Associations between DNA methylation and schizophrenia-related intermediate phenotypes – a gene set enrichment analysis

Johanna Hass ^a, Esther Walton ^a, Carrie Wright ^{b,g}, Andreas Beyer ^{c,d}, Markus Scholz ^{e,f},
Jessica Turner ^{g,h}, Jingyu Liu ^{g,m}, Michael N. Smolka ⁱ, Veit Roessner ^a, Scott R. Sponheim ^j,
Randy L. Gollub ^{k,l}, Vince D. Calhoun ^{g,m}, Stefan Ehrlich ^{a,k,l}

^a Department of Child and Adolescent Psychiatry, Faculty of Medicine of the TU Dresden, Dresden, Germany

^b Department of Neurosciences, Health Sciences Center, University of New Mexico, Albuquerque, NM, USA

^c Cellular Networks and Systems Biology, Biotechnology Center, TU Dresden, Dresden, Germany

^d University of Cologne, CECAD, Cologne, Germany

^e Institute for Medical Informatics, Statistics and Epidemiology (IMISE), University of Leipzig, Leipzig, Germany

^f LIFE (Leipzig Interdisciplinary Research Cluster of Genetic Factors, Phenotypes and Environment), University of Leipzig, Leipzig, Germany

^g The Mind Research Network, Albuquerque, NM USA

^h Psychology Department, University of New Mexico, Albuquerque, NM, USA

ⁱ Department of Psychiatry, Faculty of Medicine of the TU Dresden, Dresden, Germany

^j Department of Psychiatry and the Center for magnetic Resonance Research, University of Minnesota, Minneapolis, MN USA

^k Department of Psychiatry, Massachusetts General Hospital, Boston, MA USA

^l MGH/MIT/HMS Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA USA

^m Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM USA

Corresponding author: Stefan Ehrlich, M.D., Faculty of Medicine of the TU Dresden,
Department of Child and Adolescent Psychiatry, Translational Developmental Neuroscience
Section, Fetscherstraße 74, 01307 Dresden, Germany, Phone: +49 (0)351 458-2244, Fax:
+49 (0)351 458 -5754, Email: stefan.ehrlich@uniklinikum-dresden.de

Abstract

Multiple genetic approaches have identified microRNAs as key effectors in psychiatric disorders as they post-transcriptionally regulate expression of thousands of target genes. However, their role in specific psychiatric diseases remains poorly understood . In addition, epigenetic mechanisms such as DNA methylation, which affect the expression of both microRNAs and coding genes, are critical for our understanding of molecular mechanisms in schizophrenia.

Using clinical, imaging, genetic, and epigenetic data of 103 patients with schizophrenia and 111 healthy controls of the Mind Clinical Imaging Consortium (MCIC) study of schizophrenia, we conducted gene set enrichment analysis to identify markers for schizophrenia-associated intermediate phenotypes. Genes were ranked based on the correlation between DNA methylation patterns and each phenotype, and then searched for enrichment in 221 predicted microRNA target gene sets.

We found the predicted hsa-miR-219a-5p target gene set to be significantly enriched for genes (*EPHA4*, *PKNOX1*, *ESR1*, amongst others) whose methylation status is correlated with hippocampal volume independent of disease status. Our results were strengthened by significant associations between hsa-miR-219a-5p target gene methylation patterns and hippocampus-related neuropsychological variables. IPA pathway analysis of the respective predicted hsa-miR-219a-5p target genes revealed associated network functions in behaviour and developmental disorders.

Altered methylation patterns of predicted hsa-miR-219a-5p target genes are associated with a structural aberration of the brain that has been proposed as a possible biomarker for schizophrenia. The (dys)regulation of microRNA target genes by epigenetic mechanisms may confer additional risk for developing psychiatric symptoms. Further study is needed to understand possible interactions between microRNAs and epigenetic changes and their impact on risk for brain-based disorders such as schizophrenia.

Keywords: DNA methylation; schizophrenia; microRNA targets; GSEA; intermediate
phenotype

1 Introduction

Schizophrenia is a highly heritable mental disorder [1], and after the decoding of the human genome, efforts to identify schizophrenia-associated genes have intensified. However, schizophrenia is genetically complex and likely involves thousands of genetic risk variants with small contributions to the disorder [2]. Recent meta-analyses of large genetic studies have identified novel genetic variants associated with schizophrenia or schizophrenia-related phenotypes [3, 4] but investigations have largely failed in finding replicable schizophrenia-specific susceptibility genes due to the lack of an etiologically-related phenotypic definition and the aforementioned polygeneity [5]. Furthermore, modulating environmental factors, which can affect gene expression through epigenetic mechanisms, have not been well studied. At present, DNA methylation is the most frequently studied epigenetic mechanism [6]. The covalent binding of a methyl group to the 5' end of cytosine in DNA CpG dinucleotides can silence the gene and inhibit gene expression [7]. It has been proposed that measuring DNA methylation changes may provide a better measure of environmental impacts than evaluating environmental factors themselves [8, 9]. Due to their dependence on genetic and environmental factors, epigenetic markers integrate information from both causal contributors to phenotypes. Therefore, understanding epigenetic signatures may help to improve our understanding of the pathogenesis of schizophrenia [10, 11].

Since intermediate phenotypes are thought to be more proximal to the underlying biology of schizophrenia than varying clinical constructs [12, 13], in this study we analyzed DNA methylation in conjunction with three different widely acknowledged schizophrenia-related intermediate phenotypes: reduced hippocampal volume, reduced superior temporal gyrus (STG) thickness, and dorsolateral prefrontal cortex (DLPFC) working memory (WM) load-dependent neural activity (% BOLD signal change). A reduction of hippocampal volume in schizophrenia patients has repeatedly been demonstrated [14–17] and abnormalities of hippocampal structure and function in schizophrenia have been associated with deficits in

1 memory and executive function [18]. The STG is involved in auditory processing [19] and
2 social cognitive processes [20]. Previous studies have shown structural alterations of the
3 STG in schizophrenia patients [21–24], which have also been associated with auditory
4 hallucinations [25–27]. Another well-studied intermediate phenotype is DLPFC activity during
5 WM processing [28–31]. Schizophrenia patients are thought to require additional neural
6 resources to achieve a comparable WM performance as healthy individuals. Depending on
7 the WM load this can result in increased DLPFC response, which has been termed “neural
8 inefficiency” [32–36].

9
10 An important class of molecules that we expect to affect risk for schizophrenia
11 possibly via modulation of intermediate phenotypes for schizophrenia are microRNAs
12 (miRNAs) and their target genes [37]. MiRNAs are small non-coding RNAs that can post-
13 transcriptionally silence the expression of a large number of target genes **either through**
14 **translational inhibition or degradation of target mRNAs** [38]. A set of miRNA target
15 genes regulated by a single miRNA likely constitutes a biological network of functionally-
16 associated molecules. Consequently, dysregulation of a single miRNA could be powerfully
17 influential for a polygenic disorder such as schizophrenia. In line with that, the largest
18 schizophrenia genome-wide association studies (GWAS) conducted to date identified
19 variants within the hsa-miR-137 host gene, *MIR137HG*, as well as variants within validated
20 miR-137 target genes to be among the top risk variants for schizophrenia [39, 40],
21 suggesting a potential role for the miRNA in the disorder. Several studies suggest that one of
22 the identified *MIR137HG* variants is associated with schizophrenia-related intermediate
23 phenotypes such as dorsolateral prefrontal cortex hyperactivation [41] and dorsolateral
24 prefrontal-hippocampal functional connectivity [42]. Furthermore, Wright et al. [43] found
25 significant enrichment for schizophrenia-associated genes among the list of potential and
26 experimentally validated miR-137 targets, as well as significant enrichment of targets within
27 schizophrenia-relevant canonical pathways, such as those involved in neuronal function and
28 development. In an imaging genetics approach, Potkin et al. [44] discovered gene regulatory
29 networks of GWAS-identified risk variants for schizophrenia that are assumed to be regulated

by several miRNAs, including miR-137 and others (miR-448, miR-218, miR-182, miR-518C, miR-200B, miR-429, miR-374, miR-369-3P, miR-27A, and miR-27B). Apart from miR-137, other miRNAs (such as miR-15, miR-219, miR-508) also have extensive evidence of their potential involvement in the pathophysiology of mental disorders [45, 46].

MiRNA-mediated regulation of target genes is highly correlated with miRNA target-gene specific promoter methylation [47]. Simultaneous changes of DNA methylation combined with miRNA dysregulation could thus potentiate effects on “downstream” genes (i.e. genes in the regulatory pathway of a miRNA) and various phenotypes. So far, there is little knowledge about the cooperative regulation of gene expression through miRNA targeting and DNA methylation. Analyzing miRNA target gene networks (instead of single gene analyses) and their epigenetic alterations may further deepen our understanding of the biological pathways underlying a complex illness such as schizophrenia.

