GENETICS AND GENOMICS

Genome-wide DNA methylation risk scores for schizophrenia derived from blood and brain tissues further explain the genetic risk in patients stratified by polygenic risk scores for schizophrenia and bipolar disorder

Kazutaka Ohi,1,2 Mihoko Shimada,3 Midori Soda,4 Daisuke Nishizawa,5 Daisuke Fujikane,1 Kentaro Takai,1 Ayumi Kuramitsu,1 Yukimasa Muto,1 Shunsuke Sugiyama,1 Junko Hasegawa,5 Kiyoyuki Kitaichi,4 Kazutaka Ikeda,5 Toshiki Shioiri1

ABSTRACT

Background Genetic and environmental factors contribute to the pathogenesis of schizophrenia (SZ) and bipolar disorder (BD). Among genetic risk groups stratified by combinations of Polygenic Risk Score (PRS) deciles for SZ, BD and SZ versus BD, genetic SZ risk groups had high SZ risk and prominent cognitive impairments. Furthermore, epigenetic alterations are implicated in these disorders. However, it was unclear whether DNA Methylation Risk Scores (MRSs) for SZ risk derived from blood and brain tissues were associated with SZ risk, particularly the PRS-stratified genetic SZ risk group.

Methods Epigenome-wide association studies (EWASs) of SZ risk in whole blood were preliminarily conducted between 66 SZ patients and 30 healthy controls (HCs) and among genetic risk groups (individuals with low genetic risk for SZ and BD in HCs (n=30) and in SZ patients (n=11), genetic BD risk in SZ patients (n=25) and genetic SZ risk in SZ patients (n=30)) stratified by combinations of PRSs for SZ, BD and SZ versus BD. Next, differences in MRSs based on independent EWASs of SZ risk in whole blood, postmortem frontal cortex (FC) and superior temporal gyrus (STG) were investigated among our case–control and PRS-stratified genetic risk status groups.

Results Among case–control and genetic risk status groups, 33 and 351 genome-wide significant differentially methylated positions (DNPs) associated with SZ were identified, respectively, many of which were hypermethylated. Compared with the low genetic risk in HCs group, the genetic SZ risk in SZ group had 39 genome-wide significant DMPs, while the genetic BD risk in SZ group had only six genome-wide significant DMPs. The MRSs for SZ risk derived from whole blood, FC and STG were higher in our SZ patients than in HCs in whole blood and were particularly higher in the genetic SZ risk in SZ group than in the low genetic risk in HCs and genetic BD risk in SZ groups. Conversely, the MRSs for SZ risk based on our whole-blood EWASs among genetic risk groups were also associated with SZ in the FC and STG. There were no correlations between the MRSs and PRSs.

WHAT IS ALREADY KNOWN ON THIS TOPIC
⇒ Genetic and environmental factors contribute to the pathogenesis of schizophrenia (SZ) and bipolar disorder (BD). Genetic SZ risk groups stratified by combinations of Polygenic Risk Scores (PRSs) for SZ, BD and SZ versus BD had high SZ risk and prominent cognitive impairments.

WHAT THIS STUDY ADDS
⇒ Genome-wide DNA Methylation Risk Scores (MRSs) for the risk of SZ would be associated with the risk of SZ beyond tissue differences between the blood and brain. Furthermore, the MRSs for the risk of SZ explained the risk among genetic risk groups even after stratifying by PRSs for SZ, BD and SZ versus BD.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY
⇒ MRSs would be genetic markers for assessing an individual’s risk of SZ independent of PRSs.

Conclusions These findings suggest that the MRS is a potential genetic marker in understanding SZ, particularly in patients with a genetic SZ risk.

INTRODUCTION

Schizophrenia (SZ) and bipolar disorder (BD) are clinically and genetically heterogeneous psychiatric disorders with a lifetime prevalence of approximately 1%1 2 and an estimated heritability of approximately 80%.3 4 Genetic and environmental factors contribute to the pathogenesis of both disorders. Recent approaches to understanding the causes of these disorders have focused on explaining the genetic contribution to the disorder. To date, the latest and largest-scale genome-wide association study (GWAS) of SZ by the Psychiatric Genomics

Correspondence to Dr Kazutaka Ohi, Department of Psychiatry, Gifu University Graduate School of Medicine, Gifu, Japan; ohikazutaka.h@yg.fu-u.ac.jp

Received 9 November 2023
Accepted 14 December 2023

© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. Published by BMJ.


BMJ Mental Health 2024;27:1–11. doi:10.1136/bmjment-2023-300936
Consortium wave 3 (PGC3) has identified 287 distinct genomic loci for SZ in 79 755 SZ patients and 243 649 controls. Similarly, a large-scale GWAS of BD by the PGC3 identified 64 independent genomic loci implicated in BD in 41 917 BD patients and 371 549 controls. Genetic commonalities between SZ and BD were indicated by a shared genetic variation and a high degree of polygenicity of approximately 70%-80% between SZ and BD. In contrast, the PGC has identified two distinct genetic loci differentiating SZ from BD (SZ vs BD), that is, SZ-specific genetic loci. The Polygenic Risk Scores (PRSs) that quantify the combined effects of the single-nucleotide polymorphisms (SNPs) across the whole genome for differentiating SZ from BD were associated with the risk of SZ and low premorbid intelligence.

We indicated that the PRSs for SZ, BD and SZ versus BD based on the GWASs from the PGC were higher in SZ patients than in healthy controls (HCs). On the other hand, the PRSs between the case–control groups highly overlapped. Therefore, we divided participants into decile category groups based on PRSs and found that individuals in higher deciles for SZ PRSs had up to approximately fourfold, sixfold and twofold higher risk of SZ and PRSs for BD and SZ versus BD than those in lower deciles for SZ PRSs, respectively. In addition, we investigated differences in cognitive impairments that were a core feature of SZ and BD and that were more prominent in SZ patients than in BD patients among three genetic risk groups: the genetic SZ risk, genetic BD risk and low genetic risks for SZ and BD groups, stratified by combinations of the PRSs for SZ, BD and SZ versus BD. In support of the evidence for a genetic correlation between cognitive impairments and the risk of SZ but not BD, we found that individuals who had a high genetic susceptibility for SZ in SZ patients displayed worse cognitive performances than those who had a high genetic susceptibility for BD in SZ patients. These findings suggest that increased genetic loading specifically for SZ but not for BD in SZ patients is associated with more severe cognitive impairments, and stratifying patients based on combinations of PRSs would have substantial clinical utility.

