. *Int. J. Neuropsychopharmacol.* **16**, 289–299.
Dalton V. S., Kolshus E. and McLoughlin D. M. 2014 Epigenetics and depression: return of the repressed. *J. Affect Disord.* **155**, 1–12.
Darmanis S., Sloan S. A., Zhang Y., Enge M., Caneda C., Shuer L. M. *et al*. 2015 A survey of human brain transcriptome diversity at the single cell level. *Proc Natl. Acad. Sci. USA* **112**, 7285.
Dong E., Guidotti A., Zhang H. and Pandey S. C. 2018 Prenatal stress leads to chromatin and synaptic remodeling and excessive alcohol intake comorbid with anxiety-like behaviors in adult offspring. *Neuropharmacology* **140**, 76–85.
Frei O., Holland D., Smeland O. B., Shadrin A. A., Fan C. C., Maeland S. *et al*. 2019 Bivariate causal mixture model quantifies polygenic overlap between complex traits beyond genetic correlation. *Nat. Commun.* **10**, 2417.
Guo L., Lin W., Zhang Y., Li W. and Wang J. 2019 BEST: a web server for brain expression Spatiotemporal pattern analysis. *BMC Bioinformatics* **20**, 632.
Hannon E., Spiers H., Viana J., Pidsley R., Burrage J., Murphy T. M. *et al*. 2016 Methylation QTLs in the developing brain and their enrichment in schizophrenia risk loci. *Nat. Neurosci.* **19**, 48–54.
Hawrylycz M. J., Lein E. S., Guillozet-Bongaarts A. L., Shen E. H., Ng L., Miller J. A. *et al*. 2012 An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391–399.
Heyes S., Pratt W. S., Rees E., Dahimene S., Ferron L., Owen M. J. and Dolphin A. C. 2015 Genetic disruption of voltage-gated calcium channels in psychiatric and neurological disorders. *Prog. Neurobiol.* **134**, 36–54.
Horowitz-Kraus T., Woodburn M., Rajagopal A., Versace A. L., Kowatch R. A., Bertocci M. A. *et al*. 2018 Decreased functional connectivity in the fronto-parietal network in children with mood disorders compared to children with dyslexia during rest: An fMRI study. *Neuroimage Clin.* **18**, 582–590.
Hunt S. E., McLaren W., Gil L., Thormann A., Schuilenburg H., Sheppard D. *et al*. 2018 Ensembl variation resources. *Database*, https://doi.org/10.1093/database/bay119.
Hyun K., Jeon J., Park K. and Kim J. 2017 Writing, erasing and reading histone lysine methylations. *Exp. Mol. Med.* **9**, e324.
Jambhekar A., Dhall A. and Shi Y. 2019 Roles and regulation of histone methylation in animal development. *Nat. Rev. Mol. Cell Biol.* **20**, 625–641.
Johnson M. R., Shkura K., Langley S. R., Delahaye-Duriez A., Srivastava P., Hill W. D. *et al*. 2016 Systems genetics identifies a convergent gene network for cognition and neurodevelopmental disease. *Nat. Neurosci.* **19**, 223–232.
Kang H. J., Kawasawa Y. I., Cheng F., Zhu Y., Xu X., Li M. *et al*. 2011 Spatiotemporal transcriptome of the human brain. *Nature* **478**, 483–489.
Landrum M. J., Lee J. M., Benson M., Brown G. R., Chao C., Chitipiralla S. *et al*. 2018 ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.* **46**, D1062–D1067.
Langfelder P. and Horvath S. 2008 WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559.
Lee P. H., O'Dushlaine C., Thomas B. and Purcell S. M. 2012 INRICH: interval-based enrichment analysis for genome-wide association studies. *Bioinformatics* **28**, 1797–1799.
MacArthur J., Bowler E., Cerezo M., Gil L., Hall P., Hastings E. *et al*. 2017 The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res.* **45**, D896–D901.
McCullumsmith R. E., O'Donovan S. M., Drummond J. B., Benesh F. S., Simmons M., Roberts R. *et al*. 2016 Cell-specific abnormalities of glutamate transporters in schizophrenia: sick astrocytes and compensating relay neurons? *Mol. Psychiat.* **21**, 823–830.
Miklowitz D. J., Efthimiou O., Furukawa T. A., Scott J., McLaren R., Geddes J. R. *et al.* 2020 Adjunctive psychotherapy for bipolar disorder: a systematic review and component network meta-analysis. *JAMA Psychiat*. (https://doi.org/https://doi.org/10.1001/jamapsychiatry.2020.2993).
Miller J. A., Ding S. L., Sunkin S. M., Smith K. A., Ng L., Szafer A. *et al*. 2014 Transcriptional landscape of the prenatal human brain. *Nature* **508**, 199–206.
Nestler E. J., Pena C. J., Kundakovic M., Mitchell A. and Akbarian S. 2016 Epigenetic basis of mental illness. *Neuroscientist* **22**, 447–463.
Network and Pathway Analysis Subgroup of Psychiatric Genomics Consortium 2015 Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. *Nat. Neurosci.* **18**, 199–209.
Oksenberg N. and Ahituv N. 2013 The role of AUTS2 in neurodevelopment and human evolution *Trends Genet.* **29**, 600-608.
Peter C. J. and Akbarian S. 2011 Balancing histone methylation activities in psychiatric disorders. *Trends Mol. Med.* **17**, 372–379.
Petrelli F., Pucci L. and Bezzi P. 2016 Astrocytes and microglia and their potential link with autism spectrum disorders. *Front. Cell. Neurosci.*, https://doi.org/10.3389/fncel.2016.00021.