In the present study we conducted gene set enrichment analyses (GSEA) using the predicted¹ miRNA target gene sets provided by the Molecular Signatures Database v4.0 (<http://www.broadinstitute.org/gsea/msigdb/>) of the GSEA toolbox [48]. In contrast to recent studies that mostly investigated enrichment in gene expression data sets [49–51], we explored DNA methylation in schizophrenia patients and healthy controls to identify potential associations between network level epigenetic changes in predicted miRNA target gene sets and widely studied intermediate phenotypes for schizophrenia. Gene set enrichment analysis holds the advantage that pathways can be reliably detected even when effect sizes of individual genes are small or signal-to-noise ratio is low, which is of importance especially for polygenic disorders such as schizophrenia. To the best of our knowledge, this approach – combining DNA methylation and intermediate phenotypes in a gene set enrichment analysis – has not been applied previously in the field of schizophrenia. Since we were interested in phenotypes associated with a brain disorder, we only included CpG sites (and corresponding

¹ The identification of miRNA target genes is largely based on bioinformatic algorithms. **However, a sizable proportion of genes are also experimentally validated targets of the respective microRNA (see last column of Table 3 for an example).**

genes) for which at least moderate correlation in DNA methylation between blood and brain tissue can be assumed [52].

2 Material and Methods

2.1 Participants

Imaging, genetic, epigenetic and behavioral data from participants of the Mind Clinical Imaging Consortium (MCIC) study of schizophrenia from four participating sites (the University of New Mexico (UNM), the University of Minnesota (UM), Massachusetts General Hospital (MGH), and the University of Iowa (UI)) were used to determine DNA methylation and genetic polymorphisms in cryo-conserved blood samples and to analyze structural and functional intermediate phenotypes. Out of a total of 328 participants, blood samples were available for 234 participants. DNA methylation and genetic data of 214 participants passed epi-/genetic quality control procedures, resulting in a final dataset of 103 schizophrenia patients and 111 healthy controls after imaging quality control steps (see below). Patients had a Diagnostic and Statistical Manual of Mental Disorders (*DSM-IV*) diagnosis of schizophrenia, established using a Structured Clinical Interview for DSM disorders (SCID) [53] and a review of case files by trained clinicians. In the initial cohort, controls were matched to the patient group for age, sex, and parental education and were excluded if they had a history of a medical or Axis I psychiatric diagnosis. For additional details about the participants and clinical measures, see Ehrlich et al. [14], Gollub et al. [54] and Supplementary Material (SM) 1.1.

2.2 Behavioral task

The Sternberg Item Recognition Paradigm (SIRP) is a working memory task, previously shown to consistently activate the DLPFC in healthy controls and schizophrenia patients [35]. The SIRP was administered during six 46 second blocks per run for three 360

second runs. In each block, a memory set composed of one (load 1), three (load 3), or five (load 5) digits was presented (two blocks per load condition). The Encode phase was followed by a presentation of 14 digits, one at a time (the Probe phase) and participants responded to each probe to indicate whether or not the probe digit was in the memory set. For additional details about the paradigm, see Roffman et al. [55] and SM 1.2. The stimuli and responses were presented and collected using E-prime software (EPrime v1.1, Psychology Software Tools, Inc., Pittsburgh, PA). Participants were excluded from further analysis, if they completed a block with less than a 75% accuracy rate and/or with more than 6 probes not answered within a block.

2.3 Image Acquisition and Processing

Structural magnetic resonance imaging (MRI) data was acquired with either a 1.5T Siemens Sonata (UNM, MGH, and UI) or a 3T Siemens Trio (UMN). Cortical reconstruction and volumetric segmentation based on high resolution structural MRI scans was performed with the FreeSurfer surface reconstruction software (<http://surfer.nmr.mgh.harvard.edu>). Hippocampal volume and STG thickness are part of the FreeSurfer standard output, mean volume and thickness values for each participant were calculated using measures from the left and right hemisphere, respectively.

Functional MRI data was acquired with either a 1.5T Siemens Sonata (UNM) or a 3T Siemens Trio (UMN, MGH, and UI) and was analyzed using fMRIB Software Library (FSL) (<http://www.fmrib.ox.ac.uk/fsl>). A Functional Imaging Linear Model [56] was fit to model the Probe phases of each subject's preprocessed functional time series. We used the linear Contrasts Of Parameter Estimate (COPE) Probe-load 5 versus Probe-load 1 and refer to responses to this condition as 'load-dependent' activation [55]. For a region-of-interest (ROI) analysis, we derived DLPFC ROIs from FreeSurfer cortical parcellations based on each participants individual cortical anatomy as described previously [14, 57]. We obtained indices of activation for the left and right DLPFC ROI using the COPE obtained from the second-level fixed-effects analysis for each participant and calculated mean DLPFC activation in

percent signal change. For additional details about data acquisition, quality control and preprocessing, see Walton et al. [58], SM 1.3 and 1.4.

2.4 DNA methylation data: Sample processing and quality control

Blood samples were obtained from 234 participants and sent to the Harvard Partners Center for Genetics and Genomics for DNA extraction. All DNA extraction, bisulphite modification and hybridization steps were done blinded to group assignment, the latter two were performed at the Mind Research Network Neurogenetics Core Lab on an Infinium HumanMethylation27 BeadChip using Illumina Infinium Methylation Assay. Built-in controls were used to evaluate the quality of individual arrays.

Samples and CpG sites were filtered according to coverage (see SM 1.5). After excluding one intensity outlier identified by a quantile color bias adjustment of the red and green fluorescence channel intensities, data were normalized using the R package “wateRmelon” of the Bioconductor project [59]. Finally, we removed five samples displaying sex mismatch resulting in a dataset of 229 participants. After excluding participants, who had no imaging data or imaging/behavioral/genetic data of insufficient quality (see above), genome-wide methylation beta-values were extracted from a final dataset of 103 schizophrenia patients and 111 healthy controls. For more information on DNA methylation preprocessing and intensity data extraction, see SM 1.5.

We corrected our data statistically for the effects of chip and bisulphite sequencing (BS) conversion efficiency by first estimating coefficients in a linear regression model and then computing residuals (corrected beta-values). In a similar way, we removed the effects of sex and age [60–62] as well as population stratification as follows: DNA methylation has been shown to be moderately influenced by population structure, although to a much lesser degree than genetic data [63]. By means of principal and independent component analyses, Liu et al. [63] found that the population structure present in DNA methylation data is very well captured by the first genetic principal component. Hence, to avoid confounding effects due to population stratification, we applied principal component analysis (PCA) to our genotype data

(Illumina HumanOmni-Quad BeadChip) using EIGENSTRAT of the EIGENSOFT 3.0 software package [64, 65], extracted ten principal components (PC) and then included the first component in our regression models. For more details, see [15].

Since DNA methylation is known to be tissue-specific [66–68], the question arises whether we can draw conclusions on brain-based intermediate phenotypes based on methylation values measured in peripheral blood cells. Walton et al. (unpublished results) established a list of CpG sites with well correlated DNA methylation beta-values by comparing brain and blood samples from patients with treatment-resistant epilepsy in a within-subject design. Accordingly, we excluded all CpG sites with a low correlation (Spearman's rank correlation coefficient < 0.5) between brain and blood tissue-specific methylation beta-values. The final methylation data set comprised 214 subjects and 3,086 CpG sites.

2.5 Statistics

Basic demographic characteristics were compared across diagnostic groups and acquisition site-specific scanner field strengths using Welch's t tests. Chi-square statistics were used to examine differences in categorical variables. Alpha was set to 0.05 for all analyses. Sample characteristic analyses were carried out with IBM SPSS Statistics 21.

Each of the three intermediate phenotypes (hippocampal volume, STG thickness, and WM load-dependent neural activity (% BOLD signal change) in the DLPFC) was controlled for the effects of age, sex and site-specific scanner field strength in a linear regression model. Standardized residuals of each phenotype were used in all (gene set enrichment) analyses. For hippocampal volume, we also controlled for intracranial volume[69].

We used GSEA [48] to identify predicted miRNA target gene sets whose methylation patterns correlated with the intermediate phenotypes (Figure 1): First, each gene's methylation pattern was correlated with the respective phenotype across all patients and controls. If there was more than one CpG site per gene, i.e. more than one beta-value associated with a gene name, median methylation was used (GSEA standard procedure).

Correlation was determined using the Pearson product-moment correlation coefficient (GSEA standard procedure for continuous phenotypes). Next, genes were ranked based on those correlations and tested if the ones targeted by a given miRNA had methylation-phenotype correlations larger than expected by chance. For this, predicted target gene sets of 221 miRNAs were obtained from the Molecular Signatures Database (<http://www.broadinstitute.org/gsea/msigdb/>) by selecting genes containing the respective 3'-UTR miRNA binding motifs, and then filtered for genes with available methylation data. We adopted the current annotation and terminology of miRNAs according to miRBase v21.

There are three key elements of the GSEA method: (1) Calculation of an Enrichment Score (ES), that reflects the degree to which a predicted miRNA target gene set is overrepresented at the top or bottom of the list of genes ranked according to the positive or negative correlation between methylation and phenotype; (2) Estimation of Significance Level of ES using an empirical phenotype-based permutation test procedure; (3) Adjustment for Multiple Hypothesis Testing, including normalization of the ES to account for different gene set sizes and calculation of the false discovery rate (FDR) to control the proportion of false positives. For further details about the GSEA method, please see Subramanian et al. [48]. Since we applied the gene set enrichment analysis as an exploratory approach, we used an initial FDR cutoff of 0.25 to determine significance of our results (GSEA default setting for permutation), which is consistent with many previous analyses using GSEA [70–73]. Using a more stringent FDR cutoff may substantially increase the chances to overlook potentially significant results [74]. To control for the number of tested intermediate phenotypes ($n = 3$), we lowered the FDR threshold to 0.083 ($= 0.25/3$).