Heritability estimates from GWASs explain a part of the heritability predicted by twin studies, suggesting that genetic markers other than SNPs may contribute to the risk of these disorders. Indeed, a high burden of both rare genetic variants and rare copy number variants has been found in SZ patients. Furthermore, environmental risks, such as prenatal infection, obstetric complications and childhood maltreatment, contribute to susceptibility to SZ and BD via interaction with the genetic risk. Epigenetic processes, which refer to functional alterations in the genome that do not alter the nucleotide sequence, also have a role in susceptibility to SZ and BD. DNA methylation (DNAm) is an epigenetic modification that is affected by both genetic and environmental effects and can regulate increased or decreased gene expression. DNAm is the addition of a methyl group to the carbon-5 position of a cytosine residue in DNA. DNAm is most commonly observed within the context of cytosine-phosphate-guanine dinucleotides (CpGs) but also within the context of cytosine-phosphate-H dinucleotides (CpHs, where H = cytosine, adenine or thymine) in the current study. There is some evidence that specific DNAm is potentially reversible after the onset as well as through the life course in brain disorders. Altered DNAm patterns were revealed in SZ patients and in BD patients. Furthermore, multiple differentially methylated loci in peripheral blood have been identified in monozygotic twins discordant for SZ and BD. Some epigenetic profiles differ in the brain compared with peripheral tissue; however, other DNAm sites have common patterns, which would make these regions ideal as potential biomarkers for psychiatric disorders. DNAm studies of psychiatric disorders have been more widely performed in peripheral tissues because they can be readily obtained from living patients. As interindividual variation in DNAm was reflected across both the brain and blood, it was suggested that peripheral tissues may have utility in epidemiological studies of psychiatric disorders.

To date, a large-scale meta-analysis of epigenome-wide association studies (EWASs) of blood-derived DNAm patterns for SZ has been conducted in a total of 1714 individuals from three independent cohorts, and the researchers identified 1223 differential DNAm sites associated with SZ in their discovery cohort (p < 5.00 x 10^-8), with over 90% of these DNAm sites having a consistent direction of effect across the discovery and two independent replication cohorts. In general, methodology to calculate the PRS from GWAS summary statistics was established. To estimate the Methylation Risk Score (MRS) from EWAS summary statistics, a similar methodology as that for the PRS could be applied to methylated data, although the methodology for the MRS has not yet been established. MRSs for the risk of SZ based on differential DNAm sites in the EWAS meta-analysis of SZ in whole blood were calculated at both stringent and liberal significance thresholds of individual DNAm sites (P_c < 1.00 x 10^-5) in SZ patients relative to 59 BD patients and 55 HCs in whole blood, suggesting that the genome-wide methylation patterns may represent distinct pathophysiology between SZ and BD. However, the number of DNAm sites for the MRS calculation, especially at liberal thresholds, was limited (P_c < 0.5, n = 1208) because their EWAS meta-analysis was only based on marginally associated DNAm sites in the discovery cohort (p < 1.00 x 10^-5). In addition to peripheral whole blood, EWASs in brain tissues, including the frontal cortex (FC) and superior temporal gyrus (STG), have previously been reported in SZ patients. Common differential DNAm sites between peripheral blood and brain tissues in SZ patients have been reported. However, it is unknown whether genome-wide MRSs for SZ derived from brain tissues other than peripheral blood are also associated with the risk of SZ in whole blood. Furthermore, it is unclear whether the MRSs derived from whole blood and brain tissues are more strongly associated with SZ in the genetic risk groups stratified by PRSs. There were few overlaps between the significant differentially methylated positions (DMPs) associated with SZ case status and DMPs associated with the PRS for SZ. We hypothesised that the high MRSs for the risk of SZ in the brain, as well as blood tissues, would be associated with the risk of SZ in case–control groups. Moreover, given that the disorder-associated and PRS-associated DMPs were distinct and that increased genetic loading specifically for SZ but not for BD is associated with more severe cognitive impairments in SZ patients, we hypothesised that the high MRSs for the risk of SZ in the brain and blood tissues would further explain the risk of SZ in the genetic risk groups stratified by PRSs.

In this study, to identify differential DNAm sites associated with the risk of SZ, we preliminarily performed EWASs of the risk of SZ in peripheral blood among 66 SZ patients and 30 HCs consisting of genetic risk groups (individuals with low genetic risk in HCs (n = 30) and in SZ (n = 11), genetic BD risk in SZ (n = 25) and genetic SZ risk in SZ (n = 30) stratified by combinations of PRS deciles for SZ, BD and SZ versus BD. Next, we calculated the MRSs based on differential DNAm sites at different P thresholds in EWASs of current smoking and risk of SZ derived from whole blood, postmortem FC and...
STG samples. After confirming that the current smoking-derived MRS could predict current smokers, we investigated whether the MRSs for the risk of SZ in whole blood, postmortem FC and STG tissues were associated with case-control status and PRS-stratified genetic risk status. We conversely examined whether the MRSs based on our EWASs of the risk of SZ in peripheral blood between case-control status and among the genetic risk status groups were associated with the diagnostic status of SZ in postmortem FC and STG. Furthermore, we explored whether MRSs and PRSs were distinct genetic markers.

