




OPEN

## Distinct epigenetic signatures between adult-onset and late-onset depression

Hirota Yamagata<sup>1,2</sup>, Hiroyuki Ogihara<sup>3</sup>, Koji Matsuo<sup>1,7</sup>, Shusaku Uchida<sup>1,2,8</sup>, Ayumi Kobayashi<sup>1</sup>, Tomoe Seki<sup>1,2</sup>, Masaaki Kobayashi<sup>1</sup>, Kenichiro Harada<sup>1</sup>, Chong Chen<sup>1</sup>, Shigeo Miyata<sup>4,5</sup>, Masato Fukuda<sup>5</sup>, Masahiko Mikuni<sup>5,6,9</sup>, Yoshihiko Hamamoto<sup>3</sup>, Yoshifumi Watanabe<sup>1,10</sup> & Shin Nakagawa<sup>1</sup>

The heterogeneity of major depressive disorder (MDD) is attributed to the fact that diagnostic criteria (e.g., DSM-5) are only based on clinical symptoms. The discovery of blood biomarkers has the potential to change the diagnosis of MDD. The purpose of this study was to identify blood biomarkers of DNA methylation by strategically subtyping patients with MDD by onset age. We analyzed genome-wide DNA methylation of patients with adult-onset depression (AOD; age  $\geq 50$  years, age at depression onset  $< 50$  years;  $N = 10$ ) and late-onset depression (LOD; age  $\geq 50$  years, age at depression onset  $\geq 50$  years;  $N = 25$ ) in comparison to that of 30 healthy subjects. The methylation profile of the AOD group was not only different from that of the LOD group but also more homogenous. Six identified methylation CpG sites were validated by pyrosequencing and amplicon bisulfite sequencing as potential markers for AOD in a second set of independent patients with AOD and healthy control subjects ( $N = 11$ ). The combination of three specific methylation markers achieved the highest accuracy (sensitivity, 64%; specificity, 91%; accuracy, 77%). Taken together, our findings suggest that DNA methylation markers are more suitable for AOD than for LOD patients.

Depressive disorders are among the most common psychiatric diseases. However, it is often difficult to discriminate major depressive disorder (MDD) from other depressive disorders based on operational diagnostic criteria (e.g., Diagnostic And Statistical Manual of Mental Disorders, Fifth Edition [DSM-5]) because these criteria mainly rely on patients' symptoms and do not include laboratory findings. Moreover, MDD is probably a heterogeneous entity because the pathophysiology of depression includes multiple factors such as genetics and environment<sup>1</sup>. The onset age is a key factor to heterogeneity in various diseases, for example, diabetes mellitus<sup>2</sup> and asthma<sup>3</sup>. Systematic reviews reported that patients with late-onset depression (LOD) have a lower frequency of family history<sup>4</sup>, decrease in cognitive executive function<sup>5</sup>, and more frequent and pronounced white matter hyperintensities in magnetic resonance imaging<sup>6</sup> compared to patients with early-onset depression, suggesting different pathophysiologies. Recently, we discovered a validated gene expression marker for MDD in patients aged 50 years or older with confirmed MDD diagnosis, without a manic episode, and onset before the age of

<sup>1</sup>Division of Neuropsychiatry, Department of Neuroscience, Yamaguchi University Graduate School of Medicine, 1-1-1 Minami-kogushi, Ube, Yamaguchi 755-8505, Japan. <sup>2</sup>Core Research for Evolutional Science and Technology (CREST), Japan Science and Technology Agency (JST), 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan. <sup>3</sup>Division of Electrical, Electronic and Information Engineering, Graduate School of Sciences and Technology for Innovation, Yamaguchi University, 2-16-1 Tokiwadai, Ube, Yamaguchi 755-8611, Japan. <sup>4</sup>Department of Genetic and Behavioral Neuroscience, Gunma University Graduate School of Medicine, 3-39-22 Showa-machi, Maebashi, Gunma 371-8511, Japan. <sup>5</sup>Department of Psychiatry and Neuroscience, Gunma University Graduate School of Medicine, 3-39-22 Showa-machi, Maebashi, Gunma 371-8511, Japan. <sup>6</sup>Hakodate Watanabe Hospital, 1-31-1 Yunokawacho, Hakodate, Hokkaido 042-8678, Japan. <sup>7</sup>Present address: Department of Psychiatry, Faculty of Medicine, Saitama Medical University, 38 Morohongo, Moroyama, Iruma, Saitama 350-0495, Japan. <sup>8</sup>Present address: SK Project, Medical Innovation Center, Kyoto University Graduate School of Medicine, 53 Shogoin-Kawahara-cho, Sakyo-ku, Kyoto 606-8507, Japan. <sup>9</sup>Present address: Department of Psychiatry, Hokkaido University, Graduate School of Medicine, Kita 15, Nishi 7, Kita-ku, Sapporo 060-8638, Japan. <sup>10</sup>Present address: Southern TOHOKU Research Institute for Neuroscience, Southern TOHOKU General Hospital, 7-115 Yatsuyamada, Koriyama, Fukushima 963-8052, Japan. ✉email: gata@yamaguchi-u.ac.jp