3 Results

3.1 Sample characteristics

Patients and controls did not significantly differ in demographic variables such as age, sex, parental socio-economic status and Annett handedness score. Schizophrenia patients

had a significantly lower WRAT3-RT score and showed significantly lower mean hippocampal volume and mean STG thickness. We found no significant difference in DLPFC activity during SIRP, but schizophrenia patients performed significantly lower in the WM paradigm (Table 1). For site-specific demographics, clinical variables and intermediate phenotypes see SM Table 1.

3.2 Gene set enrichment analysis of DNA methylation

We correlated each gene's methylation pattern with the phenotype variation across the study population and subsequently identified miRNAs whose predicted target genes are more correlated than expected by chance (see Materials and Methods). This approach assumes that target genes of a miRNA are functionally linked and thus more likely to be methylated in concert than random sets of genes.

For hippocampal volume, we found three predicted miRNA target gene sets significantly enriched (FDR < 25%) based on a positive correlation between methylation data and the phenotype (see Table 2). Hsa-miR-219a-5p² was the target gene set with the highest normalized ES (NES = 1.9368). The associated FDR q-value was 0.0498. Given this q-value, the hsa-miR-219a-5p gene set was the only target gene set passing multiple testing correction. See Figure 2 for a graphical view of the ES (with detailed information on how to interpret enrichment plots in SM 2.1) and refer to Table 3 for a full list of genes included in the predicted hsa-miR-219a-5p target gene set.

For STG thickness, the predicted miR-513 target gene set was significantly enriched (NES = 1.7636) with a FDR q-value of 0.2404, again based on a positive correlation between methylation data and phenotype. None of the predicted miRNA target gene sets was significantly enriched when analyzing the methylation data in association with working-memory load-dependent neural activity in the DLPFC.

² Hsa = human (homo sapiens); see SM Table 2 for details on the miR-219 gene family and corresponding mature miRNAs

3.3 Additional analyses

We carried out separate gene set enrichment analyses for schizophrenia patients and healthy controls to investigate potential subgroup effects. Again, the predicted hsa-miR-219a-5p target gene set was enriched for hippocampal volume in schizophrenia patients as well as in healthy controls (nominal p-value of 0.016 and 0.038, respectively; based on a positively ranked list metric).

In a subsequent model, we also adjusted for lifetime exposure to antipsychotic drugs in schizophrenia patients (by including this variable as an additional covariate in the phenotype regression analysis), and again found the predicted hsa-miR-219a-5p target gene set enriched with a nominal p-value of 0.016.

Since our main finding is related to hippocampal volume, we aimed to further explore this association by investigating hippocampus-related neuropsychological variables as additional phenotypes. In particular, we tested for a methylation-diagnosis interaction effect (using ANOVA) on indices of episodic memory (Logical Memory II subtest of the WMS-III [75]) and manipulation of material in working memory (Letter Number Sequencing [LNS] subtest of the WAIS-III [76]), both of which have been found to be important to memory formation and medial temporal lobe function [77]. We found a significant interaction effect of *EPHA4* methylation (cg15808558) and diagnosis on LNS ($F(1) = 7.120$, $p = 0.008$). Similarly, we found an interaction effect of trend level significance for *PKNOX1* methylation (cg14204791) and diagnosis on LNS ($F(1) = 3.799$, $p = 0.053$).

Ingenuity pathway analysis (IPA, Ingenuity® Systems, CA, USA, www.ingenuity.com; see SM 1.6 for method details) of predicted hsa-miR-219a-5p target genes that were most highly correlated with hippocampal volume (Table 3), revealed a functional role of these genes in cellular development ($p = 2.19 \times 10^{-4}$), gene expression ($p = 2.38 \times 10^{-4}$) and cell-to-cell signaling and interaction ($p = 4.70 \times 10^{-4}$), as well as associated network functions in behavior and developmental disorders (Ingenuity network score = 21; according to the fit of the network to the hsa-miR-219a-5p leading edge genes) [78]. Including all predicted hsa-miR-219a-5p target genes of Table 3, there is supporting evidence for an association with

axonal guidance signaling pathways ($p = 0.0128$), and again associated network functions in behavior (Ingenuity network score = 24) as well as cell-to-cell signaling and interaction (Ingenuity network score = 16).

4 Discussion

Results of this study, which employed a GSEA approach to methylation data in schizophrenia for the first time, suggests that the methylation state of hsa-miR-219a-5p targets correlates with hippocampal volume – an intermediate phenotype for schizophrenia. By focusing on CpG sites with highly correlated blood and brain tissue-specific methylation beta-values, we were able to ensure that the identified potential targets impact brain-related biological processes, which was confirmed by the IPA pathway analysis.

As the first identified functional mature product of the miR-219 precursor family, hsa-miR-219a-5p was originally referred to as hsa-miR-219. Hsa-miR-219a-5p is coded by two different stem loop sequences, hsa-miR-219a-1 and hsa-miR-219a-2, in two locations, 6p21 (which is a putative susceptibility locus for schizophrenia [79]) and 9q34³, respectively (see SM Table 2 for details). This may add to the complexity of regulation of hsa-miR-219a-5p expression and subsequently its target genes. Identified as miR-219 in previous studies, hsa-miR-219a-5p has also been implicated in circadian clock regulation [81], oligodendrocyte differentiation and myelination [82] as well as disruption of NMDA receptor signaling [83] – all processes that have been associated with schizophrenia: A disruption of circadian rhythm, in particular sleep-onset and maintenance insomnia, is a common phenomenon in schizophrenia [84]. In line with Zhao et al. [82], a large number of diffusion tensor imaging studies provide evidence for reduced white matter integrity in schizophrenia [85–90] and there is ample evidence for NMDA receptor hypofunction in schizophrenia [91, 92].

Furthermore, hsa-miR-219-5p was found to be the most highly enriched miRNA in

³ Human genome Browser at UCSC (GRCh37/hg19 assembly) [80]

1 synaptic fractions (5-fold), while at the same time it is the most strongly down-
2 regulated miRNA in schizophrenia synaptosomes (70% decrease) [46]. Taken together,
3 these findings and our own data suggest a major role of the miR-219 precursor family
4 miRNAs, and especially hsa-miR-219a-5p, in brain development, abnormal behavior
5 and schizophrenia in particular.
6
7
8
9

10 One of hsa-miR-219a-5p's predicted target genes that was identified as a leading
11 edge gene (i.e. genes that contribute most to the ES) in our gene set enrichment analysis, is
12 the ephrin type-A receptor 4 (*EPHA4*) gene, which belongs to the ephrin receptor subfamily
13 of the protein-tyrosine kinase family and is expressed predominantly in the hippocampus [93,
14 94]. EPH and EPH-related receptors have been implicated in mediating developmental
15 events, particularly in the nervous system, including axon guidance [95, 96]. EPH signaling
16 was found to be crucial for hippocampal circuit plasticity [97] and to be associated with
17 synaptic long-term potentiation [98]. In addition, we found a significant interaction effect
18 between *EPHA4* methylation (cg15808558) and diagnosis on memory performance in our
19 sample, which suggests a potential influence of an altered hsa-miR-219a-5p target
20 methylation also at the behavioral level. **A reduced performance on memory tests is one**
21 **of the most stable cognitive findings in schizophrenia [99, 100].**
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36

37 Another predicted target gene of hsa-miR-219a-5p is *PKNOX1* (PBX/knotted 1
38 homeobox 1) and its overexpression in the hippocampus was associated with facilitated
39 learning and memory in a mouse model for the trisomy of the Abcg1-U2af1 region [101].
40 Again, we found an interaction effect between *PKNOX1* methylation (cg14204791) and
41 diagnosis on memory performance implying that an altered *PKNOX1* expression in
42 schizophrenia patients or individuals at risk (i.e. related to an aberrant methylation pattern)
43 may further add to the cognitive deficits which are hallmarks of schizophrenia.
44
45
46
47
48
49
50
51
52

53 The estrogen receptor 1, encoded by the *ESR1* gene and another predicted hsa-miR-
54 219a-5p target, is a ligand-activated transcription factor [102] regulating cellular proliferation
55 and differentiation in target tissues [103, 104]. **Only recently, ESR1 has been**
56 **experimentally validated to be a target gene of hsa-mir-219-5p [105].** Although not
57
58
59
60
61
62
63
64
65

explicitly in hippocampus, Knickmeyer et al. [106] observed smaller volumes for an *ESR1* variant in multiple regions of frontal and parietal lobes. **Weickert et al. [107] found several variants of the *ESR1* gene and *ESR1* expression being associated with schizophrenia (although a negative genetic study does exist as well [108]). Additional studies conducted by the same group also found reduced *ESR1* mRNA levels in cortical and subcortical brain regions in patients suffering from major mental illness including a reduction in hippocampal *ESR1* mRNA in patients with schizophrenia [109, 110].** Estrogen, which binds to estrogen receptor 1 encoded by the *ESR1* gene, modulates human emotion [111] and frontal cortical activity during cognitive tasks [112, 113], both of which are altered in schizophrenia and other mental disorders [107]. Furthermore, sex differences have been described for nearly all features of schizophrenia [114]. Clinical studies suggest that estrogens exert psychoprotective effects in schizophrenia patients [115] which may be mediated by an interaction effect between estrogen and brain-derived neurotrophic factor (BDNF) [114].