METHODS

Study sample description

Among previous study subjects (n=439) composed of 173 patients with SZ, 70 of their unaffected first-degree relatives and 196 HCs,14 the current study sample (n=96) was extracted based on combinations of PRS deciles for SZ, BD and SZ versus BD (see the Extraction of genetic risk groups section). The current study sample (n=96) was composed of individuals with low genetic risks of SZ and BD in HCs (n=30, low genetic risk in HCs), those with low genetic risks of SZ and BD in SZ patients (n=11, low genetic risk in SZ), those with high genetic risk of BD in SZ patients (n=25, genetic BD risk in SZ) and those with high genetic risk of SZ in SZ patients (n=30, genetic SZ risk in SZ) (table 1). A detailed description of participant recruitment and diagnosis has been provided previously.12 14 41 42 Briefly, all participants were of Japanese descent and had no biological first-degree or second-degree relatives among the participants. The participants were recruited from the Schizophrenia Non-Affected Relative Project.12–14 41–46 The patients were recruited from Kanazawa Medical University Hospital. Patients were diagnosed with SZ according to the criteria in the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5). HCs were evaluated using the Structured Clinical Interview for DSM-IV-Non-Patient version (SCID-NP) to exclude individuals who had current or past contact with psychiatric services, had received psychiatric medication or had any family history of neuropsychiatric disorder within their first-degree or second-degree relatives.
Genotyping, imputation and PRS calculations for SZ, BD and SZ versus BD
A detailed description of the genotyping, quality control (QC) and imputation procedures applied in the study sample has been provided previously. Briefly, peripheral venous blood was collected from the participants, and genomic DNA was extracted from the whole-blood samples. Genotyping was performed using Infinium OmniExpressExome-8 V1.1 or V1.6 BeadChips (Illumina, San Diego, California, USA). After the QC and imputation procedures, 8,741,088 SNPs were retained for PRS analyses.

We calculated three PRSs for SZ, BD and SZ versus BD in our study sample using three publicly available GWAS data-sets (PGC3 SZ,5 PGC3 BD and SZ vs BD). According to the standard manufacturer’s instructions, methylation quantification and QC

Methylation quantification and QC
According to the standard manufacturer’s instructions, methylation quantification was undertaken using the Illumina Infinium Human Methylation EPIC BeadChip (Illumina, San Diego, California, USA) on whole-blood DNA samples from 96 individuals with low genetic risk, genetic BD risk and genetic SZ risk. To avoid bias due to slides or position on slides, DNA samples were randomised on the BeadChip.

Quantification was completed using the Qubit fluorimeter, and bisulphite conversions were completed using the EZ DNA Methylation workflow. It produced a methylation score as a β value.

Table 1 Demographic characteristics among individuals with low genetic risks for SZ and BD in HCs (low genetic risk in HCs), those with low genetic risks for SZ and BD in SZ patients (low genetic risk in SZ), those with high genetic risk for BD in SZ patients (genetic BD risk in SZ) and those with high genetic risk for SZ in SZ patients (genetic SZ risk in SZ)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Low genetic risk in HCs (n=30)</th>
<th>Low genetic risk in SZ (n=11)</th>
<th>Genetic BD risk in SZ (n=25)</th>
<th>Genetic SZ risk in SZ (n=30)</th>
<th>P value (F or χ²)</th>
<th>Post hoc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>41.8±13.2</td>
<td>43.5±8.4</td>
<td>43.4±13.0</td>
<td>45.5±16.3</td>
<td>0.78 (0.4)</td>
<td>–</td>
</tr>
<tr>
<td>Sex (male/female)</td>
<td>18/12</td>
<td>4/7</td>
<td>12/13</td>
<td>14/16</td>
<td>0.53 (2.2)*</td>
<td>–</td>
</tr>
<tr>
<td>Education (years)</td>
<td>16.2±5.5</td>
<td>11.5±1.8</td>
<td>12.5±2.5</td>
<td>12.4±2.1</td>
<td>6.19×10⁻⁴ (14.8)</td>
<td>Low HC&gt;low SZ, BD, SZ</td>
</tr>
<tr>
<td>Estimated premorbid IQ</td>
<td>108.5±8.0</td>
<td>95.5±15.5</td>
<td>99.8±12.8</td>
<td>96.8±8.6</td>
<td>3.80×10⁻⁴ (6.8)</td>
<td>Low HC&gt;low SZ, BD, SZ</td>
</tr>
<tr>
<td>Age at onset (years)</td>
<td>24.5±10.0</td>
<td>25.4±9.5</td>
<td>29.3±11.4</td>
<td>28.1±13.3</td>
<td>0.28 (1.3)</td>
<td>–</td>
</tr>
<tr>
<td>Duration of illness (years)</td>
<td>19.0±9.6</td>
<td>18.0±14.0</td>
<td>16.2±14.2</td>
<td>0.81 (0.2)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>CPZ-eq (mg/day)</td>
<td>–</td>
<td>659.9±598.3</td>
<td>474.4±424.5</td>
<td>517.6±444.9</td>
<td>0.55 (0.6)</td>
<td>–</td>
</tr>
<tr>
<td>BPD-eq (mg/day)</td>
<td>–</td>
<td>0.4±0.9</td>
<td>0.9±2.1</td>
<td>0.4±1.3</td>
<td>0.49 (0.7)</td>
<td>–</td>
</tr>
<tr>
<td>PANSS positive symptoms</td>
<td>–</td>
<td>16.5±4.8</td>
<td>15.9±5.7</td>
<td>16.1±5.2</td>
<td>0.96 (0.1)</td>
<td>–</td>
</tr>
<tr>
<td>PANSS negative symptoms</td>
<td>–</td>
<td>23.2±7.5</td>
<td>19.5±8.2</td>
<td>19.3±6.4</td>
<td>0.30 (1.2)</td>
<td>–</td>
</tr>
<tr>
<td>Current smoker (±)</td>
<td>4/26</td>
<td>3/8</td>
<td>6/19</td>
<td>4/26</td>
<td>0.54 (2.2)*</td>
<td>–</td>
</tr>
<tr>
<td>B cells</td>
<td>0.04±0.02</td>
<td>0.06±0.03</td>
<td>0.04±0.03</td>
<td>0.05±0.02</td>
<td>0.42 (0.9)</td>
<td>–</td>
</tr>
<tr>
<td>CD4+ T cells</td>
<td>0.14±0.06</td>
<td>0.11±0.05</td>
<td>0.14±0.07</td>
<td>0.13±0.05</td>
<td>0.48 (0.8)</td>
<td>–</td>
</tr>
<tr>
<td>CD8+ T cells</td>
<td>0.05±0.04</td>
<td>0.02±0.03</td>
<td>0.04±0.04</td>
<td>0.04±0.04</td>
<td>0.20 (1.6)</td>
<td>–</td>
</tr>
<tr>
<td>Monocytes</td>
<td>0.07±0.02</td>
<td>0.08±0.03</td>
<td>0.07±0.03</td>
<td>0.08±0.02</td>
<td>0.43 (0.9)</td>
<td>–</td>
</tr>
<tr>
<td>Neutrophils</td>
<td>0.63±0.09</td>
<td>0.65±0.07</td>
<td>0.63±0.11</td>
<td>0.63±0.09</td>
<td>0.88 (0.2)</td>
<td>–</td>
</tr>
<tr>
<td>Natural killer cells</td>
<td>0.07±0.05</td>
<td>0.08±0.04</td>
<td>0.07±0.05</td>
<td>0.08±0.04</td>
<td>0.71 (0.5)</td>
<td>–</td>
</tr>
</tbody>
</table>

