Gene expression-based biological test for major depressive disorder: an advanced study

Shin-ya Watanabe1
Shusuke Numata1
Jun-ichi Iga2
Makoto Kinoshita1
Hidehiro Umehara1
Kazuo Ishii3
Tetsuro Ohmori1

1Department of Psychiatry, Institute of Biomedical Sciences, Tokushima University Graduate School, Tokushima, 2Department of Neuropsychiatry, Molecules and Function, Ehime University Graduate School of Medicine, Ehime, 3Department of Applied Biological Science, Faculty of Agriculture, Tokyo University of Agriculture and Technology, Tokyo, Japan

Correspondence: Shusuke Numata
Department of Psychiatry, Institute of Biomedical Sciences, Tokushima University Graduate School, 3-8-15 Kuramoto-cho, Tokushima 770 8503, Japan
Tel +81 886 33 7130
Fax +81 886 33 7131
Email shu-numata@umin.ac.jp

Purpose: Recently, we could distinguish patients with major depressive disorder (MDD) from nonpsychiatric controls with high accuracy using a panel of five gene expression markers (ARHGAP24, HDACS, PDGF, PRN, and SLC6A4) in leukocyte. In the present study, we examined whether this biological test is able to discriminate patients with MDD from those without MDD, including those with schizophrenia and bipolar disorder.

Patients and methods: We measured messenger ribonucleic acid expression levels of the aforementioned five genes in peripheral leukocytes in 17 patients with schizophrenia and 36 patients with bipolar disorder using quantitative real-time polymerase chain reaction (PCR), and we combined these expression data with our previous expression data of 25 patients with MDD and 25 controls. Subsequently, a linear discriminant function was developed for use in discriminating between patients with MDD and without MDD.

Results: This expression panel was able to segregate patients with MDD from those without MDD with a sensitivity and specificity of 64% and 67.9%, respectively.

Conclusion: Further research to identify MDD-specific markers is needed to improve the performance of this biological test.

Keywords: depressive disorder, biomarker, gene expression, schizophrenia, bipolar disorder

Introduction

Major depressive disorder (MDD) is a highly prevalent psychiatric disorder1 and a leading cause of disease burden.2 Currently, a diagnosis of MDD is made on the basis of clinical symptoms that are exhibited by patients. There are no established practical laboratory tests that can be used to discriminate patients with MDD from those without MDD.

Gene expression profiling of peripheral blood has emerged as a useful method not only to investigate the pathogenesis of MDD, but also to identify relevant biomarkers.3–5 Recently, we examined the diagnostic performance of a multi-assay leukocyte gene expression-based test comprising 40 candidate genes, and developed a panel of 5 gene expression markers (ARHGAP24, HDACS, PDGF, PRN, and SLC6A4), which are able to distinguish patients with MDD from nonpsychiatric control subjects with high accuracy in discriminant analysis.6 However, it is not clear whether this panel of gene expression markers is useful when subjects with other psychiatric disorders, such as those with schizophrenia and bipolar disorder, are included. In some cases, it may be difficult for clinicians to accurately diagnose MDD at the patient’s first visit. Li et al7 have reported that a diagnosis of MDD was changed to bipolar disorder in 7.6%–12% of patients initially diagnosed with MDD, with a mean time to change of 1.89–2.98 years.7 Ruggero et al8 have reported that 16.4% of participants classified as having MDD
with psychosis at baseline were later rediagnosed with schizophrenia or schizoaffective disorder.

In the present study, we measured five gene expression levels of subjects with schizophrenia and bipolar disorder, and we evaluated whether our gene expression-based markers for MDD are able to discriminate patients with MDD from those without MDD (controls, schizophrenia, and bipolar disorder) in the combined dataset of our previous study and the present study.

Materials and methods
Subjects
Seventeen medication-free patients with schizophrenia and 39 patients with bipolar disorder were recruited from the Tokushima University Hospital in Japan. Diagnoses of schizophrenia and bipolar disorder were made by at least two experienced psychiatrists according to DSM-IV criteria on the basis of extensive clinical interviews and a review of medical records. All of the patients with bipolar disorder were treated with lithium, anticonvulsants, atypical antipsychotics, antidepressants, or combinations of the above. None of the patients had any other medical disorder. The patients had not used nonsteroidal anti-inflammatory agents or steroids within at least 2 months before the study initiation. Demographic data for the participants are shown in Table 1. All subjects who participated in this study were of unrelated Japanese origin and signed written informed consent forms, this study was also approved by the Institutional Ethics Committee of the University of Tokushima Graduate School.