To further identify possible functional mechanisms, we performed IPA pathway analysis including all predicted hsa-miR-219a-5p target genes of Table 3, and found several of those potential target genes to be involved in axonal guidance signaling, a process known to be involved in neuronal development **and a prerequisite for normal cognitive functioning** which has been repeatedly shown to be disrupted in schizophrenia [116, 117]. Furthermore, IPA pathway analysis revealed predicted hsa-miR-219a-5p target genes to be associated with cellular development, gene expression and cell-to-cell signaling and cell-cell interaction, all of which are of substantial interest in the etiology of psychiatric disorders. However, the exact mechanisms of how the expression of schizophrenia-associated risk genes may be concurrently regulated through miRNA activity and DNA methylation in terms of possible synergistic effects remain to be clarified and could be studied *in vitro* or using animal models.

The findings of our study have to be considered in the light of the following limitations: First, complete epigenetic and phenotypic data was available for only 214 participants after

quality assurance steps, lowering the statistical power of our analyses. Our findings warrant replication using larger samples. In addition, we analyzed a limited number of CpG sites following the approach of Walton et al. (unpublished results). Although this approach may lead to false negative results, it avoids inclusion of CpG sites with methylation levels that cannot be predicted from blood samples. Second, antipsychotic medication has been shown to affect DNA methylation [118]. We accounted for the effect of medication on DNA methylation in the sample of schizophrenia patients by using cumulative lifetime exposure to antipsychotic drugs as an additional covariate (see 3.3. Additional analyses). Results of this sub-analysis strengthen our main finding. Third, the motif gene sets of the Molecular Signature Database used in our GSEA procedure, in particular the miRNA target gene sets, contain bioinformatically predicted target genes (based on 3'-UTR miRNA binding motifs) that are not experimentally validated for the most part (**see also Table 3**). Therefore, some of the miRNA target genes which we have scrutinized here may in fact not be regulated by the predicted miRNAs [119]. Fourthly, gene set enrichment analysis as carried out by GSEA ignores gene-gene correlation, which may lead to an overestimation of significance levels and eventually to false positives. Finally, Gleeleher et al. [120] stated, that gene set analysis could be severely biased when applied to genome-wide methylation data because lengths and expression levels of genes frequently differ between gene sets. To overcome this potential bias, GSEA applies sample label permutations which was shown to be robust to the differences in the number of probes per gene [121, 122]. In summary, we suggest GSEA as a very useful method for hypothesis generation as it is robust to a low signal-to-noise ratio and our subsequent statistical analyses substantiated the GSEA main findings.

5 Conclusions

We found hsa-miR-219a-5p target methylation to be associated with hippocampal volume – a schizophrenia-related intermediate phenotype. The (dys)regulation of miRNAs and miRNA targets by epigenetic mechanisms has been found to interfere with physiological

1 brain functions [123–125] and might convey additional risk for structural brain changes and,
2 possibly, psychiatric symptoms. However, further study is needed to deepen our
3
4 understanding of the complex genetic and epigenetic architecture underlying
5
6 neuropsychiatric disorders. An integrated analysis of miRNA and epigenetic control could
7
8 highlight genes that are regulated by DNA methylation and are targeted by miRNAs with a
9
10 potential use as clinical markers, eventually leading to the identification of novel targets for
11
12 therapeutic approaches.
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Contributors

SE designed the study, wrote the protocol and supervised data analysis and the writing of the manuscript. JH conducted the statistical data analysis, managed the literature searches and wrote the first draft of the manuscript. EW and CW helped with the statistical analysis and assisted in manuscript preparation. AB and MS helped with the analysis and interpretation of the data and revised the manuscript. JT and JL assisted with the design of our analysis approach and revised the manuscript. MNS, VR and VDC helped with the interpretation of the data and assisted in manuscript preparation. SRS and RLG were responsible for the MCIC study design, supervision of subject assessments and data collection. All authors contributed to and have approved the final article.

Acknowledgement

The authors wish to express their gratitude to the many individuals who contributed to the MCIC study of schizophrenia.

Role of funding source

This work was supported by the National Institutes of Health (NIH/NCRR P41RR14075, NIBIB 2R01EB000840 and COBRE 5P20RR021938/P20GM103472), Department of Energy (DE-FG02-99ER62764), The Mind Research Network, Morphometry BIRN (1U24, RR021382A), Function BIRN (U24RR021992-01, NIH.NCRR MO1 RR025758-

01), NARSAD Young Investigator Grant (SE), and the Deutsche Forschungsgemeinschaft
(Research Fellowship to SE). All funding sources had no involvement in study design; in the
collection, analysis and interpretation of data; in the writing of the report; and in the decision
to submit the article for publication.

References

1. Sullivan PF, Kendler KS, Neale MC: **Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies.** *Arch Gen Psychiatry* 2003, **60**:1187–1192.
2. O'Donovan MC, Craddock NJ, Owen MJ: **Genetics of psychosis; insights from views across the genome.** *Hum Genet* 2009, **126**:3–12.
3. Ripke S, O'Dushlaine C, Chambert K, Moran JL, Kähler AK, Akterin S, Bergen SE, Collins AL, Crowley JJ, Fromer M, Kim Y, Lee SH, Magnusson PKE, Sanchez N, Stahl EA, Williams S, Wray NR, Xia K, Bettella F, Borglum AD, Bulik-Sullivan BK, Cormican P, Craddock N, de Leeuw C, Durmishi N, Gill M, Golimbet V, Hamshere ML, Holmans P, Hougaard DM, et al.: **Genome-wide association analysis identifies 13 new risk loci for schizophrenia.** *Nat Genet* 2013, **45**:1150–9.
4. Stein JL, Medland SE, Vasquez AA, Hibar DP, Senstad RE, Winkler AM, Toro R, Appel K, Bartecek R, Bergmann Ø, Bernard M, Brown AA, Cannon DM, Chakravarty MM, Christoforou A, Domin M, Grimm O, Hollinshead M, Holmes AJ, Homuth G, Hottenga J-J, Langan C, Lopez LM, Hansell NK, Hwang KS, Kim S, Laje G, Lee PH, Liu X, Loth E, et al.: **Identification of common variants associated with human hippocampal and intracranial volumes.** *Nat Genet* 2012, **44**:552–561.
5. Cross-Disorder Group of the Psychiatric Genomics Consortium, Lee SH, Ripke S, Neale BM, Faraone SV, Purcell SM, Perlis RH, Mowry BJ, Thapar A, Goddard ME, Witte JS, Absher D, Agartz I, Akil H, Amin F, Andreassen OA, Anjorin A, Anney R, Anttila V, Arking DE, Asherson P, Azevedo MH, Backlund L, Badner JA, Bailey AJ, Banaschewski T, Barchas JD, Barnes MR, Barrett TB, Bass N, et al.: **Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs.** *Nat Genet* 2013, **45**:984–994.
6. Roth TL, Lubin FD, Sodhi M, Kleinman JE: **Epigenetic mechanisms in schizophrenia.** *Biochim Biophys Acta* 2009, **1790**:869–877.
7. Curradi M, Izzo A, Badaracco G, Landsberger N: **Molecular Mechanisms of Gene Silencing Mediated by DNA Methylation.** *Mol Cell Biol* 2002, **22**:3157–3173.
8. Popov NT, Stoyanova VK, Madzhirova NP, Vachev TI: **Epigenetic aspects in schizophrenia etiology and pathogenesis.** *Folia Med (Plovdiv)* 2012, **54**:12–16.
9. Nishioka M, Bundo M, Kasai K, Iwamoto K: **DNA methylation in schizophrenia: progress and challenges of epigenetic studies.** *Genome Med* 2012, **4**:96.
10. Grayson DR, Guidotti A: **The dynamics of DNA methylation in schizophrenia and related psychiatric disorders.** *Neuropsychopharmacology* 2013, **38**:138–166.
11. Liu J, Chen J, Ehrlich S, Walton E, White T, Perrone-Bizzozero N, Bustillo J, Turner JA, Calhoun VD: **Methylation patterns in whole blood correlate with symptoms in schizophrenia patients.** *Schizophr Bull* 2014, **40**:769–776.
12. Gottesman II, Gould TD: **The endophenotype concept in psychiatry: etymology and strategic intentions.** *Am J Psychiatry* 2003, **160**:636–645.
13. Meyer-Lindenberg A, Weinberger DR: **Intermediate phenotypes and genetic mechanisms of psychiatric disorders.** *Nat Rev Neurosci* 2006, **7**:818–827.
14. Ehrlich S, Morrow EM, Roffman JL, Wallace SR, Naylor M, Bockholt HJ, Lundquist A,

Yendiki A, Ho B-C, White T, Manoach DS, Clark VP, Calhoun VD, Gollub RL, Holt DJ: **The COMT Val108/158Met polymorphism and medial temporal lobe volumetry in patients with schizophrenia and healthy adults.** *Neuroimage* 2010, **53**:992–1000.

15. Hass J, Walton E, Kirsten H, Liu J, Priebe L, Wolf C, Karbalai N, Gollub R, White T, Roessner V, Müller KU, Paus T, Smolka MN, Schumann G, IMAGEN Consortium, Scholz M, Cichon S, Calhoun V, Ehrlich S: **A Genome-Wide Association Study Suggests Novel Loci Associated with a Schizophrenia-Related Brain-Based Phenotype.** *PLoS ONE* 2013, **8**:e64872.