Six cell types (B cells, CD4+ T cells, CD8+ T cells, monocytes, neutrophils and natural killer cells) in peripheral blood were estimated using the Illumina HumanMethylation EPIC microarray. Complete demographic information was not obtained for all participants (estimated premorbid IQ in low genetic risk in HCs, n=26, in genetic BD risk in SZ, n=24 and in genetic SZ risk in SZ, n=29). Means±SDs are shown.

Six cell types (B cells, CD4+ T cells, CD8+ T cells, monocytes, neutrophils and natural killer cells) in peripheral blood were estimated using the Illumina HumanMethylation EPIC microarray. Complete demographic information was not obtained for all participants (estimated premorbid IQ in low genetic risk in HCs, n=26, in genetic BD risk in SZ, n=24 and in genetic SZ risk in SZ, n=29). Means±SDs are shown. 

P values<0.05 are shown in boldface, and post hoc analysis was performed. 

Table 1. Demographic characteristics among individuals with low genetic risks for SZ and BD in HCs (low genetic risk in HCs), those with low genetic risks for SZ and BD in SZ patients (low genetic risk in SZ), those with high genetic risk for BD in SZ patients (genetic BD risk in SZ) and those with high genetic risk for SZ in SZ patients (genetic SZ risk in SZ).
(U) probes defined as $\beta = M/(M+U)$. The R package ‘mefil’ (Efficient algorithms for analysing DNAm data) for methylation QC procedures was used to extract signal intensities and perform background correction and normalisation of methylation $\beta$-values for individual DNAm sites \cite{47} (https://github.com/perishky/mefil). The IDAT files were loaded into the R environment using mefil. We set parameters for the QC of the raw data as follows: (i) beadnum.samples.threshold (maximum threshold on the fraction of probes with too few detected beads) = 0.1; (ii) detectionp.samples.threshold (maximum threshold on the fraction of undetected probes) = 0.1; (iii) beadnum.cpgs.threshold (used to identify poor-quality probes in terms of the fraction of samples in which the probe has too few detected beads) = 0.1; (iv) detectionp.cpgs.threshold (used to identify poor-quality probes in terms of the fraction of samples in which the probe is undetected) = 0.1; (v) sex.outlier.sd (number of standard deviations to determine whether a sample is a sex outlier) = 10; (vi) snp.concordance.threshold (concordance threshold to include SNPs to calculate sample concordance using SNPs extracted from the methylation data and the Infinium OmniExpressExome data) = 0.95; and (vii) sample.genotype.concordance.threshold (concordance threshold to determine whether a sample is an outlier) = 0.8.

According to these QC parameters, no samples were excluded prior to performing quantile normalisation. Individual probes were excluded if (1) >10% of samples showed detection p values >0.01 for the probes (n = 1585) or (2) >10% of samples demonstrated less than three beads at DNAm sites (n = 806). Next, we determined the number of principal components (PCs) of control probes for adjustment of the technical measurement variance to use for functional normalisation. Functional normalisation was performed after including the PCs (in this case, the first nine PCs) and the effect of slides. Probes at sex chromosomes were removed for further analysis. After application of these QC procedures, 844,032 DNAm sites and 96 individuals remained for methylation analysis.

Six blood cell counts for B cells, CD4+ T cells, CD8+ T cells, monocytes, neutrophils and natural killer cells were estimated derived from the methylation data using meffil’s ‘blood gse35069 complete’ profile reference.

**EWASs between SZ patients and HCs and among genetic risk groups**

Normalised $\beta$-values were converted to M values via logit transformation for EWASs \cite{48} using the minfi R package (https://github.com/hansenlab/minfi). DMPs between SZ patients and HCs and among the PRS-stratified genetic risk groups were investigated using a linear regression analysis with each M value at DNAm sites as the dependent variable, diagnostic status or genetic risk status as the independent variable, and age, sex, current smoking status, blood cell counts and surrogate variables as covariates using the meffil R package. A genome-wide significance threshold of $p < 9.00 \times 10^{-8}$ was used as the significance threshold that was proposed to control for the false positive rate for EPIC array DNAm studies.\cite{49}