Tissue processing, RNA purification, and sample preparation for real-time PCR analysis
PAX gene blood ribonucleic acid (RNA) tubes (Qiagen, Tokyo, Japan) and PAX gene Blood RNA kits (Qiagen) were used according to the manufacturer’s recommendations to extract total RNA from peripheral leukocytes taken from whole blood samples. After assessing RNA quality and quantity, individual total RNA samples (2 μm each), random (N6) primers, and Quantiscript Reverse Transcriptase (Qiagen, Tokyo, Japan) were used to synthesize complementary deoxyribonucleic acids (cDNAs).

Gene expression analysis
Quantitative real-time polymerase chain reaction (RT-PCR) analysis was performed using an ABI 7500 Fast Real-Time PCR System (Applied Biosystems, Foster City, CA, USA) using the TaqMan gene expression probes (Applied Biosystems) Hs00229928_mL (ARHGAP24), Hs00608366_m1 (HDAC5), Hs00219196_m1 (PDGFC), Hs1920617_s1 (PRNP), Hs0984356_m1 (SLC6A4), and Hs99999905_m1 (GAPDH). GAPDH was used as the reference gene, as in our previous study. All measurements were performed in duplicate. The ∆∆Ct method was used to determine relative expression of each gene in the subjects. The obtained value for each gene (Ct) was used to calculate the relative level of gene expression (∆Ct) by normalization to the reference gene, GAPDH. For validation of our expression data, we used 10 samples from the current study and measured GAPDH expression according to the same protocol, and we observed a high correlation between the two datasets (r²=0.986, P=1.11e–08). Expression data for 25 patients with MDD and 25 controls were obtained from our previous study.

Statistical analyses and construction of the discrimination score
Statistical analyses were performed using R ver3.1.1. One-way analysis of variance (ANOVA) was used to test for differences in age among the groups, and Fisher’s exact test was used to examine differences in sex distribution among the groups. The significance of differential expression among groups was examined using Kruskal–Wallis one-way ANOVA and the Bonferroni test was used for post hoc comparisons. A linear discriminant function was developed for discrimination between patients with MDD and without

Table 1 Demographic data of participants

<table>
<thead>
<tr>
<th>Demographic factor</th>
<th>MDD</th>
<th>Schizophrenia</th>
<th>Bipolar disorder</th>
<th>Control</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>25</td>
<td>17</td>
<td>39</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Male, n</td>
<td>7</td>
<td>8</td>
<td>15</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Female, n</td>
<td>18</td>
<td>9</td>
<td>24</td>
<td>16</td>
<td>0.652</td>
</tr>
<tr>
<td>Age</td>
<td>43.00±13.99</td>
<td>31.52±10.18</td>
<td>51.05±11.94</td>
<td>40.40±11.878</td>
<td>1.90e–06</td>
</tr>
<tr>
<td>HAM-D score</td>
<td>22.4±7.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Data are expressed as n or mean (±SD).
Abbreviations: SD, standard deviation; MDD, major depressive disorder; HAM-D, Hamilton Rating Scale for Depression.
MDD, as in our previous study. A discrimination score (D-score) was calculated for each subject by multiplying the coefficients of the linear discriminants, obtained using the lda() function of the MASS package in R, to the standardized values of expression after Z-score transformation of the five selected genes. A D-score was calculated for each subject. The D-score indicates the probability of membership to the group of patients with MDD or the group of subjects without MDD (schizophrenia or bipolar disorder or controls). The highest membership probability for each subject determined their classification into the MDD group or the other groups. A D-score was calculated for each sample as follows:

$$D\text{-score} = -0.73980239 \times PDGFC + 0.84889084 \times PRNP - 0.42559503 \times ARHGAP24 - 0.07806493 \times SLC6A4 + 1.48110014 \times HDAC5 + 0.005.$$

In this plot, as an aid in understanding the discriminant, to set a linear decision boundary at the zero level, the value of the constant term was adjusted. Any subject whose D-score > 0 was identified as “patients with MDD”, and those with a D-score < 0 was identified as “subjects without MDD”.