16. Van Erp TGM, Saleh PA, Huttunen M, Lönngqvist J, Kaprio J, Salonen O, Valanne L, Poutanen V-P, Standertskjöld-Nordenstam C-G, Cannon TD: **Hippocampal volumes in schizophrenic twins.** *Arch Gen Psychiatry* 2004, **61**:346–353.

17. Velakoulis D, Wood SJ, Wong MTH, McGorry PD, Yung A, Phillips L, Smith D, Brewer W, Proffitt T, Desmond P, Pantelis C: **Hippocampal and amygdala volumes according to psychosis stage and diagnosis: a magnetic resonance imaging study of chronic schizophrenia, first-episode psychosis, and ultra-high-risk individuals.** *Arch Gen Psychiatry* 2006, **63**:139–149.

18. Antonova E, Sharma T, Morris R, Kumari V: **The relationship between brain structure and neurocognition in schizophrenia: a selective review.** *Schizophr Res* 2004, **70**:117–45.

19. Chen Y-H, Edgar JC, Huang M, Hunter MA, Epstein E, Howell B, Lu BY, Bustillo J, Miller GA, Cañive JM: **Frontal and superior temporal auditory processing abnormalities in schizophrenia.** *Neuroimage Clin* 2013, **2**:695–702.

20. Bosia M, Riccaboni R, Poletti S: **Neurofunctional correlates of theory of mind deficits in schizophrenia.** *Curr Top Med Chem* 2012, **12**:2284–2302.

21. Edgar JC, Chen YH, Lanza M, Howell B, Chow VY, Heiken K, Liu S, Wootton C, Hunter MA, Huang M, Miller GA, Cañive JM: **Cortical thickness as a contributor to abnormal oscillations in schizophrenia?** *Neuroimage Clin* 2013, **4**:122–9.

22. Honea R, Crow TJ, Passingham D, Mackay CE: **Regional deficits in brain volume in schizophrenia: a meta-analysis of voxel-based morphometry studies.** *Am J Psychiatry* 2005, **162**:2233–45.

23. Kasai K, Shenton ME, Salisbury DF, Hirayasu Y, Lee CU, Ciszewski AA, Yurgelun-Todd D, Kikinis R, Jolesz FA, McCarley RW: **Progressive decrease of left superior temporal gyrus gray matter volume in patients with first-episode schizophrenia.** *Am J Psychiatry* 2003, **160**:156–64.

24. Yoshida T, McCarley RW, Nakamura M, Lee K, Koo M-S, Bouix S, Salisbury DF, Morra L, Shenton ME, Niznikiewicz MA: **A prospective longitudinal volumetric MRI study of superior temporal gyrus gray matter and amygdala-hippocampal complex in chronic schizophrenia.** *Schizophr Res* 2009, **113**:84–94.

25. Barta PE, Pearlson GD, Powers RE, Richards SS, Tune LE: **Auditory hallucinations and smaller superior temporal gyral volume in schizophrenia.** *Am J Psychiatry* 1990, **147**:1457–62.

26. Flaum M, O'Leary DS, Swayze VW 2nd, Miller DD, Arndt S, Andreasen NC: **Symptom dimensions and brain morphology in schizophrenia and related psychotic disorders.** *J Psychiatr Res* 1995, **29**:261–276.

27. Nenadic I, Smesny S, Schlösser RGM, Sauer H, Gaser C: **Auditory hallucinations and brain structure in schizophrenia: voxel-based morphometric study.** *Br J Psychiatry* 2010, **196**:412–413.
28. Kim DI, Sui J, Rachakonda S, White T, Manoach DS, Clark VP, Ho BC, Schulz SC, Calhoun VD: **Identification of imaging biomarkers in schizophrenia: a coefficient-constrained independent component analysis of the mind multi-site schizophrenia study.** *Neuroinformatics* 2009, **8**:213–29.
29. Manoach DS: **Prefrontal cortex dysfunction during working memory performance in schizophrenia: reconciling discrepant findings.** *Schizophr Res* 2003, **60**:285–98.
30. Potkin SG, Turner JA, Brown GG, McCarthy G, Greve DN, Glover GH, Manoach DS, Belger A, Diaz M, Wible CG, Ford JM, Mathalon DH, Gollub R, Lauriello J, O’Leary D, van Erp TGM, Toga AW, Preda A, Lim KO, FBIRN: **Working memory and DLPFC inefficiency in schizophrenia: the FBIRN study.** *Schizophr Bull* 2009, **35**:19–31.
31. Walton E, Geisler D, Lee PH, Hass J, Turner JA, Liu J, Sponheim SR, White T, Wassink TH, Roessner V, Gollub RL, Calhoun VD, Ehrlich S: **Prefrontal Inefficiency Is Associated With Polygenic Risk for Schizophrenia.** *Schizophr Bull* 2013:sbt174.
32. Brauns S, Gollub RL, Roffman JL, Yendiki A, Ho B-C, Wassink TH, Heinz A, Ehrlich S: **DISC1 is associated with cortical thickness and neural efficiency.** *Neuroimage* 2011, **57**:1591–1600.
33. Ehrlich S, Brauns S, Yendiki A, Ho B-C, Calhoun V, Schulz SC, Gollub RL, Sponheim SR: **Associations of cortical thickness and cognition in patients with schizophrenia and healthy controls.** *Schizophr Bull* 2011, **38**:1050–62.
34. Karlsgodt KH, Sanz J, van Erp TG, Bearden CE, Nuechterlein KH, Cannon TD: **Re-evaluating dorsolateral prefrontal cortex activation during working memory in schizophrenia.** *Schizophr Res* 2009, **108**:143–50.
35. Manoach DS, Press DZ, Thangaraj V, Searl MM, Goff DC, Halpern E, Saper CB, Warach S: **Schizophrenic subjects activate dorsolateral prefrontal cortex during a working memory task, as measured by fMRI.** *Biol Psychiatry* 1999, **45**:1128–1137.
36. Potkin SG, Turner JA, Guffanti G, Lakatos A, Fallon JH, Nguyen DD, Mathalon D, Ford J, Lauriello J, Macciardi F: **A Genome-Wide Association Study of Schizophrenia Using Brain Activation as a Quantitative Phenotype.** *Schizophr Bull* 2009, **35**:96–108.
37. Hansen T, Olsen L, Lindow M, Jakobsen KD, Ullum H, Jonsson E, Andreassen OA, Djurovic S, Melle I, Agartz I, Hall H, Timm S, Wang AG, Werge T: **Brain expressed microRNAs implicated in schizophrenia etiology.** *PLoS ONE* 2007, **2**:e873.
38. Selbach M, Schwanhäusser B, Thierfelder N, Fang Z, Khanin R, Rajewsky N: **Widespread changes in protein synthesis induced by microRNAs.** *Nature* 2008, **455**:58–63.
39. Schizophrenia Working Group of the Psychiatric Genomics Consortium: **Biological insights from 108 schizophrenia-associated genetic loci.** *Nature* 2014, **511**:421–427.
40. Kwon E, Wang W, Tsai L-H: **Validation of schizophrenia-associated genes CSMD1, C10orf26, CACNA1C and TCF4 as miR-137 targets.** *Mol Psychiatry* 2013, **18**:11–12.
41. Van Erp TG, Guella I, Vawter MP, Turner J, Brown GG, McCarthy G, Greve DN, Glover

- GH, Calhoun VD, Lim KO, Bustillo JR, Belger A, Ford JM, Mathalon DH, Diaz M, Preda A, Nguyen D, Macciardi F, Potkin SG: **Schizophrenia miR-137 locus risk genotype is associated with dorsolateral prefrontal cortex hyperactivation.** *Biol Psychiatry* 2014, **75**:398–405.
42. Liu B, Zhang X, Hou B, Li J, Qiu C, Qin W, Yu C, Jiang T: **The impact of MIR137 on dorsolateral prefrontal-hippocampal functional connectivity in healthy subjects.** *Neuropsychopharmacology* 2014, **39**:2153–60.
43. Wright C, Turner JA, Calhoun VD, Perrone-Bizzozero N: **Potential Impact of miR-137 and Its Targets in Schizophrenia.** *Front Genet* 2013, **4**:58.
44. Potkin SG, Macciardi F, Guffanti G, Fallon JH, Wang Q, Turner JA, Lakatos A, Miles MF, Lander A, Vawter MP, Xie X: **Identifying gene regulatory networks in schizophrenia.** *Neuroimage* 2010, **53**:839–847.
45. Beveridge NJ, Gardiner E, Carroll AP, Tooney PA, Cairns MJ: **Schizophrenia is associated with an increase in cortical microRNA biogenesis.** *Mol Psychiatry* 2010, **15**:1176–1189.
46. Smalheiser NR, Lugli G, Zhang H, Rizavi H, Cook EH, Dwivedi Y: **Expression of microRNAs and Other Small RNAs in Prefrontal Cortex in Schizophrenia, Bipolar Disorder and Depressed Subjects.** *PLoS One* 2014, **9**.
47. Taguchi Y-H: **MicroRNA-mediated regulation of target genes in several brain regions is correlated to both microRNA-targeting-specific promoter methylation and differential microRNA expression.** *BioData Min* 2013, **6**:11.
48. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP: **Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.** *Proc Natl Acad Sci USA* 2005, **102**:15545–15550.
49. Harris LW, Lockstone HE, Khaitovich P, Weickert CS, Webster MJ, Bahn S: **Gene expression in the prefrontal cortex during adolescence: implications for the onset of schizophrenia.** *BMC Medical Genomics* 2009, **2**:28.
50. Lai Y, Zhang F, Nayak TK, Modarres R, Lee NH, McCaffrey TA: **Concordant integrative gene set enrichment analysis of multiple large-scale two-sample expression data sets.** *BMC Genomics* 2014, **15 Suppl 1**:S6.
51. Schmitt A, Leonardi-Essmann F, Durrenberger PF, Parlapani E, Schneider-Axmann T, Spanagel R, Arzberger T, Kretschmar H, Herrera-Marschitz M, Gruber O, Reynolds R, Falkai P, Gebicke-Haerter PJ: **Regulation of immune-modulatory genes in left superior temporal cortex of schizophrenia patients: a genome-wide microarray study.** *World J Biol Psychiatry* 2011, **12**:201–215.
52. Nishioka M, Bundo M, Kasai K, Iwamoto K: **DNA methylation in schizophrenia: progress and challenges of epigenetic studies.** *Genome Med* 2012, **4**:96.
53. First M, Spitzer A, Gibbon M, Williams J: *Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Nonpatient Edition.* New York: New York State Psychiatric Institute; 2002.
54. Gollub RL, Shoemaker JM, King MD, White T, Ehrlich S, Sponheim SR, Clark VP, Turner JA, Mueller BA, Magnotta V, O'Leary D, Ho BC, Brauns S, Manoach DS, Seidman L, Bustillo