**MRSs for current smoking and risk of SZ derived from peripheral whole blood, postmortem FC and postmortem STG**

Because current cigarette smoking status largely affects DNAm patterns,\cite{50} we calculated the MRS for current cigarette smoking in our study sample using previously described methods\cite{38} as a positive control. To identify multiple risk DNAm sites for current smoking and their p values and effect sizes (coefficient beta values; %changes in DNAm), the available EWAS dataset of current compared with never smokers\cite{50} was used as a discovery sample for the MRS. Mean effect sizes at 187 DMPs were calculated between the discovery and replication panels in the EWAS. To remove highly correlated DNAm sites in our study sample, comethylated regions were calculated using the CoMeBack method,\cite{52} which identifies comethylated regions based on a Spearman correlation cut-off of 0.3 and the proximity of DNAm sites (DNAm sites within a 2 kb proximity window). In the two or more correlated DNAm sites within 2 kb, a single DNAm site demonstrating the strongest association with smoking as indicated by the meta-analysis p value was retained. After extracting DNAm sites identical to the discovery EWAS sample and applying the CoMeBack method, 147 independent DNAm sites remained in our study sample. We calculated the MRS constructed from DNAm sites showing a nominal association with current smoking in the discovery EWAS under a liberal significance threshold ($P_x \leq 1$). For each participant included in our study sample, an MRS was calculated by weighting the scores for the ‘risk DNAm sites’ by the effect sizes observed for the discovery EWAS. The score, consisting of the methylation degree of risk DNAm sites multiplied by the effect sizes, was summed over all the DNAm sites in the $P_x$ DNAm set for individuals in the study sample.

We calculated the MRSs for the risk of SZ derived from whole blood in our study sample using summary statistics from the EWAS meta-analysis of SZ in whole blood using the Illumina Infinium HumanMethylation450 BeadChip\cite{53} as a discovery EWAS sample. The EWAS meta-analysis was carried out across three independent cohorts (860 SZ patients and 854 HCs) comprising a discovery cohort, a large replication cohort and a secondary replication cohort of 96 monozygotic twin pairs discordant for SZ.\cite{54} The meta-analysis was restricted to 1223 DNAm sites identified at a relaxed threshold of $p < 5.00 \times 10^{-5}$ in the discovery cohort. In the discovery EWAS, 1223 risk DNAm sites for the risk of SZ and their p values and effect sizes were identified. After extracting DNAm sites identical to the discovery EWAS sample and applying the CoMeBack method, 1139 independent DNAm sites remained in our study sample. MRSs constructed from DNAm sites showed a nominal association with SZ in the discovery EWAS meta-analysis under the liberal significance thresholds: $P_x < 0.001$, $P_x < 0.01$, $P_x < 0.05$, $P_x < 0.1$, $P_x < 0.2$, $P_x < 0.5$ and $P_x \leq 1$. In the $P_x$ DNAm sets, the MRSs were calculated for each individual in the study sample.

Furthermore, to calculate the MRSs for the risk of SZ derived from postmortem FC and STG in our study sample, we used raw IDAT files for postmortem FC in SZ patients and HCs using the Illumina Infinium HumanMethylation450 BeadChip (GSE61107_RAW)\cite{55} and postmortem planum temporale of the STG grey matter in the right hemisphere in SZ patients and HCs using the Illumina Infinium HumanMethylationEPIC BeadChip (GSE144910_RAW)\cite{56} from the Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/). According to the same methylation quantification, QC, EWASs and MRS calculation procedures, including the CoMeBack method mentioned above, except for estimations of cells count for neuronal (NeuN+) and nonneuronal (NeuN−) cells using meffil’s ‘guintivano dlpfc’ profile reference, we performed EWASs of the risk of SZ in FC (21 SZ patients and 24 HCs after QC, online supplemental table 1) and in STG (43 SZ patients and 43 HCs after QC, online supplemental table 1), with age, sex, race (only GSE144910), postmortem interval hours and information on NeuN+ and NeuN− cells as covariates. Using summary statistics of the EWASs, the MRSs for the risk of SZ derived from FC and STG were calculated in our study sample. The number of DNAm sites at the seven $P_x$ thresholds in MRS calculations for the risk of SZ derived from whole blood, FC and STG are shown in online supplemental table 2.
Statistical analyses
All statistical analyses were performed using IBM SPSS Statistics V28.0 software (IBM). Differences in continuous variables, such as age and years of education, among genetic risk groups were analysed using an analysis of variance. Post hoc tests with Fisher’s least significant difference were used to evaluate significant differences among the groups. Differences in categorical variables, such as sex and smoking status, were analysed using Pearson’s χ² test. To examine whether the MRs based on each PT were associated with current smoking, case-control or genetic risk statuses in our study sample, we performed logistic regression analyses with smoking status (current smokers and non-smokers), diagnostic status (SZ patients and HCs) or genetic risk statuses (low genetic risk in HCs and genetic BD risk in SZ, or low genetic risk in HCs and genetic SZ risk in SZ) as the dependent variable, each MR based on EWASs as the independent variables, and age, sex, current smoking status (when examining case-control and genetic risk status groups) and blood cell counts as covariates. The variance in the smoking status or risk of SZ explained by the MRs is indicated by Nagelkerke’s pseudo-R². To determine the variance accounted for only by the MRs, we subtracted the Nagelkerke’s pseudo-R² of only covariates from those of these models. To illustrate the results in the figures, MRs corrected for covariates were z-standardised. The nominal significance level for all statistical tests was set at p<0.05. The MRs at each PT were highly correlated with each other and were not independent. Therefore, the p values based on different PT values were not corrected. A Bonferroni-corrected p value threshold of p<8.33×10⁻⁴ (α=0.05/three tissue-based MRs for the risk of SZ and three comparisons of diagnostic and genetic risk status groups) was used to avoid type I error.

RESULTS
EWASs of the risk of SZ between case–control status and among genetic risk groups stratified by PRSs
Although the sample sizes of our study cohort were relatively small, we first preliminarily performed EWAS of the risk of SZ in peripheral blood between 66 SZ patients and 30 HCs. We identified 33 genome-wide significant DMPs associated with SZ (figure 2 and online supplemental table 3A, p<9.00×10⁻⁸). Patients with SZ had 19 increased DMPs and 14 decreased DMPs compared with HCs, indicating that 57.6% of DMPs were hypermethylated in SZ patients.