**Results**

Demographic data of the participants are shown in Table 1. There were no significant differences among the groups in terms of sex distribution, while there were significant differences in age among the groups. We measured mRNA expression levels of five genes (ARHGAP24, HDAC5, PDGFC, PRNP, and SLC6A4) using quantitative RT-PCR. The expression levels of each of these five genes are shown in Table 2. The expression levels of these five genes were not associated with sex or age in multivariable analyses.

All of the aforementioned five genes had similar changes in expression in MDD, schizophrenia, and bipolar disorder compared to controls and significantly changed among the groups. Compared to controls, significantly increased mRNA levels of the ARHGAP24 gene were observed in schizophrenia and bipolar disorder ($P=8.5 \times 10^{-3}$, 0.017, respectively). Compared to controls, significantly increased mRNA levels of the HDAC5 gene were observed in schizophrenia and bipolar disorder ($P=4.9 \times 10^{-3}$, 1.6 $\times 10^{-3}$, respectively). Significantly increased mRNA levels of the PDGFC gene were observed in MDD, schizophrenia, and bipolar disorder compared to controls ($P=0.015$, 5.8 $\times 10^{-3}$, 0.042, respectively). Although mRNA levels of the PRNP gene significantly changed among groups, the significant differences in the comparison between each patient group and control group were not observed. Significantly increased mRNA levels of the SLC6A4 gene were seen in MDD, schizophrenia, and bipolar disorder compared to controls ($P=0.0238$, 4.9 $\times 10^{-3}$, 0.003, respectively).

The D-scores for each subject are graphically depicted in Figure 1. The index in Figure 1 indicates the order of the samples sorted by their D-scores. Of the 25 patients with MDD, 16 had an MDD-positive D-score test. Of the other 81 subjects (17 with schizophrenia, 39 with bipolar disorder, and 25 controls), 52 had an MDD-negative D-score. We were thus able to segregate the patients with MDD from those without MDD with a sensitivity of 64% and specificity of 67.9%. The discriminating capacities of each single gene and a combination of five genes are shown in Tables S1 and S2.

**Discussion**

We recently developed a multiplex expression-based biological test for MDD, which demonstrated good sensitivity, and specificity exceeding 80% in a cohort comprising patients with MDD and nonpsychiatric controls. In the present study, we examined the diagnostic performance of this panel in a cohort, including subjects with other psychiatric disorders, and demonstrated moderate diagnostic performance with a sensitivity and specificity of 64% and 67.9%, respectively, in differentiating patients with MDD from those without MDD.

Multiple biomarkers can serve better than a single marker in reducing the impact of the variation between populations.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>ΔC₅₀ values of five genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene</td>
<td>MDD (n=25)</td>
</tr>
<tr>
<td>PDGFC</td>
<td>8.17±0.59</td>
</tr>
<tr>
<td>PRNP</td>
<td>6.59±1.23</td>
</tr>
<tr>
<td>ARHGAP24</td>
<td>6.47±0.60</td>
</tr>
<tr>
<td>SLC6A4</td>
<td>9.08±1.05</td>
</tr>
<tr>
<td>HDAC5</td>
<td>4.48±0.50</td>
</tr>
</tbody>
</table>

**Note:** Data are expressed as mean (±SD) unless otherwise noted.

**Abbreviations:** BP, bipolar disorder; Con, control; MDD, major depressive disorder; Sc, schizophrenia; SD, standard deviation.
Several blood-based multiplex diagnostic biomarkers for MDD have been reported. Spijker et al. demonstrated that 7 lipopolysaccharide stimulation-based gene expression markers discriminated 16 patients with MDD from 13 controls with a sensitivity and specificity of 87.5% and 61.5%, respectively. This result was replicated in an independent set of samples with a sensitivity and specificity of 76.9% and 71.4%, respectively (13 patients with MDD and 14 controls). Numata et al. demonstrated that the top 18 DNA methylation markers of peripheral leukocytes distinguished 20 patients with MDD from 19 controls. This result was replicated in an independent set of samples with high accuracy (12 patients with MDD and 12 controls).