	HC	AOD	LOD
<b>(a)</b>			
No. of subjects	30	10	25
Sex (Male/Female)	13/17	5/5	9/16
Age (years)	63.3 ± 8.0	56.0 ± 6.8*	61.8 ± 7.5
Age at onset (years)	–	35.7 ± 8.9	57.8 ± 6.7 <sup>#</sup>
Total duration of all MDD episodes (months)	–	7.0 ± 6.4	13.2 ± 16.4
Number of episodes	–	3.0 ± 1.7	2.1 ± 2.1
Number of WBCs	4928 ± 1430	4757 ± 1495	5109 ± 1773
Neutrophils (%)	58.1 ± 8.7	58.4 ± 9.1	63.9 ± 6.6*
Lymphocytes (%)	33.2 ± 8.1	33.1 ± 7.5	28.3 ± 6.4
Eosinophils (%)	3.0 ± 1.7	3.0 ± 2.1	2.4 ± 2.0
Basophils (%)	0.45 ± 0.25	0.51 ± 0.29	0.55 ± 0.53
Monocytes (%)	5.2 ± 1.3	5.0 ± 1.5	4.8 ± 1.0
SIGH-D	0.7 ± 0.9	21.7 ± 2.2*	22.4 ± 4.3*
Equivalent dose of imipramine (mg)	–	221 ± 131	183 ± 147
<b>(b)</b>			
No. of subjects	10	10	10
Sex (Male/Female)	5/5	5/5	5/5
Age (years)	59.7 ± 6.2	56.0 ± 6.8	59.2 ± 4.6
Age at onset (years)	–	35.7 ± 8.9	55.5 ± 4.5 <sup>#</sup>
Total duration of all MDD episodes (months)	–	7.0 ± 6.4	21.8 ± 22.5
Number of episodes	–	3.0 ± 1.7	2.3 ± 2.5
Number of WBCs	5171 ± 1180	4757 ± 1495	5333 ± 1667
Neutrophils (%)	59.9 ± 9.0	58.4 ± 9.1	60.5 ± 6.3
Lymphocytes (%)	32.3 ± 8.8	33.1 ± 7.5	32.3 ± 5.7
Eosinophils (%)	2.4 ± 1.7	3.0 ± 2.1	1.9 ± 2.0
Basophils (%)	0.48 ± 0.28	0.51 ± 0.29	0.42 ± 0.30
Monocytes (%)	4.9 ± 1.5	5.0 ± 1.5	4.9 ± 1.1
SIGH-D	0.7 ± 0.7	21.7 ± 2.2*	23.0 ± 4.7*
Equivalent dose of imipramine (mg)	–	221 ± 131	218 ± 198

**Table 1.** Participants' demographics and characteristics in Group 1a and 1b (discovery and training samples, respectively). Characteristics of the participants for the genome-wide methylation array (a) and for the validation using pyrosequencing and amplicon bisulfite sequencing (b). Statistical analyses performed using the EZR software (version 1.37) (<http://www.jichi.ac.jp/saitama-sct/SaitamaHP.files/statmedEN.html>). Data are shown as the mean ± standard deviation. \* $p < 0.05$  compared to HC. <sup>#</sup> $p < 0.05$  compared to AOD. AOD, adult-onset depression; HC, healthy controls; LOD, late-onset depression; MDD, major depressive disorder; SIGH-D, Structured Interview Guide of the Hamilton Depression Rating Scale score; WBCs, white blood cells.

50 years<sup>7</sup>. Other reports also suggest heterogeneity of MDD because leukocytic gene-expression profiles differ between adult-onset depression (AOD) and LOD<sup>7,8</sup>.

DNA methylation is one of the most common epigenetic modifications, controlling gene expression<sup>9,10</sup>. DNA methylation is known to control the resilience to early life stress and depression-like behavior in mice<sup>11</sup>. In humans, several reports show changed DNA methylation in the peripheral blood of patients with depression<sup>12,13</sup>. As an indicative result, we reported alterations in mRNAs of DNA methyltransferases<sup>14</sup>. However, few studies have focused on stratifying MDD by DNA methylation.