Next, we performed EWAS of the risk of SZ among four PRS-stratified genetic risk groups (low genetic risk in HCs<low genetic risk in SZ<genetic BD risk in SZ<genetic SZ risk in SZ). Ten genome-wide significant DMPs associated with SZ were identified (figure 2 and online supplemental table 3B, p<9.00×10⁻⁸). As the sample size of the low genetic risk in SZ was limited and the order among four genetic risk groups might be incorrect, the low genetic risk in SZ group was excluded, and we repeatedly performed EWAS of the risk of SZ among three genetic risk groups (low genetic risk in HCs<genetic BD risk in SZ<genetic SZ risk in SZ). After excluding the low genetic risk in the SZ group, more genome-wide significant DMPs associated with genetic SZ risk (n=351 DMPs) were identified (figure 2 and online supplemental table 3C, p<9.00×10⁻⁸). Genetic SZ risk in the SZ group had 277 increased DMPs and 79 decreased DMPs compared with the other groups, indicating that 78.9% of DMPs were hypermethylated in genetic SZ risk in the SZ group.

Furthermore, we investigated DMPs among the three genetic risk groups. Compared with low genetic risk in HCs, genetic SZ risk in SZ had 39 genome-wide significant DMPs (online supplemental figure 1 and table 3D, p<9.00×10⁻⁸), while genetic BD risk in SZ had only six genome-wide significant DMPs (online supplemental figure 1 table 3E, p<9.00×10⁻⁸). In contrast, there were no genome-wide significant DMPs between genetic SZ risk in SZ and genetic BD risk in SZ (online supplemental figure 1, p>9.00×10⁻⁸).

Effects of the MRS for current smoking on current smoking status in our study participants
To confirm whether the calculation method of the MRS works well, the MRS at PT≤1 for current smokers based on the EWAS of current smokers was calculated in our participants as a positive control. We confirmed that the MRS for current smoking was significantly associated with current smoking status in our

Figure 2  Epigenome-wide association studies (EWASs) of the risk of SZ between SZ patients and HCs and among four (low genetic risk in HCs, low genetic risk in SZ, genetic BD risk in SZ and genetic SZ risk in SZ) genetic risk groups. Manhattan and volcano plots of DMPs associated with the risk of SZ are indicated. The dotted blue lines in the Manhattan plots indicate a p value of 1.00×10⁻¹²; the dotted red lines in the Manhattan and volcano plots indicate a p value of 9.00×10⁻⁸. Red and blue circles in the volcano plots indicate hypermethylated and hypomethylated DMPs associated with SZ risk, respectively. BD, bipolar disorder; DMPs, differentially methylated positions; HCs, healthy controls; SZ, schizophrenia.
Differences in MRSs for the risk of SZ derived from peripheral blood between case–control status and among the PRS-stratified genetic risk groups

We investigated the effects of MRSs for the risk of SZ derived from peripheral blood at different PT levels on case–control (SZ patients and HCs) and genetic risk (low genetic risk in HCs and genetic BD risk in SZ or low genetic risk in HCs and genetic SZ risk in SZ) statuses (figure 3). The MRSs based on EWAS of SZ risk in peripheral blood were nominally higher in our SZ patients than in HCs (figure 3, a maximum at PT≤1: R$^2=0.066$, p=0.033). In addition, the MRSs for the risk of SZ derived from peripheral blood were nominally higher in genetic SZ risk than in low genetic risk in HCs (figure 3, a maximum at PT<0.05: R$^2=0.080$, p=0.045) and genetic BD risk in SZ (PT<0.05: R$^2=0.106$, p=0.033). In contrast, there were no significant differences in the MRSs between low genetic risk in HCs and genetic BD risk in SZ (figure 3, p>0.05).

Differences in MRSs for the risk of SZ derived from postmortem FC between case–control status and among genetic risk groups stratified by PRSs

To confirm whether the methylation status of SZ in other tissues was also associated with the risk of SZ in whole blood, we next investigated differences in the MRSs based on EWAS of the risk of SZ in the FC among case–control and genetic risk status groups (figure 4). The MRSs related to the risk of SZ derived from the FC were significantly higher in SZ patients than in HCs (figure 4, a maximum at PT≤1: R$^2=0.120$, p=4.57×10$^{-3}$). In addition, the MRSs were significantly higher in genetic SZ risk than in low genetic risk in HCs (figure 4, a maximum at PT≤1: R$^2=0.226$, p=3.26×10$^{-3}$) and genetic BD risk in SZ (PT≤1: R$^2=0.169$, p=0.015), while there were no significant copyright.

Figure 3  Effects of MRSs for the risk of SZ derived from peripheral blood at different P, levels on case–control and PRS-stratified genetic risk statuses. Box plots between case–control and among genetic risk status groups indicate the individual MRSs related to the risk of SZ derived from peripheral blood at PT≤1 in our participants. *p<0.05. BD, bipolar disorder; HCs, healthy controls; MRSs, Methylation Risk Scores; PRS, Polygenic Risk Score; SZ, schizophrenia.

Figure 4  Effects of MRSs for the risk of SZ derived from postmortem frontal cortex on case–control status and PRS-stratified genetic risk statuses. Box plots between case–control and among genetic risk status groups indicate the individual MRSs related to the risk of SZ derived from the frontal cortex at PT≤1 in our participants. *p<0.05, **p<0.01, ***p<0.001. BD, bipolar disorder; HCs, healthy controls; MRSs, Methylation Risk Scores; PRS, Polygenic Risk Score; SZ, schizophrenia.
Differences in MRSs for the risk of SZ derived from postmortem STG between case–control status and among genetic risk groups stratified by PRSs

We further investigated whether the MRSs related to the risk of SZ derived from the STG were associated with case–control and genetic risk statuses (figure 5). The MRSs for the risk of SZ derived from the STG were significantly higher in SZ patients than in HCs (figure 5, a maximum at $P_T<0.1$: $R^2=0.238$, $p=2.37\times10^{-5}$). Furthermore, the MRSs were significantly higher in genetic SZ risk in SZ than in low genetic risk in HCs (figure 5, a maximum at $P_T<0.1$: $R^2=0.284$, $p=1.10\times10^{-5}$) and genetic BD risk in SZ ($P_T<0.1$: $R^2=0.181$, $p=8.22\times10^{-5}$). In addition, the MRSs were nominally higher in genetic BD risk in STG in SZ than in low genetic risk in HCs (figure 5, a maximum at $P_T<0.01$: $R^2=0.185$, $p=8.53\times10^{-5}$).