Zheng et al. demonstrated that 17 plasma nuclear magnetic resonance-based markers distinguished 58 patients with MDD from 42 control subjects. This result was replicated in an independent set of samples with a sensitivity and specificity of 92.8% and 83.3%, respectively (14 patients with MDD and 12 controls). The same group also demonstrated that 17 gas chromatography–mass spectrometry-based markers of peripheral mononuclear cells discriminated 50 subjects with MDD from 50 controls, yielding an area under the curve (AUC) of 0.93. This result was replicated in an independent set of samples with an AUC of 0.87 (58 patients with MDD and 56 controls).

Liu et al. demonstrated that 4 plasma ultra-high-performance liquid chromatography equipped with quadrupole-time-of-flight mass spectrometry-based markers distinguished 60 patients with MDD from 59 control subjects with good sensitivity and specificity. This result was replicated in an independent set of samples with an AUC exceeding 0.8 (75 patients with MDD and 52 controls). Papakostas et al. demonstrated that 9 serum protein-based markers, which were measured using ELISA, distinguished 36 patients with MDD from 43 nondepressed subjects with high accuracy. This result was replicated in an independent set of samples with a sensitivity and specificity of 91.1% and 81%, respectively (34 patients with MDD and 43 controls). Furthermore, the authors later conducted a test using an improved model and confirmed their results in a larger independent sample set with an AUC of 0.96. These results suggest that multiple blood biomarkers can be useful tools for diagnosing MDD. However, previous biological tests for MDD have examined whether biomarkers can distinguish individuals with MDD from nonpsychiatric controls. To our knowledge, this is the first study to evaluate the utility of a biological diagnostic test for MDD in the condition where not only controls, but also subjects with other psychiatric disorders are included in the analysis. As expected, we found that the sensitivity of the test in discriminating patients with MDD from those without MDD, including those with schizophrenia and bipolar disorder, was lower in the present study compared to its sensitivity in discriminating between patients with MDD and controls in our previous study. On the other hand, the best AUCs of MDD, schizophrenia, and bipolar disorder in the combined dataset were 0.87, 0.97, and 0.84, suggesting that this expression panel was not specific to MDD. Further research to identify MDD-specific markers is needed to improve the performance of this biological test.

Our expression panel consisted of five genes, ARHGAP24, HDAC5, PDGFC, PRNP, and SLC6A4. Altered expressions of the ARHGAP24, PDGFC, and PRNP genes in MDD were originally discovered using cDNA microarray in our previous studies. The ARHGAP24 gene encodes a GTPase-activating protein, and dysregulation of this gene inhibits axon and dendrite outgrowth and branching. PDGFC is a member of the platelet-derived growth factor (PDGF) family and is critical for neuronal survival in the central nervous system.

Figure 1 The distribution of D-scores.

Notes: The distribution of D-scores indicated a sensitivity and specificity of 64% and 67.9%, respectively, in differentiating between MDD patients and subjects without MDD. This result was replicated in an independent set of samples with an AUC exceeding 0.8 (75 patients with MDD and 52 controls).

Abbreviations: D-score, discrimination score; MDD, major depressive disorder.

stable and subgroups. Several blood-based multiplex diagnostic biomarkers for MDD have been reported. Spijker et al. demonstrated that 7 lipopolysaccharide stimulation-based gene expression markers discriminated 16 patients with MDD from 13 controls with a sensitivity and specificity of 87.5% and 61.5%, respectively. This result was replicated in an independent set of samples with a sensitivity and specificity of 76.9% and 71.4%, respectively (13 patients with MDD and 14 controls). Numata et al. demonstrated that the top 18 DNA methylation markers of peripheral leukocytes distinguished 20 patients with MDD from 19 controls. This result was replicated in an independent set of samples with high accuracy (12 patients with MDD and 12 controls).