We postulate that epigenetic factors have a more significant influence on AOD than on LOD. The aim of this study was the analysis of the differences in methylation profiles of patients with AOD and LOD. We also identified DNA methylation markers from probes of patients with AOD.

## Results

The demographic and clinical data of Group 1a (discovery samples) are shown in Table 1a. Participants with AOD were significantly younger than those in the healthy control (HC) group. The mean percentage of neutrophils was higher in the LOD than the HC group.

We identified specific methylation sites associated with AOD and LOD compared to HC from 411,383 probes using the Fisher ratio (F-ratio), which compares the variance between classes relative to the variance within classes. An increase in the F-ratio generally suggests that a feature distinguishes better between compared classes. The comparison between the top-ranked probes of the AOD and LOD groups and the HC group are shown in Table 2 and Supplementary Table S1. The F-ratios of the top-ranked probes distinguishing the AOD from the HC group were higher than those discriminating between the LOD and HC groups. Similarly, the F-ratios of the

ID	Fisher Ratio	HC average	AOD average	CHR	MAPINFO	UCSC_REFGENE_NAME	UCSC_REFGENE_GROUP	RELATION_TO_UCSC_CPG_ISLAND
<b>(a)</b>								
cg15794987	3.85	0.954	0.937	8	6,534,897			
cg07763047	3.71	0.046	0.025	5	160974837	GABRB2;GABRB2	5'UTR;5'UTR	Island
cg21347377	3.59	0.014	0.006	10	96,305,691	HELLS	1stExon	Island
cg17458347	3.53	0.099	0.083	1	42,928,946			Island
cg14567489	3.41	0.021	0.014	20	43,977,078	SDC4	TSS200	Island
cg02221310	3.33	0.013	0.004	13	41,363,578	SLC25A15;SLC25A15	5'UTR;1stExon	Island
cg16579770	3.29	0.934	0.917	16	72,058,938	DHODH	3'UTR	
cg01320648	3.22	0.548	0.517	4	123091375	KIAA1109	TSS1500	
cg05974695	3.19	0.026	0.017	11	130786536	SNX19	TSS200	Island
cg15525503	3.13	0.020	0.008	1	3,773,166	DFFB;KIAA0562	TSS1500;5'UTR	N_Shore
cg26655295	3.08	0.045	0.034	2	677,162	TMEM18	Body	Island
cg12867237	3.07	0.146	0.128	11	44,087,540	ACCS;ACCS	TSS200;TSS1500	Island
cg16234726	3.06	0.937	0.922	14	102099176			N_Shore
cg20452738	3.04	0.922	0.947	1	226828650	ITPKB	Body	N_Shore
cg23543123	2.99	0.023	0.013	4	48,485,569	SLC10A4;SLC10A4	5'UTR;1stExon	Island
cg22115465	2.91	0.983	0.979	7	2,170,343	MAD1L1;MAD1L1;MAD1L1	Body;Body;Body	
cg04147027	2.90	0.030	0.022	6	82,462,442	FAM46A	TSS200	Island
cg15971980	2.89	0.571	0.676	6	150254442			
cg24934561	2.89	0.064	0.055	1	3,663,902	KIAA0495	TSS200	Island
cg23531049	2.81	0.845	0.892	15	42,075,277	MAPKBP1;MAPKBP1	Body;Body	
ID	Fisher Ratio	HC average	LOD average	CHR	MAPINFO	UCSC_REFGENE_NAME	UCSC_REFGENE_GROUP	RELATION_TO_UCSC_CPG_ISLAND
<b>(b)</b>								
cg06557376	3.14	0.959	0.949	17	8,384,607	MYH10	Body	
cg23854456	2.87	0.936	0.920	19	39,806,075	LRFN1	TSS200	S_Shore
cg09597022	2.46	0.884	0.855	6	33,288,332	DAXX;DAXX;DAXX;DAXX	Body;Body;Body;Body	N_Shore
cg04521026	2.21	0.958	0.948	2	240067196	HDAC4	Body	
cg15390741	2.06	0.986	0.976	1	51,822,535	EPS15;EPS15	Body;Body	
cg06034437	2.03	0.521	0.474	3	10,183,196	VHL;VHL	TSS200;TSS200	N_Shore
cg25996553	1.99	0.889	0.869	8	13,133,084	DLC1	Body	N_Shore
cg18394552	1.92	0.648	0.577	5	159428643			
cg04262179	1.89	0.975	0.971	1	178482030	C1orf49;C1orf49;C1orf49;C1orf49	TSS200;TSS200;TSS200;TSS200	
cg03473408	1.88	0.769	0.696	6	30,588,689	MRPS18B	Body	S_Shelf
cg27318635	1.83	0.949	0.937	16	3,781,749	CREBBP;CREBBP	Body;Body	S_Shore
cg16011480	1.78	0.151	0.139	10	97,667,594	C10orf131	TSS200	Island
cg27088822	1.75	0.799	0.752	12	121731226	CAMKK2;CAMKK2;CAMKK2;CAMKK2;CAMKK2;CAMKK2	5'UTR;5'UTR;5'UTR;5'UTR;5'UTR;5'UTR	N_Shelf
cg13081381	1.72	0.480	0.502	13	82,264,192			
cg01345727	1.71	0.759	0.695	13	52,734,018	NEK3;NEK3;NEK3;NEK3	TSS1500;TSS200;TSS200;TSS200	S_Shore
cg11855759	1.70	0.205	0.234	17	9,548,802	USP43	TSS200	Island
cg22389270	1.70	0.263	0.244	8	101162769	POLR2K	TSS200	
cg11558107	1.69	0.932	0.912	15	60,700,124			
cg20420249	1.66	0.808	0.758	1	172378938	DNM3;DNM3	3'UTR;3'UTR	
cg15587947	1.66	0.861	0.844	19	10,220,948	PPAN-P2RY11;P2RY11;PPAN	Body;TSS1500;Body	Island