Conversely, we revealed that the MRSs for the risk of SZ based on our EWAS among genetic risk status in whole blood were nominally higher in SZ patients than in HCs in both the FC (online supplemental figure 3, a maximum at $P_T<0.001$: $R^2=0.089$, $p=0.027$) and STG (online supplemental figure 3, a maximum at $P_T<0.05$: $R^2=0.090$, $p=0.018$). In contrast, the MRSs for the risk of SZ based on our EWASs between case–control status in whole blood were nominally higher in SZ patients than HCs in the FC only (online supplemental figure 3, a maximum at $P_T\leq 1$: $R^2=0.102$, $p=0.023$).

Correlations between the MRSs and PRSs

To explore whether the MRSs and PRSs were independent genetic markers, we investigated the correlations between MRSs for SZ derived from peripheral blood, FC and STG and PRSs for SZ, BD and SZ versus BD in our study participants. There were no significant correlations between the MRSs and PRSs (online supplemental figure 4, $p>0.01$), except for marginal positive correlations of the PRSs for SZ with the MRSs for SZ derived from the FC (Spearman’s $r=0.21$, $p=0.042$) and STG ($r=0.22$, $p=0.028$).

DISCUSSION

This is the first study to investigate the MRSs for the risk of SZ derived from blood and brain tissues to explain the risk of SZ between SZ patients and HCs, particularly genetic SZ risk among the PRS-stratified genetic risk groups. In the preliminary EWASs of SZ risk between the case–control status and among the PRS-stratified genetic risk status, genome-wide significant DMPs related to SZ were higher in the PRS-stratified genetic SZ risk group than in the overall SZ patient and genetic BD risk groups. The MRSs for the risk of SZ derived from peripheral blood, postmortem FC and STG were higher in our SZ patients, particularly in the PRS-stratified genetic SZ risk in SZ, than in the low genetic risk in HCs and genetic BD risk in SZ. In contrast, the MRSs for the risk of SZ based on our preliminary EWASs of SZ risk among the PRS-stratified genetic risk status in peripheral blood were associated with the diagnostic status of SZ in postmortem FC and STG. Furthermore, we indicated that the MRSs and PRSs were distinct genetic markers in our participants. These findings suggest that alterations in genome-wide DNAm in SZ patients beyond tissue differences between the blood and brain would contribute to the pathogenesis of SZ even after stratifying participants into genetic risk groups by PRSs.

To date, large-scale GWASs by PGC have successfully identified multiple loci associated with SZ, BD and SZ versus BD, and these GWAS summary statistics are available for PRS calculations. Methodology for the PRS calculation has been established, while that for the MRS calculation has not been established, and there are several issues related to the MRS calculations that need to be solved. The number of DNAm sites is approximately two times higher with the Illumina Infinium Human Methylation EPIC than with the 450 BeadChips. Although EWASs of SZ risk derived from blood and brain tissues have also been performed, the sample sizes for EWASs were much smaller than those for GWASs. Furthermore, complete summary statistics for most EWASs are not available. In contrast, partial summary statistics above a certain $P_T^*$, for example, $P_T<5.00\times10^{-5}$, or raw (IDAT) files from the methylation array platform are available. Using the partial summary statistics, or the IDAT files, we calculated the MRSs after removing highly correlated DNAm sites in our study sample using the CoMeBack method in a similar way.
as pruning for PRS calculations. We confirmed that the MRS for current smoking was consistent with actual current smoking status, with current smokers having higher scores than current nonsmokers, and revealed that the MRSs for SZ risk calculated using the partial summary statistics or the IDAT files were associated with case-control and genetic risk status outcomes in participants. These findings suggest that the MRSs could be available as one of the genome-wide genetic profiles.

There were no significant correlations between the MRSs and PRSs. Studies to elucidate the pathophysiology of SZ and BD have traditionally focused on genetic factors, including PRSs or environmental factors, while more recent work has highlighted an additional role for epigenetic processes, including DNAm, in mediating susceptibility. DNAm has been identified as a key mechanism for the environmental regulation of gene expression and is essential for human development and the pathogenesis of disorders. We indicated that MRSs based on the accumulation of DNAm patterns were associated with diagnostic status, particularly genetic risk status stratified by PRSs based on the accumulation of SNPs. These findings suggest that the PRSs and MRSs might independently and additively contribute to the risk of disorder via regulation of gene expression and might be independent genetic markers for assessing an individual’s risk of disorder across the entire genome.

Although the degrees of individual DNAm are not necessarily concordant between blood and brain tissues, common differential DNAm sites between peripheral blood and brain tissues have been reported in SZ patients. In EWAS of the risk of SZ between low genetic risk in HCs and genetic SZ risk in SZ, we explored the blood–brain epigenetic concordance of individual genome-wide significant DMPs using the BECon tool (https://redgar598.shinyapps.io/BECon/) (online supplemental figure 5). The interindividual variations between blood and brain tissues at the DMPs were limited. However, it was unclear whether genome-wide DNAm profiles across the genome derived from blood and brain tissues commonly contribute to the risk of SZ. Furthermore, we demonstrated that the risk of SZ and genetic SZ risk in our whole-blood participants were affected by the MRSs for SZ risk derived from the FC and STG as well as whole blood, while the risk of SZ in the FC and STG was also affected by MRSs for genetic SZ risk derived from our whole-blood participants. The variances (Nagelkerke’s pseudo-$R^2$) in the risk of SZ explained by the MRSs derived from whole blood, FC, and STG were different: 6.6% for whole blood, 12.0% for FC, and 23.8% for STG. As shown in online supplemental table 2, the number of DNAm sites for MRS calculations was different among MRSs derived from whole blood, FC and STG depending on the partial and complete summary statistics or the type of methylation array platform. Thus, the different variances would be due to differences in the number of DNAm sites for MRS calculations rather than differences between blood and brain tissues.