Zheng et al. demonstrated that 17 plasma nuclear magnetic resonance-based markers distinguished 58 patients with MDD from 42 control subjects. This result was replicated in an independent set of samples with a sensitivity and specificity of 92.8% and 83.3%, respectively (14 patients with MDD and 12 controls). The same group also demonstrated that 17 gas chromatography–mass spectrometry-based markers of peripheral mononuclear cells discriminated 50 subjects with MDD from 50 controls, yielding an area under the curve (AUC) of 0.93. This result was replicated in an independent set of samples with an AUC of 0.87 (58 patients with MDD and 56 controls).

Liu et al. demonstrated that 4 plasma ultra-high-performance liquid chromatography equipped with quadrupole-time-of-flight mass spectrometry-based markers distinguished 60 patients with MDD from 59 control subjects with good sensitivity and specificity. This result was replicated in an independent set of samples with an AUC exceeding 0.8 (75 patients with MDD and 52 controls).

Papakostas et al. demonstrated that 9 serum protein-based markers, which were measured using ELISA, distinguished 36 patients with MDD from 43 nondepressed subjects with high accuracy. This result was replicated in an independent set of samples with a sensitivity and specificity of 91.1% and 81%, respectively (34 patients with MDD and 43 controls). Furthermore, the authors later conducted a test using an improved model and confirmed their results in a larger independent sample set with an AUC of 0.96. These results suggest that multiple blood biomarkers can be useful tools for diagnosing MDD. However, previous biological tests for MDD have examined whether biomarkers can distinguish individuals with MDD from nonpsychiatric controls. To our knowledge, this is the first study to evaluate the utility of a biological diagnostic test for MDD in the condition where not only controls, but also subjects with other psychiatric disorders are included in the analysis. As expected, we found that the sensitivity of the test in discriminating patients with MDD from those without MDD, including those with schizophrenia and bipolar disorder, was lower in the present study compared to its sensitivity in discriminating between patients with MDD and controls in our previous study. On the other hand, the best AUCs of MDD, schizophrenia, and bipolar disorder in the combined dataset were 0.87, 0.97, and 0.84, suggesting that this expression panel was not specific to MDD. Further research to identify MDD-specific markers is needed to improve the performance of this biological test.

Our expression panel consisted of five genes, ARHGAP24, HDAC5, PDGFC, PRNP, and SLC6A4. Altered expressions of the ARHGAP24, PDGFC, and PRNP genes in MDD were originally discovered using cDNA microarray in our previous studies. The ARHGAP24 gene encodes a GTPase-activating protein, and dysregulation of this gene inhibits axon and dendrite outgrowth and branching. PDGFC is a member of the platelet-derived growth factor (PDGF) family and is critical for neuronal survival in the central nervous system.
is involved in disrupted-in-schizophrenia-1 expression,\(^2^1\) gamma aminobutyric acid dysfunction,\(^2^2\) and white matter integrity in bipolar disorder.\(^2^3\) The PRNP gene encodes the prion protein, which has a protective role in several neurological conditions and affects N-methyl-D-aspartate receptor function in hippocampal neurons.\(^2^4\) Mice with disruptions of this gene show depressive-like behavior.\(^2^5\) In addition, cellular prion protein-positive glial cells have been shown to be significantly reduced in the cingulate gyrus of patients with bipolar disorder or MDD.\(^2^6\) HDAC5 is a member of the histone deacetylase family, which remove acetyl groups from histones to produce a less-accessible chromatin structure. Chronic administration of imipramine reversed downregulation of brain-derived neurotrophic factor (BDNF) expression and increased histone acetylation at the BDNF gene promoters in social defeat stress mice, which was associated with a selective downregulation of HDAC5 expression.\(^2^7\) Previous reports have indicated increased HDAC5 expression in leukocytes of patients with MDD.\(^2^8,2^9\) Furthermore, HDAC5 expression was also associated with the treatment efficacy or clinical improvement in MDD.\(^3^0\) The SLC6A4 gene is one of the most investigated genes in relation to MDD.\(^3^1,3^2\) Several studies have consistently demonstrated increased expressions of the SLC6A4 gene in leukocytes of MDD\(^3^3-3^5\) and schizophrenia.\(^3^6\) In PET imaging studies, SLC6A4-binding potential has been shown to be elevated in subjects with MDD who have more severe symptoms.\(^3^7\)