**Table 2.** The 20 top-ranked DNA methylation sites in the AOD versus HC group (a) and the LOD versus HC group (b). AOD, adult-onset depression; HC, healthy controls; LOD, late-onset depression.

top-ranked RefSeq genes differentiating the AOD from the HC group were also higher than those distinguishing the LOD from the HC group (Supplementary Table S2).

To validate the methylation levels of these probe sites determined by the genome-wide DNA methylation array, we analyzed some identified sites of Group 1b participants using pyrosequencing. The demographic and clinical data of Group 1b are shown in Table 1b. The sensitivity to detect methylation by pyrosequencing is

Probe ID	Cut-off points	Performance for training samples			Performance for test samples		
		Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy
cg07584066	0.308 >	0.70	0.70	0.70	0.82	0.73	0.77
cg07763047	0.018 >	1.00	0.40	0.70	0.91	0.46	0.68
cg10294474	0.789 <	1.00	0.60	0.80	0.55	0.36	0.46
cg15971980	0.691 <	0.90	0.80	0.85	0.36	0.46	0.41
cg16579770	0.889 >	0.80	0.70	0.75	0.73	0.64	0.68
cg20903900	0.828 >	1.00	0.50	0.75	0.55	0.73	0.64

**Table 3.** Sensitivity, specificity, and accuracy in the independent second set (Group 2) for discrimination between AOD and HC. The optimal cut-off points of Group 1b (training samples) were determined by maximizing the Youden index. Sensitivity, specificity, and discrimination accuracy for Group 2 (independent test samples) were calculated. AOD, adult-onset depression; HC, healthy controls.

	HC	AOD
No. of subjects	11	11
Sex (Male/Female)	5/6	5/6
Age (years)	53.8 ± 9.1	53.5 ± 7.6
Age at onset (years)	–	30.3 ± 9.6
SIGH-D	1.8 ± 1.9	21.3 ± 7.6*
Equivalent dose of imipramine (mg)	–	155 ± 91

**Table 4.** Participants demographics and characteristics in Group 2 (independent test samples). Data are shown as the mean ± standard deviation. \* $p < 0.05$  compared to HC. AOD, adult-onset depression; HC, healthy controls; SIGH-D, Structured Interview Guide of the Hamilton Depression Rating Scale score.

about 5%<sup>15</sup>. Therefore, we chose the four probes cg15971980, cg07584066, cg20903900, and cg10294474 whose averaged methylation levels in the AOD group were significantly elevated by at least 5% compared to those in the HC group. The methylation levels of these sites determined by pyrosequencing significantly correlated with those determined by the methylation beads array. The correlation coefficients between the methylation levels of beads array and pyrosequencing,  $p$ -values, and scatter plots are shown in Supplementary Table S3a and Fig. S1, respectively. These results indicated that the findings of the methylation beads array were valid.

However, many of the target methylation sites with a high discrimination rate between AOD and HC presented only small differences of less than 5% in the methylation beads array analysis (Table 2). To validate small differences in methylation levels, we analyzed them additionally by amplicon bisulfite sequencing. We checked the six identified probes cg15794987, cg07763047, cg21347377, cg17458347, cg14567489, and cg16579770. The site cg15971980 was used as the positive control of the amplicon bisulfite sequencing procedure. The methylation levels in the amplicon bisulfite sequencing of the two sites cg07763047 and cg16579770 were significantly correlated with those in the methylation beads array (Supplementary Table S3b). Scatter plots demonstrate the relationships between clinical parameters and each methylation level (Supplementary Figs. S2–7). The methylation levels of the cg07763047 and cg10294474 sites correlated with the age at onset (Supplementary Table S4).