JR, Lauriello J, Bockholt J, Lim KO, Rosen BR, Schulz SC, Calhoun VD, Andreasen NC: **The MCIC collection: a shared repository of multi-modal, multi-site brain image data from a clinical investigation of schizophrenia.** *Neuroinformatics* 2013, **11**:367–388.

55. Roffman JL, Gollub RL, Calhoun VD, Wassink TH, Weiss AP, Ho BC, White T, Clark VP, Fries J, Andreasen NC, Goff DC, Manoach DS: **MTHFR 677C → T genotype disrupts prefrontal function in schizophrenia through an interaction with COMT 158Val → Met.** *Proc Natl Acad Sci USA* 2008, **105**:17573–17578.

56. Woolrich MW, Ripley BD, Brady M, Smith SM: **Temporal autocorrelation in univariate linear modeling of FMRI data.** *Neuroimage* 2001, **14**:1370–1386.

57. Yendiki A, Greve DN, Wallace S, Vangel M, Bockholt J, Mueller BA, Magnotta V, Andreasen N, Manoach DS, Gollub RL: **Multi-site characterization of an fMRI working memory paradigm: reliability of activation indices.** *Neuroimage* 2010, **53**:119–131.

58. Walton E, Turner J, Gollub RL, Manoach DS, Yendiki A, Ho B-C, Sponheim SR, Calhoun VD, Ehrlich S: **Cumulative genetic risk and prefrontal activity in patients with schizophrenia.** *Schizophr Bull* 2013, **39**:703–711.

59. Pidsley R, Y Wong CC, Volta M, Lunnon K, Mill J, Schalkwyk LC: **A data-driven approach to preprocessing Illumina 450K methylation array data.** *BMC Genomics* 2013, **14**:293.

60. Boks MP, Derks EM, Weisenberger DJ, Strengman E, Janson E, Sommer IE, Kahn RS, Ophoff RA: **The relationship of DNA methylation with age, gender and genotype in twins and healthy controls.** *PLoS ONE* 2009, **4**:e6767.

61. Liu J, Morgan M, Hutchison K, Calhoun VD: **A study of the influence of sex on genome wide methylation.** *PLoS ONE* 2010, **5**:e10028.

62. Zhang FF, Cardarelli R, Carroll J, Fulda KG, Kaur M, Gonzalez K, Vishwanatha JK, Santella RM, Morabia A: **Significant differences in global genomic DNA methylation by gender and race/ethnicity in peripheral blood.** *Epigenetics* 2011, **6**:623–629.

63. Liu J, Hutchison K, Perrone-Bizzozero N, Morgan M, Sui J, Calhoun V: **Identification of genetic and epigenetic marks involved in population structure.** *PLoS ONE* 2010, **5**:e13209.

64. Patterson N, Price AL, Reich D: **Population structure and eigenanalysis.** *PLoS Genet* 2006, **2**:e190.

65. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D: **Principal components analysis corrects for stratification in genome-wide association studies.** *Nat Genet* 2006, **38**:904–909.

66. Grunau C, Hindermann W, Rosenthal A: **Large-scale methylation analysis of human genomic DNA reveals tissue-specific differences between the methylation profiles of genes and pseudogenes.** *Hum Mol Genet* 2000, **9**:2651–2663.

67. Herzog E, Galvez J, Roks A, Stolk L, Verbiest M, Eilers P, Cornelissen J, Steegers E, Steegers-Theunissen R: **Tissue-specific DNA methylation profiles in newborns.** *Clin Epigenetics* 2013, **5**:8.

68. Zhang B, Zhou Y, Lin N, Lowdon RF, Hong C, Nagarajan RP, Cheng JB, Li D, Stevens M, Lee HJ, Xing X, Zhou J, Sundaram V, Elliott G, Gu J, Shi T, Gascard P, Sigaroudinia M, Tlsty

- TD, Kadlecsek T, Weiss A, O'Geen H, Farnham PJ, Maire CL, Ligon KL, Madden PA, Tam A, Moore R, Hirst M, Marra MA, Zhang B, et al.: **Functional DNA methylation differences between tissues, cell types, and across individuals discovered using the M&M algorithm.** *Genome Res* 2013, **23**:1522–40.
69. Whitwell JL, Crum WR, Watt HC, Fox NC: **Normalization of cerebral volumes by use of intracranial volume: implications for longitudinal quantitative MR imaging.** *AJNR Am J Neuroradiol* 2001, **22**:1483–1489.
70. Wu JQ, Ruth Sassé T, Wolkenstein G, Conceicao V, Miranda Saksena M, Soedjono M, Perera SS, Wang B, Dwyer DE, Saksena NK: **Transcriptome analysis of primary monocytes shows global down-regulation of genetic networks in HIV viremic patients versus long-term non-progressors.** *Virology* 2013, **435**:308–319.
71. Li Z-H, Xu H, Zheng W, Lam SH, Gong Z: **RNA-sequencing analysis of TCDD-induced responses in zebrafish liver reveals high relatedness to in vivo mammalian models and conserved biological pathways.** *PLoS ONE* 2013, **8**:e77292.
72. Xiong Q, Ancona N, Hauser ER, Mukherjee S, Furey TS: **Integrating genetic and gene expression evidence into genome-wide association analysis of gene sets.** *Genome Res* 2012, **22**:386–397.
73. Vahey MT, Wang Z, Kester KE, Cummings J, Heppner DG, Nau ME, Ofori-Anyinam O, Cohen J, Coche T, Ballou WR, Ockenhouse CF: **Expression of genes associated with immunoproteasome processing of major histocompatibility complex peptides is indicative of protection with adjuvanted RTS,S malaria vaccine.** *J Infect Dis* 2010, **201**:580–589.
74. Yu C-C, Furukawa M, Kobayashi K, Shikishima C, Cha P-C, Sese J, Sugawara H, Iwamoto K, Kato T, Ando J, Toda T: **Genome-wide DNA methylation and gene expression analyses of monozygotic twins discordant for intelligence levels.** *PLoS ONE* 2012, **7**:e47081.
75. Wechsler D: *Wechsler Memory Scale. 3rd Ed.* San Antonio, TX: Psychological Corporation; 1997.
76. Wechsler D: *Wechsler Adult Intelligence Scale. 3rd Ed.* San Antonio, TX: Psychological Corporation; 1997.
77. Axmacher N, Haupt S, Cohen MX, Elger CE, Fell J: **Interference of working memory load with long-term memory formation.** *Eur J Neurosci* 2009, **29**:1501–1513.
78. Long F, Liu H, Hahn C, Sumazin P, Zhang MQ, Zilberstein A: **Genome-wide prediction and analysis of function-specific transcription factor binding sites.** *In Silico Biol* 2004, **4**:395–410.
79. Roig B, Virgos C, Franco N, Martorell L, Valero J, Costas J, Carracedo A, Labad A, Vilella E: **The discoidin domain receptor 1 as a novel susceptibility gene for schizophrenia.** *Mol Psychiatry* 2007, **12**:833–841.
80. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler and D: **The Human Genome Browser at UCSC.** *Genome Res* 2002, **12**:996–1006.
81. Cheng H-YM, Papp JW, Varlamova O, Dziema H, Russell B, Curfman JP, Nakazawa T, Shimizu K, Okamura H, Impey S, Obrietan K: **microRNA modulation of circadian-clock period and entrainment.** *Neuron* 2007, **54**:813–829.