There are several limitations to the present study. First, our sample size was not large, and we did not confirm our results in an independent cohort. Second, the patients with bipolar disorder were receiving various medications. In addition, we did not evaluate the mood status of these patients with bipolar disorder as well as that of the patients with schizophrenia. Further studies will be needed in consideration of medications and mood status. Third, we did not take physical conditions, such as body mass index and smoking status, into consideration, due to lack of information. Fourth, we combined our previous expression data of subjects with MDD and controls with the present expression data of schizophrenia and bipolar disorder although we conducted experiments according to the same protocol. Fifth, all subjects who participated in this study were of Japanese origin. Further studies will be needed in other ethnic cohorts. Finally, we measured relative mRNA expression levels. An absolute quantitation across laboratories will be needed to develop a common and reliable biological test.\(^3^8\)

**Conclusion**

We were able to differentiate patients with MDD from subjects without MDD, including those with schizophrenia and bipolar disorder, with moderate sensitivity and specificity using gene expression changes in our panel. As we developed this panel based on MDD candidate genes, further research to identify MDD-specific markers will be needed to improve the performance of this biological test.

**Acknowledgments**

This work was supported in part by Japan Agency for Medical Research and development, AMED (TO), Grant-in-Aid for Scientific Research (C) (No 15K098090) (SN), Grant-in-Aid for Scientific Research (C) (No 15K09808) (J-II), Grant-in-Aid for Scientific Research (C) (No 26330325) (KI), and Grant-in-Aid for Young Scientists (B) (No 16K19769) (S-yW).

The authors would like to thank all the volunteers who understood the study purpose and participated in this study, and the physicians who helped them to collect clinical data and blood samples. The authors would also like to thank Mrs Akemi Okada for her technical assistance.

**Author contributions**

SN and TO designed the study and acquired the data. S-yW acquired the data, KI, MK, HU, and S-yW analyzed the data. J-II, SN, and S-yW wrote the draft of this paper. All authors contributed toward data analysis, drafting and revising the paper and agree to be accountable for all aspects of the work.

**Disclosure**

The authors report no conflicts of interest in this work.

**References**


# Supplementary materials

## Table S1 The discriminating capacities of each single gene

<table>
<thead>
<tr>
<th>Gene</th>
<th>AUC</th>
<th>95% CI</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDGFC</td>
<td>0.52</td>
<td>0.401</td>
<td>0.633</td>
<td>68</td>
</tr>
<tr>
<td>PRNP</td>
<td>0.56</td>
<td>0.431</td>
<td>0.688</td>
<td>52</td>
</tr>
<tr>
<td>ARHGA24</td>
<td>0.51</td>
<td>0.377</td>
<td>0.633</td>
<td>48</td>
</tr>
<tr>
<td>SLC6A4</td>
<td>0.51</td>
<td>0.383</td>
<td>0.626</td>
<td>56</td>
</tr>
<tr>
<td>HDAC5</td>
<td>0.61</td>
<td>0.504</td>
<td>0.722</td>
<td>88</td>
</tr>
</tbody>
</table>

**Abbreviations:** AUC, area under the curve; CI, confidence interval; Min, minimum; Max, maximum.

## Table S2 The discriminating capacities of each combination of five genes

<table>
<thead>
<tr>
<th>Gene number</th>
<th>Candidate gene</th>
<th>Lowest Wilks’ Lambda</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>PRNP, HDAC5</td>
<td>0.9453054</td>
<td>56</td>
<td>65.4</td>
</tr>
<tr>
<td>3</td>
<td>PDGFC, PRNP, HDAC5</td>
<td>0.900726</td>
<td>60</td>
<td>71.8</td>
</tr>
<tr>
<td>4</td>
<td>PDGFC, PRNP, ARGAP2, HDAC5</td>
<td>0.8887443</td>
<td>64</td>
<td>64.1</td>
</tr>
<tr>
<td>5</td>
<td>PDGFC, PRNP, ARGAP2, SLC6A4, HDAC5</td>
<td>0.8884201</td>
<td>64</td>
<td>67.9</td>
</tr>
</tbody>
</table>