We focused on discriminating between the AOD and HC groups, while the methylation levels from Group 1b were used as the training sample. Each receiver operating characteristic (ROC) curve of the six cytosine methylation sites cg15971980, cg07584066, cg20903900, cg10294474, cg07763047, and cg16579770 in Group 1b was obtained by calculating their sensitivity and specificity. All values for cut-off point, sensitivity, specificity, and discrimination accuracy are displayed in Table 3. All six cytosine methylation sites demonstrated a high discrimination accuracy ( $\geq 70\%$ ).

Next, we employed a different, independent data set from Group 2 (test sample) to validate the above established cut-off points in discriminating between AOD and HC participants. The demographic and clinical data of Group 2 are shown in Table 4, whereas sensitivity, specificity, and discrimination accuracy are presented in Table 3. The methylation of cg07584066, cg16579770, and cg07763047 demonstrated also in Group 2 a relatively high accuracy ( $\geq 68\%$ ).

Next, we assessed the discrimination performance in Group 2 using combinations of two cytosine methylation sites. We constructed the Bayes classifier using the training samples from Group 1b; the independent test samples from Group 2 were diagnosed using this Bayes classifier. The values for sensitivity, specificity, and discrimination accuracy of these combinations are shown in Table 5. The combination of cg16579770 and cg07763047 demonstrated with 73% a high accuracy. We further assessed the discrimination performance using combinations of three cytosine methylations (Table 5). The group consisting of cg07584066, cg16579770, and cg07763047 exhibited the highest accuracy (sensitivity, 64%; specificity, 91%; accuracy, 77%). Furthermore, we visualized the discrimination results for the test samples with the selected probe set, i.e., the target marker set (Supplementary Fig. S8a, b). The axes  $d_1(x)$  and  $d_2(x)$ , which are described in detail in the “Statistical analysis” paragraph of our Methods section, are the distances between a test sample  $x$  from Group 2 and the distribution of

	Sensitivity of test samples	Specificity of test samples	Accuracy of test samples
cg16579770, cg07763047	0.64	0.82	0.73
cg16579770, cg20903900	0.45	0.73	0.59
cg16579770, cg10294474	0.64	0.45	0.55
cg07763047, cg20903900	0.27	0.82	0.55
cg15971980, cg07584066	0.27	0.82	0.55
cg20903900, cg07584066	0.18	0.91	0.55
cg16579770, cg07584066, cg07763047	0.64	0.91	0.77
cg16579770, cg15971980, cg07584066	0.64	0.73	0.68
cg16579770, cg07763047, cg20903900	0.36	0.91	0.64
cg16579770, cg10294474, cg07584066	0.45	0.82	0.64
cg16579770, cg20903900, cg07584066	0.55	0.73	0.64

**Table 5.** Sensitivity, specificity, and accuracy using combinations of two or three DNA methylation sites for discrimination between AOD and HC. AOD, adult-onset depression; HC, healthy controls.

the HC and AOD samples from Group 1b, respectively. Our data show that the triple combination of cg07584066, cg16579770, and cg07763047 discriminates better than the dual combination of cg16579770 and cg07763047.

To confirm that the spread of the AOD distribution is smaller than that of the LOD distribution, the values of the determinant were calculated for both classes using cg16579770 and cg07763047 as specific AOD markers. The value of the determinant signifies the spread of the corresponding distribution. The ratio of AOD to HC was 0.066, whereas the ratio of LOD to HC was 3.187 (both  $p < 0.001$ ; Supplementary Fig. S9). These results suggest that AOD was more homogeneous than HC, but LOD was less homogenous than HC.

## Discussion

Under the assumption that AOD is more strongly influenced by DNA methylation than LOD, we searched for methylation markers in AOD patients using a genome-wide DNA methylation array. As expected, the methylation levels of specific sites discriminated better between AOD and HC than between LOD and HC. Besides, we showed that the spread of the distribution in AOD is smaller than that in LOD.

The identification of specific DNA methylations may facilitate subtyping mental disorders. For example, several DNA methylation sites have been implicated in an increased risk of suicidal behavior in patients with bipolar disorder<sup>16</sup>. The DNA methylation of *FKBP5* was reported to be a potential indicator for the responsiveness to treatment by mindfulness-based stress reduction in post-traumatic stress disorder<sup>17</sup>. To our knowledge, the current study is the first demonstrating that the methylation profiles differ between AOD and LOD. Our results suggest that the vulnerability for AOD may be attributed to DNA methylation. LOD is associated with other illnesses, such as menopausal syndrome or cerebrovascular diseases rather than genetic risk<sup>18</sup>. Thus, DNA methylation markers may be suitable to diagnose AOD rather than LOD.