Rhoades R., Jackson F. and Teng S. 2019 Discovery of rare variants implicated in schizophrenia using next-generation sequencing. *J. Transl. Genet. Genomics*, https://doi.org/10.20517/jtgg.2018.26.

Ricq E. L., Hooker J. M. and Haggarty S. J. 2016 Toward development of epigenetic drugs for central nervous system disorders: Modulating neuroplasticity via H3K4 methylation. *Psychiatry Clin. Neurosci.* **70**, 536–550.

Rubinstein W. S., Maglott D. R., Lee J. M., Kattman B. L., Malheiro A. J., Ovetsky M. *et al*. 2013 The NIH genetic testing registry: a new, centralized database of genetic tests to enable access to comprehensive information and improve transparency. *Nucleic Acids Res.* **4**, D925–D935.

Ruzicka W. B., Subburaju S. and Benes F. M. 2015 Circuit- and diagnosis-specific DNA methylation changes at gamma-aminobutyric acid-related genes in postmortem human hippocampus in schizophrenia and bipolar disorder. *JAMA Psychiat.* **72**, 541–551.

Shen E., Shulha H., Weng Z. and Akbarian S. 2014 Regulation of histone H3K4 methylation in brain development and disease. *Philos. Trans. R. Soc. London Ser. B.* **369,** (doi:https://doi.org/10.1098/rstb.2013.0514).

Smeland O. B., Frei O., Fan C. C., Shadrin A., Dale A. M. and Andreassen O. A. 2019 The emerging pattern of shared polygenic architecture of psychiatric disorders, conceptual and methodological challenges. *Psychiatr. Genet.* **29**, 152–159.

Stelzer G., Rosen N., Plaschkes I., Zimmerman S., Twik M., Fishilevich S. *et al*. 2016 The genecards suite: from gene data mining to disease genome sequence analyses. *Curr. Protoc. Bioinformatics* **54**, 30–33.

Substance Abuse and Mental Health Services Administration 2019 Key substance use and mental health indicators in the United States (https://www.samhsa.gov/data/sites/default/files/reports/rpt29393/2019NSDUHFFRPDFWHTML/2019NSDUHFFR090120.htm#ami).

Swerdlow N. R. and Light G. A. 2018 Sensorimotor gating deficits in schizophrenia: advancing our understanding of the phenotype, its neural circuitry and genetic substrates. *Schizophr. Res.* **198**, 1–5.

Tarasov V. V., Svistunov A. A., Chubarev V. N., Sologova S. S., Mukhortova P., Levushkin D. *et al*. 2019 Alterations of astrocytes in the context of schizophrenic dementia. *Front. Pharmacol.* **10**, 1612.

Wessa M. and Lois G. 2015 Brain functional effects of psychopharmacological treatment in major depression: a focus on neural circuitry of affective processing. *Curr. Neuropharmacol.* **13**, 466–479.

Wu Y., Yao Y. G. and Luo X. J. 2017 SZDB: A database for schizophrenia genetic research. *Schizophr. Bull.* **43**, 459–471.

Zhang Y. W., Wang Z., Xie W., Cai Y., Xia L., Easwaran H. *et al*. 2017 Acetylation enhances TET2 function in protecting against abnormal DNA methylation during oxidative stress. *Mol. Cell* **65**, 323–335.

Zhu Y., Womer F. Y., Leng H., Chang M., Yin Z., Wei Y., Zhou Q. *et al*. 2019 The relationship between cognitive dysfunction and symptom dimensions across schizophrenia, bipolar disorder, and major depressive disorder. *Front. Psychiat.* **10**, 253.

Corresponding editor: S. Ganesh