82. Zhao X, He X, Han X, Yu Y, Ye F, Chen Y, Hoang T, Xu X, Mi Q-S, Xin M, Wang F, Appel B, Lu QR: **MicroRNA-mediated control of oligodendrocyte differentiation.** *Neuron* 2010, **65**:612–626.
83. Kocerha J, Faghihi MA, Lopez-Toledano MA, Huang J, Ramsey AJ, Caron MG, Sales N, Willoughby D, Elmen J, Hansen HF, Orum H, Kauppinen S, Kenny PJ, Wahlestedt C: **MicroRNA-219 modulates NMDA receptor-mediated neurobehavioral dysfunction.** *Proc Natl Acad Sci USA* 2009, **106**:3507–3512.
84. Karatsoreos IN: **Links between Circadian Rhythms and Psychiatric Disease.** *Front Behav Neurosci* 2014, **8**:162.
85. Ehrlich S, Geisler D, Yendiki A, Panneck P, Roessner V, Calhoun VD, Magnotta VA, Gollub RL, White T: **Associations of white matter integrity and cortical thickness in patients with schizophrenia and healthy controls.** *Schizophr Bull* 2014, **40**:665–674.
86. Miyata J, Sasamoto A, Koelkebeck K, Hirao K, Ueda K, Kawada R, Fujimoto S, Tanaka Y, Kubota M, Fukuyama H, Sawamoto N, Takahashi H, Murai T: **Abnormal asymmetry of white matter integrity in schizophrenia revealed by voxelwise diffusion tensor imaging.** *Hum Brain Mapp* 2012, **33**:1741–1749.
87. Liu X, Lai Y, Wang X, Hao C, Chen L, Zhou Z, Yu X, Hong N: **Reduced white matter integrity and cognitive deficit in never-medicated chronic schizophrenia: a diffusion tensor study using TBSS.** *Behav Brain Res* 2013, **252**:157–163.
88. Muñoz Maniega S, Lymer GKS, Bastin ME, Marjoram D, Job DE, Moorhead TWJ, Owens DG, Johnstone EC, McIntosh AM, Lawrie SM: **A diffusion tensor MRI study of white matter integrity in subjects at high genetic risk of schizophrenia.** *Schizophr Res* 2008, **106**:132–139.
89. Mori T, Ohnishi T, Hashimoto R, Nemoto K, Moriguchi Y, Noguchi H, Nakabayashi T, Hori H, Harada S, Saitoh O, Matsuda H, Kunugi H: **Progressive changes of white matter integrity in schizophrenia revealed by diffusion tensor imaging.** *Psychiatry Res* 2007, **154**:133–145.
90. Camchong J, Lim KO, Sponheim SR, Macdonald AW: **Frontal white matter integrity as an endophenotype for schizophrenia: diffusion tensor imaging in monozygotic twins and patients' nonpsychotic relatives.** *Front Hum Neurosci* 2009, **3**:35.
91. Kantrowitz J, Javitt DC: **Glutamatergic transmission in schizophrenia: from basic research to clinical practice.** *Curr Opin Psychiatry* 2012, **25**:96–102.
92. Meltzer HY, Rajagopal L, Huang M, Oyamada Y, Kwon S, Horiguchi M: **Translating the N-methyl-D-aspartate receptor antagonist model of schizophrenia to treatments for cognitive impairment in schizophrenia.** *Int J Neuropsychopharmacol* 2013, **16**:2181–2194.
93. Hawrylycz MJ, Lein ES, Guillozet-Bongaarts AL, Shen EH, Ng L, Miller JA, van de Lagemaat LN, Smith KA, Ebbert A, Riley ZL, Abajian C, Beckmann CF, Bernard A, Bertagnolli D, Boe AF, Cartagena PM, Chakravarty MM, Chapin M, Chong J, Dalley RA, Daly BD, Dang C, Datta S, Dee N, Dolbeare TA, Faber V, Feng D, Fowler DR, Goldy J, Gregor BW, et al.: **An anatomically comprehensive atlas of the adult human brain transcriptome.** *Nature* 2012, **489**:391–399.
94. Tremblay M-E, Riad M, Chierzi S, Murai KK, Pasquale EB, Doucet G: **Developmental course of EphA4 cellular and subcellular localization in the postnatal rat**

hippocampus. *J Comp Neurol* 2009, **512**:798–813.

95. Egea J, Klein R: **Bidirectional Eph-ephrin signaling during axon guidance.** *Trends Cell Biol* 2007, **17**:230–238.

96. Xu N-J, Henkemeyer M: **Ephrin reverse signaling in axon guidance and synaptogenesis.** *Semin Cell Dev Biol* 2012, **23**:58–64.

97. Galimberti I, Bednarek E, Donato F, Caroni P: **EphA4 signaling in juveniles establishes topographic specificity of structural plasticity in the hippocampus.** *Neuron* 2010, **65**:627–642.

98. Filosa A, Paixão S, Honsek SD, Carmona MA, Becker L, Feddersen B, Gaitanos L, Rudhard Y, Schoepfer R, Klopstock T, Kullander K, Rose CR, Pasquale EB, Klein R: **Neuron-glia communication via EphA4/ephrin-A3 modulates LTP through glial glutamate transport.** *Nat Neurosci* 2009, **12**:1285–1292.

99. Bowie CR, Harvey PD: **Cognitive deficits and functional outcome in schizophrenia.** *Neuropsychiatr Dis Treat* 2006, **2**:531–536.

100. Keefe RSE, Harvey PD: **Cognitive impairment in schizophrenia.** *Handb Exp Pharmacol* 2012:11–37.

101. Pereira PL, Magnol L, Sahún I, Brault V, Duchon A, Prandini P, Gruart A, Bizot J-C, Chadeaux-Vekemans B, Deutsch S, Trovero F, Delgado-García JM, Antonarakis SE, Dierssen M, Herault Y: **A new mouse model for the trisomy of the Abcg1-U2af1 region reveals the complexity of the combinatorial genetic code of down syndrome.** *Hum Mol Genet* 2009, **18**:4756–4769.

102. Lonard DM, Nawaz Z, Smith CL, O'Malley BW: **The 26S proteasome is required for estrogen receptor-alpha and coactivator turnover and for efficient estrogen receptor-alpha transactivation.** *Mol Cell* 2000, **5**:939–948.

103. Kos M, Denger S, Reid G, Gannon F: **Upstream open reading frames regulate the translation of the multiple mRNA variants of the estrogen receptor alpha.** *J Biol Chem* 2002, **277**:37131–37138.

104. Takeyama J, Suzuki T, Inoue S, Kaneko C, Nagura H, Harada N, Sasano H: **Expression and cellular localization of estrogen receptors alpha and beta in the human fetus.** *J Clin Endocrinol Metab* 2001, **86**:2258–2262.

105. Huang C, Cai Z, Huang M, Mao C, Zhang Q, Lin Y, Zhang X, Tang B, Chen Y, Wang X, Qian Z, Ye L, Peng Y, Xu H: **miR-219-5p Modulates Cell Growth of Papillary Thyroid Carcinoma by Targeting Estrogen Receptor α .** *J Clin Endocrinol Metab* 2014:jc20142883.

106. Knickmeyer RC, Wang J, Zhu H, Geng X, Woolson S, Hamer RM, Konneker T, Lin W, Styner M, Gilmore JH: **Common variants in psychiatric risk genes predict brain structure at birth.** *Cereb Cortex* 2014, **24**:1230–46.

107. Weickert CS, Miranda-Angulo AL, Wong J, Perlman WR, Ward SE, Radhakrishna V, Straub RE, Weinberger DR, Kleinman JE: **Variants in the estrogen receptor alpha gene and its mRNA contribute to risk for schizophrenia.** *Hum Mol Genet* 2008, **17**:2293–2309.

108. Martorell L, Costas J, Valero J, Gutierrez-Zotes A, Phillips C, Torres M, Brunet A, Garrido G, Carracedo A, Guillamat R, Vallès V, Guitart M, Labad A, Vilella E: **Analyses of variants located in estrogen metabolism genes (ESR1, ESR2, COMT and APOE) and**

schizophrenia. *Schizophr Res* 2008, **100**:308–315.

109. Perlman WR, Tomaskovic-Crook E, Montague DM, Webster MJ, Rubinow DR, Kleinman JE, Weickert CS: **Alteration in estrogen receptor alpha mRNA levels in frontal cortex and hippocampus of patients with major mental illness.** *Biol Psychiatry* 2005, **58**:812–824.

110. Perlman WR, Webster MJ, Kleinman JE, Weickert CS: **Reduced glucocorticoid and estrogen receptor alpha messenger ribonucleic acid levels in the amygdala of patients with major mental illness.** *Biol Psychiatry* 2004, **56**:844–852.

111. Amin Z, Canli T, Epperson CN: **Effect of estrogen-serotonin interactions on mood and cognition.** *Behav Cogn Neurosci Rev* 2005, **4**:43–58.

112. Berman KF, Schmidt PJ, Rubinow DR, Danaceau MA, Van Horn JD, Esposito G, Ostrem JL, Weinberger DR: **Modulation of cognition-specific cortical activity by gonadal steroids: a positron-emission tomography study in women.** *Proc Natl Acad Sci USA* 1997, **94**:8836–8841.

113. Gasbarri A, Tavares MCH, Rodrigues RC, Tomaz C, Pompili A: **Estrogen, cognitive functions and emotion: an overview on humans, non-human primates and rodents in reproductive years.** *Rev Neurosci* 2012, **23**:587–606.

114. Wu YC, Hill RA, Gogos A, van den Buuse M: **Sex differences and the role of estrogen in animal models of schizophrenia: interaction with BDNF.** *Neuroscience* 2013, **239**:67–83.

115. Kulkarni J, Gavrilidis E, Worsley R, Hayes E: **Role of estrogen treatment in the management of schizophrenia.** *CNS Drugs* 2012, **26**:549–557.