Several studies have examined genome-wide DNA methylation in the peripheral blood<sup>19–28</sup>. However, most of these studies were designed for symptomatic depression, did not validate their methylation markers by other methods such as pyrosequencing, or did not assess the accuracy using independent subjects. We included only patients who were diagnosed with MDD by psychiatrists, and we re-assessed six identified AOD methylation markers by pyrosequencing or amplicon bisulfite sequencing. Two methylated cytosines were validated by amplicon bisulfite sequencing, although the difference was small (< 5%) compared to HC. Furthermore, we validated the accuracy of these markers using independent subjects. In particular, the combination of cg07584066, cg16579770, and cg07763047 showed the highest accuracy. The locus cg07763047 is annotated to the gamma-aminobutyric acid (GABA) type A receptor  $\beta 2$  subunit (*GABRB2*) gene, cg16579770 to dihydroorotate dehydrogenase (*DHODH*), and cg07584066 to DEAH-box helicase 40 (*DHX40*).

*GABRB2* has been associated with schizophrenia<sup>29,30</sup> and bipolar disorder<sup>31</sup>. One study in postmortem brains of elderly patients with depression reported that the *GABRB2* expression is significantly decreased in the anterior cingulate cortex in both MDD and bipolar disorder<sup>32</sup>. The GABA and glutamate neurotransmitter systems are also implicated in MDD and suicidal behavior<sup>33</sup>. *GABRB2* transcription is regulated by epigenetic mechanisms including methylation<sup>32,34</sup>. Our findings support a role for *GABRB2* in AOD.

The enzyme *DHODH* is involved in pyrimidine biosynthesis<sup>35</sup>. Pyrimidine biosynthesis pathways are connected to cell proliferation and metabolism, therefore *DHODH* is a focus for the development of new drugs against neoplastic or immunological disorders<sup>36</sup>. Until now, *DHODH* has not been known to be directly associated with MDD. However, *DHODH* and its pathways play a fundamental role in cellular homeostasis. For example, *DHODH* is essential for T cell proliferation, and inhibition of *DHODH* is beneficial in autoimmune diseases<sup>37</sup>. It has been suggested that inflammation contributes to the pathophysiology of depression<sup>38</sup>. A recent genome-wide methylation study showed that DNA methylation is altered at multiple immune-related loci in individuals with a history of depression<sup>19</sup>. Thus, *DHODH* may be associated with depression via neuroinflammation.

The protein *DHX40* is an ATP-dependent DEX/H RNA helicase that exerts essential roles in RNA metabolism<sup>39</sup>. Unfortunately, these roles have not been well established yet; however, DEX/H RNA helicases are known to regulate ribosome biogenesis<sup>40</sup>. A recent blood transcriptome study reports that ribosomal genes are upregulated in stress<sup>41</sup>. Thus, ribosomal biogenesis may contribute to AOD pathophysiology.

A recent meta-analysis of epigenome-wide studies in middle-aged and elderly subjects associated the methylation of three CpG sites with depressive symptoms<sup>12</sup>. Unfortunately, we did not replicate these findings. In our study, we diagnosed MDD strictly and selected for the epigenome-wide analysis only MDD patients with an age of 50 years or older. By contrast, the meta-analysis consisted of nine population-based cohort studies including a cardiovascular study, most of the participants were not diagnosed with MDD, and their depressive symptoms were only assessed by self-reported questionnaires.

Several limitations of this study should be mentioned. First, the sample size was small compared to the number of analyzed CpG sites. It is likely that some false positive markers were detected or some false negative markers remained undetected. Second, we only examined methylation markers using leukocytic DNA based on future applicability for clinical use. The functions in the brain of the annotated genes and their methylation remain unclear. Third, most patients received antidepressants. We cannot exclude the possibility that antidepressants might have altered the methylation levels of candidate CpG sites. Fourth, clinical parameters might have influenced the methylation levels. A few factors were indicated to influence the AOD and HC classification by scatter plots representing the relationships between age, age of onset, number of depressive episodes, Structured Interview Guide of the Hamilton Depression Rating Scale (SIGH-D) score, white blood cell (WBC) count, ratio of neutrophils, and methylation levels. Specifically, the methylation levels of the cg07763047 and cg10294474 sites correlated with age at onset, suggesting that the detected markers were onset-age-dependent. However, other clinical factors might have influenced the methylation status. Finally, we did not validate that the methylation markers can distinguish AOD from other psychiatric diseases including schizophrenia and bipolar disorder.