Various environmental factors, including lifestyle, such as diet and smoking, and medications, affect DNAm patterns. Our analyses were performed after correcting for current smoking status. However, the MRSs might have been influenced by other demographic clinical variables. Thus, we further investigated partial correlations between MRSs derived from whole blood, FC and STG and our available clinical variables, such as duration of illness, total antipsychotic dosage in chlorpromazine equivalents and body mass index, with age, sex, smoking status and six blood cell counts as covariates in SZ patients (online supplemental table 4). Consistent with a previous study, no significant correlations between the MRSs and clinical variables were observed ($p<0.05$). These findings suggest that the MRS does not simply capture an environmental effect of lifestyle or medications but could be a novel indicator integrating environmental and genetic factors. To capture a combined effect of different environmental exposures including advanced paternal age, obstetric and perinatal complications, winter–birth, childhood adversities (i.e. physical and sexual abuse, neglect and parent death), urban upbringing, and cannabis abuse associated with SZ, aggregate environmental risk scores by summing environmental exposures were generated. The cumulative environmental risk scores were associated with increased risk of SZ. Unfortunately, as we did not have information about each environmental exposure in our participants, we could not calculate the aggregate environmental risk scores. Further study is needed to examine whether the MRSs are correlated with the aggregate environmental risk scores for SZ.

As shown in table 1, there were no significant differences in clinical severity, such as age at onset, psychiatric symptoms and antipsychotic dosage, among genetic risk groups in SZ patients stratified by PRSs. Furthermore, MRSs were not significantly correlated with any clinical severity in SZ patients (online supplemental table 4). Therefore, there might be no strong relationships between these clinical variables, including psychiatric symptoms and medications, and PRSs or MRSs at least within our SZ patients. Since PRSs have been associated with impaired intermediate phenotypes, such as cognitive functions and brain structures, MRSs might be also associated with intermediate phenotypes independent of PRSs.

There are some limitations to the interpretation of our findings. Based on the largest GWASs for SZ, BD and SZ versus BD from the PGC at the time, we extracted individuals with the genetic risk. However, the latest and largest-scale GWAS of BD that identified 298 loci has been recently preprinted. Further stratification using the latest GWAS of BD would be required. We used available partial EWAS summary statistics or raw IDAT files to calculate the MRSs as a discovery EWAS. The sample sizes in the EWAS were much smaller than those in large-scale GWASs for PRS calculation. Generally, since the statistical power of the PRS analysis depends heavily on the sample size in GWASs, further MRS studies using a larger EWAS sample would yield more reliable findings. We cannot directly compare the variance in the SZ risk explained by the MRSs derived from blood and brain tissues because the utilised number of DNAm sites was different among the MRSs due to differences in partial or complete EWAS summary statistics and 450 or EPIC arrays. The MRS derived from complete EWAS summary statistics using the Illumina EPIC array would yield more reliable results.

In conclusion, we suggest that genome-wide DNAm risk scores for the risk of SZ would contribute to the pathogenesis of SZ beyond tissue differences between the blood and brain. Furthermore, the genome-wide DNAm risk scores for the risk of SZ explained the SZ risk among genetic risk groups even after stratifying by PRSs for SZ, BD and SZ versus BD. As we indicated that there were no significant correlations between the MRSSs and PRSs and that the MRS might not be a simple genetic marker capturing an environmental effect of lifestyle or medications, we suggest that MRSs would be genetic markers for assessing an individual’s risk of SZ independent of PRSs. Our findings support a multifactorial liability model for SZ and highlight the important additive effect between epigenetic and genetic risk factors, which may lead to the development of SZ.

Acknowledgements We would like to thank all the individuals who participated in this study. We thank Dr R Nakaki and Dr G Yoshikawa (Rheiika, Tokyo, Japan) for analyzing the DNA methylation data.
Contributors KO supervised the entire project, collected the data, wrote the manuscript and was critically involved in the design, analysis and interpretation of the data. KD, MShimada, DF and TS were responsible for performing the literature review. DN, JH, MSoda, KK and KI were involved in the genotyping. MSoda, DF, KT, AK, YM, SS, KK and TS were heavily involved in the collection of the majority of the data and intellectually contributed to data interpretation. All authors contributed to and approved the final manuscript. KO acts as guarantor.

Funding This work was supported by Grants-in-Aid for Scientific Research (C) (19K08081, 21K07497 and 22K07614) and Young Scientists (B) (116K19784) from the Japan Society for the Promotion of Science (JSPS); the AMED under Grant Numbers JP21ak0105032, JP22k0031112 and JP23dk0107103; a grant from the SENSINH Medical Research Foundation (N/A); a grant from the Uehara Memorial Foundation (N/A); a grant from the Takeda Science Foundation (N/A); a grant from the YOKOYAMA Foundation for Clinical Pharmacology (YRF1807); a grant from the Smoking Research Foundation (N/A); and a grant from the Mochida Memorial Foundation for Medical and Pharmaceutical Research (N/A).

Competing interests None declared.

Patient consent for publication Consent obtained directly from patient(s).

Ethics approval Written informed consent was obtained from all participants after the procedures had been thoroughly explained. This study was performed in accordance with the World Medical Association’s Declaration of Helsinki and was approved by the Research Ethics Committees of Gifu University (2019-23) and Kanazawa Medical University (G113, G174).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. The datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated scientific statement, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iD Kazutaka Oh ̈i http://orcid.org/0000-0001-9577-9640

REFERENCES

39 McKinney BC, McClain L, Hensler CM, et al. Schizophrenia-associated differential DNA methylation in brain is distributed across the genome and Annotated to
Mad11L1, a locus at which DNA methylation and transcription phenotypes share genetic variation with schizophrenia risk. *Transl Psychiatry* 2022;12:340.