116. Gregório SP, Sallet PC, Do K-A, Lin E, Gattaz WF, Dias-Neto E: **Polymorphisms in genes involved in neurodevelopment may be associated with altered brain morphology in schizophrenia: preliminary evidence.** *Psychiatry Res* 2009, **165**:1–9.

117. Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM, Nord AS, Kusenda M, Malhotra D, Bhandari A, Stray SM, Rippey CF, Roccanova P, Makarov V, Lakshmi B, Findling RL, Sikich L, Stromberg T, Merriman B, Gogtay N, Butler P, Eckstrand K, Noory L, Gochman P, Long R, Chen Z, Davis S, Baker C, Eichler EE, Meltzer PS, et al.: **Rare Structural Variants Disrupt Multiple Genes in Neurodevelopmental Pathways in Schizophrenia.** *Science* 2008, **320**:539–543.

118. Melas PA, Rogdaki M, Ösby U, Schalling M, Lavebratt C, Ekström TJ: **Epigenetic aberrations in leukocytes of patients with schizophrenia: association of global DNA methylation with antipsychotic drug treatment and disease onset.** *FASEB J* 2012, **26**:2712–2718.

119. Shin C, Nam J-W, Farh KK-H, Chiang HR, Shkumatava A, Bartel DP: **Expanding the microRNA targeting code: functional sites with centered pairing.** *Mol Cell* 2010, **38**:789–802.

120. Geeleher P, Hartnett L, Egan LJ, Golden A, Raja Ali RA, Seoighe C: **Gene-set analysis is severely biased when applied to genome-wide methylation data.** *Bioinformatics* 2013, **29**:1851–1857.

121. Barry WT, Nobel AB, Wright FA: **Significance analysis of functional categories in gene expression studies: a structured permutation approach.** *Bioinformatics* 2005,

21:1943–1949.

122. Efron B, Tibshirani R: **On testing the significance of sets of genes.** *Ann Appl Stat* 2007, **1**:107–129.

123. Sato F, Tsuchiya S, Meltzer SJ, Shimizu K: **MicroRNAs and epigenetics.** *FEBS J* 2011, **278**:1598–1609.

124. Wang W, Kwon EJ, Tsai L-H: **MicroRNAs in learning, memory, and neurological diseases.** *Learn Mem* 2012, **19**:359–368.

125. Gräff J, Kim D, Dobbin MM, Tsai L-H: **Epigenetic regulation of gene expression in physiological and pathological brain processes.** *Physiol Rev* 2011, **91**:603–649.

126. Yang J-H, Li J-H, Shao P, Zhou H, Chen Y-Q, Qu L-H: **starBase: a database for exploring microRNA-mRNA interaction maps from Argonaute CLIP-Seq and Degradome-Seq data.** *Nucleic Acids Res* 2011, **39**(Database issue):D202–209.

127. Li J-H, Liu S, Zhou H, Qu L-H, Yang J-H: **starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data.** *Nucleic Acids Res* 2014, **42**(Database issue):D92–97.

Table 1. Demographic data, clinical variables and intermediate phenotypes

Sample	Size	Sex		Age	WRAT3-RT	Parental SES	Handedness	Medication	Hippocampal	STG	DLPFC	SIRP
		(female)							volume	thickness	activity	performance
		N	%						Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
SZ	103	28	27.2	34.55 (10.60)	46.25 (7.01)	2.87 (1.01)	1.51 (3.52)	57.30 (123.07)	4139.64 (471.04)	2.79 (0.16)	0.05 (0.10)	95.21 (4.99)
HC	111	40	36.0	32.05 (10.83)	51.03 (3.97)	2.68 (0.72)	0.79 (2.38)	-	4439.49 (427.69)	2.90 (0.16)	0.06 (0.10)	98.46 (1.62)
Total	214	68	31.8	33.25 (10.77)	48.78 (6.09)	2.77 (0.87)	1.14 (2.99)	-	4295.17 (472.51)	2.85 (0.17)	0.06 (0.10)	96.90 (3.99)

Means and standard deviations (SD) are given. SZ = patient with schizophrenia; HC = healthy control. WRAT3-RT = Wide Range Achievement Test 3 - Reading Test. Parental SES (socioeconomic status) was classified according to Hollingshead, and handedness determined using the Annett Scale of Hand Preference. DLPFC activity = mean working memory load-dependent neural activity (% BOLD signal change) in the DLPFC during the Sternberg Item Recognition Paradigm (SIRP). SIRP performance represents the mean accuracy during the working memory task across all loads (1, 3, 5). * significantly different between SZ and HC on basis of Welch's *t* test ($p < 0.05$). For DLPFC activity and SIRP performance, data were only available for 203 and 209 participants, respectively.

Table 2. MicroRNA target gene sets significantly enriched for hippocampal volume (positively ranked list metric; Pearson correlation)

Gene Set	Size	ES	NES	NOM p-val	FDR
GACAATC, HSA-MIR-219A-5P	22	0.60	1.94	0.002	0.050
GGCCAGT,MIR-193A,MIR-193B	8	0.73	1.72	0.004	0.246
TCCAGAT,MIR-516-5P	12	0.63	1.72	0.014	0.194

Names of gene sets as given by the Molecular Signatures Database of GSEA. Size = number of genes included in each gene set (restricted by the number of CpG sites and associated genes included in the MCIC data); ES = enrichment score; NES = normalized enrichment score (adjusted for different sizes of gene sets); NOM p-val = nominal p-value; FDR = false discovery rate q-value. FDR threshold = 0.083 (= 0.25/3;corrected for number of tested intermediate phenotypes).

Table 3. Full list of genes included in the hsa-miR-219a-5p gene set (as analyzed by GSEA)

Gene Name	Location	Rank in Gene List	Rank Metric Score	Running ES	HGNC Gene Name /Aliases	Experimental Evidence
KIAA0182	16q24.1	31	0.148	0.108	GSE1; Gse1 coiled-coil protein	predicted
PKNOX1	21q22.3	48	0.140	0.215	PBX/knotted 1 homeobox 1	predicted
TRAF7	16p13.3	55	0.137	0.323	TNF receptor-associated factor 7, E3 ubiquitin protein ligase	predicted
CBFA2T3	16q24	98	0.118	0.403	core-binding factor, runt domain, alpha subunit 2; translocated to, 3	predicted
EPHA4	2q36.3	179	0.102	0.456	EPH receptor A4	predicted
MKNK2	19p13.3	373	0.076	0.448	MAP kinase interacting serine/threonine kinase 2	starBase CLIP-seq
DDAH1	1p22	413	0.072	0.492	dimethylargininedimethylaminohydrolase 1	starBase CLIP-seq
ESR1	6q24-q27	416	0.072	0.549	estrogen receptor 1	Luciferase Reporter Assay [105]
AKAP13	15q24-q25	426	0.071	0.603	A kinase (PRKA) anchor protein 13	starBase CLIP-seq
PHACTR2	6q24.1	788	0.039	0.505	phosphatase and actin regulator 2	starBase CLIP-seq
SLC31A1	9q32	1028	0.022	0.438	solute carrier family 31 (copper transporter), member 1	starBase CLIP-seq
KBTBD8	3p14	1086	0.019	0.433	kelch repeat and BTB (POZ) domain containing 8	starBase CLIP-seq
ADD2	2p13.3	1148	0.016	0.423	adducin 2 (beta)	predicted

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

ERGIC1	5q35.1	1223	0.011	0.406	endoplasmic reticulum-golgi intermediate compartment (ERGIC) 1	starBase CLIP-seq
RBM24	6p22.3	1260	0.009	0.401	RNA binding motif protein 24	predicted
KIAA0240	6p21.1	1318	0.005	0.384	GLTSCR1L; GLTSCR1-like	predicted
NEK6	9q33.3-q34.11	1451	-0.003	0.339	NIMA-related kinase 6	starBase CLIP-seq
THRB	3p24.2	1551	-0.009	0.311	thyroid hormone receptor, beta	starBase CLIP-seq
SCN5A	3p21	1646	-0.015	0.290	sodium channel, voltage-gated, type V, alpha subunit	predicted
TGFBR2	3p22	1907	-0.032	0.222	transforming growth factor, beta receptor II (70/80kDa)	starBase CLIP-seq
FZD4	11q14-q21	2324	-0.063	0.124	frizzled class receptor 4	predicted
SEMA4G	10q24.31	2354	-0.066	0.167	sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, semaphoring) 4G	predicted

Genes contributing to the core enrichment of the predicted hsa-miR-219a-5p gene set (leading edge subset) are bold. HGNC = HUGO (HUMAN Genome Organization) Gene Nomenclature Committee. ES = enrichment score. **The last column indicates whether target genes were predicted based on bioinformatic algorithms (labelled as “predicted”) or are experimentally validated, e.g. by the starBase [126, 127] prediction intersection tool using CLIP-Seq (cross-linking immunoprecipitation-high-throughput sequencing) data.**

Figure Legends

Figure 1. Workflow of statistical analysis

GSEA (gene set enrichment analysis) procedure. ES = enrichment score.

Figure 2. Enrichment plot: Graphical view of the enrichment score (ES) for the predicted hsa-miR-219a-5p target gene set

Profile of the running ES and positions of gene set members on the rank ordered list (according to Pearson correlation with hippocampal volume). See SM 2.1 and GSEA documentation (<http://www.broadinstitute.org/gsea/doc/GSEAUUserGuideFrame.html>) for detailed information on how to interpret enrichment plots.

Figure1

Figure2