In summary, we detected several methylation markers to discriminate AOD from HC populations with high accuracy. The alteration of DNA methylation in AOD differed from that in LOD. The spread of the distribution in AOD was smaller than those in HC and LOD. These results suggest that methylation markers of MDD should be examined under specific consideration of the onset age.

## Materials and methods

**Subjects.** The Institutional Review Boards of Yamaguchi University Hospital and Gunma University Hospital approved this study, and all subjects were fully informed about study aims and procedures before providing written informed consent for participation. This study was carried out in accordance with the latest version of the Declaration of Helsinki.

**Participants in Group 1.** All participants were recruited between 2012 and 2013. Elderly (age  $\geq 50$  years) patients with MDD and healthy participants were enrolled as previously described<sup>8</sup>. Briefly, patients with MDD, as well as HC participants, were assessed using clinical interviews, as well as a structured interview of the International Neuropsychiatric Interview (M.I.N.I., Japanese version 5.0.0), which is structured in accord with the DSM-IV<sup>42</sup>. We also defined a depressed state by a score of more than 18 using the SIGH-D<sup>43</sup>. Patients in remission met the DSM-IV criteria for full remission. Exclusion criteria for MDD patients included current or history of substance abuse/dependence, other psychiatric illnesses, other medical conditions (e.g., neurological diseases, severe liver failure, unstable hypertension), or family history of hereditary neurological disorders.

The samples of AOD participants (age of onset  $< 50$  years) in a depressed state ( $N = 10$ ), LOD participants (onset age  $\geq 50$  years) in a depressed state ( $N = 25$ ), and HC participants ( $N = 30$ ) were analyzed by genome-wide DNA methylation assays. Their demographic data are summarized in Table 1a (Group 1a; discovery samples). The AOD participants ( $N = 10$ ), a part of the LOD group ( $N = 10$ ), and some HC participants ( $N = 10$ ) were additionally evaluated using amplicon bisulfite sequencing and pyrosequencing. Their demographic data are summarized in Table 1b (Group 1b; training samples). In Group 1a, the following values were missing: the total duration parameter in the LOD group (1); the number of episodes in the AOD group (1); and the percentages of each WBC type in the HC (4), AOD (1), and LOD (2) groups. In contrast, the following values were missing in Group 1b: the number of episodes in the AOD group (1) and the percentages of each WBC type in the HC (1) and AOD (1) groups.

**Participants in Group 2.** The second group consisted of 11 AOD and 11 HC participants at the Yamaguchi site and the Gunma site. We enrolled 5 patients with AOD and 5 HC participants at the Yamaguchi site, as well as 6 AOD and 6 HC subjects at the Gunma site<sup>8,44</sup>, and the samples were collected between 2012 and 2015. All AOD patients and HC participants met the inclusion and exclusion criteria according to the protocol described for Group 1 with the only differences being the age restriction (age  $\geq 20$  years) and the SIGH-D scores ( $\geq 14$ ). The demographic data are summarized in Table 4 (Group 2; independent test sample).

**Genome-wide DNA methylation array.** Venous blood samples were collected from all participants between 9:00 am and 12:00 pm. Genomic DNA was purified from peripheral blood cells using the QIAamp DNA Blood Midi Kit (Qiagen, Chatsworth, CA, USA) according to the manufacturer's manual. The DNA quality was determined using a Spectrophotometer U-2900 (Hitachi High-Technologies, Tokyo, Japan), and an O.D. 260/280 ratio  $\geq 1.5$  was obligatory to consider samples for further analyses. DNA quantity was determined using the Quant-iT dsDNA Assay Kit, broad range (Thermo Fisher Scientific Inc., Waltham, MA, USA). Bisulfite conversion of 500 ng genomic DNA was performed using the EZ DNA Methylation Kit (Zymo Research, Irvine, CA, USA). The DNA methylation level was assessed according to the manufacturer's protocols using Infinium HumanMethylation450 BeadChips (Illumina Inc., San Diego, CA, USA). DNA methylation data were analyzed with GenomeStudio (version 2011.1) and Methylation Module (version 1.9.0; both Illumina Inc.). The raw data were normalized using the beta-mixture quantile normalization method<sup>45</sup>. The methylation ratio of each CpG site was determined as a beta-value from 0 (completely unmethylated) to 1 (fully methylated). The probes for the

analysis of methylation markers were excluded according to the following conditions: (1) the detection p-value was less than 0.01 in at least one sample, (2) the beta-value could not be calculated for at least one probe, and (3) all probes of single-nucleotide polymorphisms (rs#). The final data included 411,383 probes of CpG sites. The probes for the RefSeq analysis were excluded as follows: (1) the detection p-value was less than 0.01 in at least one sample and (2) all probes of single-nucleotide polymorphisms (rs#). These data finally included 481,247 probes. The RefSeq genes were identified according to the accession numbers of the probes. If multiple probes were associated with a promoter region, the median methylation frequency was calculated.

**Pyrosequencing.** Pyrosequencing was performed using the standard methods cited in a previous study<sup>46</sup>. Polymerase chain reaction (PCR) and sequencing primers (see Supplementary Table S5) were designed using Pyrosequencing Assay Design Software v2.0 (Qiagen).

Genomic DNA (200 ng) was bisulfite-treated using the EZ DNA Methylation-Gold Kit (Zymo Research) according to the manufacturer's protocol. Briefly, the reaction was performed 10 min at 98 °C and 2.5 h at 64 °C, then samples were stored at 4 °C. The converted samples were washed and desulfonated with M-wash and M-desulfonation buffers on a Zymo-Spin IC Column. The converted genomic DNA was eluted by adding 20 µl elution buffer. DNA samples were stored at -20 °C until further use.

PCRs were carried out in a volume of 20 µl with 20 ng or more of converted DNA, 2.5 µl 10× Taq buffer, 5 units Hot Start Taq Polymerase (Qiagen), 2 µl 2.5 mM dNTP mixture, 1 µl 10 pmol/µl Primer-S, and 1 µl 10 pmol/µl biotinylated-Primer-As. The amplification was carried out according to the general guidelines suggested for pyrosequencing: denaturation at 95 °C for 10 min, followed by 45 cycles at 95 °C for 30 s, at 48 °C for 30 s, and at 72 °C for another 30 s, and a final extension at 72 °C for 5 min. The ssDNA template was prepared from 16 to 18 µl biotinylated PCR product using Streptavidin Sepharose High Performance beads (GE Healthcare, Chicago, IL, USA) following the PSQ 96 sample preparation guide. Gel images of the PCR products are shown in Supplementary Fig. S10. For the analysis, 15 pmol of the respective sequencing primer were added. Sequencing was performed using a PyroMark ID system with the PyroMark Gold Reagents kit (Qiagen) according to the manufacturer's instruction without further optimization. The methylation percentage was calculated as the average of the degree of methylation at each CpG site formulated in pyrosequencing.

**Amplicon bisulfite sequencing.** For validation of the whole-genome methylation array data, we selected the loci of seven identified genes. Each primer for amplicon bisulfite sequencing was designed using the Methyl Primer Express software v1.0 (Thermo Fisher Scientific Inc.). Sodium bisulfite conversions of 500 ng genomic DNA were performed using the MethylEasy Xceed Rapid DNA Bisulphite Modification Kit (Genetic Signatures, New South Wales, Australia). PCR amplification of bisulfite-treated DNA was performed using TAKARA Epi-Taq HS (Takara Bio Inc., Kusatsu, Japan). The list of primer sequences and the summary of the PCR conditions are presented in Supplementary Table S6. The PCR products were directly cleaned up using NucleoSpin Gel (Takara Bio Inc.) and a PCR Clean-up kit (Macherey–Nagel GmbH & Co. KG, Düren, Germany). Gel images of the PCR products are shown in Supplementary Fig. S11. The sequence libraries were produced using the TruSeq Nano DNA LT Library Prep Kit (Illumina Inc.). The qualities of the libraries were checked using an Agilent 2100 BioAnalyzer (Agilent Technologies Inc., Santa Clara, CA, USA). Sequencing of 250 bp paired-end reads was generated using the Illumina MiSeq platform with MiSeq Reagent Kit ver3 (Illumina Inc.). Sequence data were analyzed using Genedata Expressionist for Genomic Profiling ver9.1.4a (Genedata Inc., Basel, Switzerland). Each read was trimmed to 150 bp, and mapping was performed relative to the h19/GRCh37 assembly of the human genome using Bismark Bisulfite Mapper ver0.14.4<sup>47</sup>. Percentages of methylcytosines were calculated using Bismark Bisulfite Mapper ver0.14.4.

**Statistical analysis.** Age distributions, total durations of all MDD episodes, and numbers of episodes of the participants were evaluated using the unpaired Student's t-test or the one-way analysis of variance (ANOVA) with the Tukey's post hoc test, whereas the sex distribution was analyzed by Fisher's exact test using the EZR software (version 1.37)<sup>48</sup>. The number of WBCs and the percentage of each WBC type were evaluated using the one-way ANOVA with the Tukey's post hoc test. Missing values were deleted using pairwise deletion methods.

We selected candidate methylation markers from 411,383 probes of CpG sites using the F-ratio as previously reported<sup>49</sup>. The F-ratio has been used as a popular criterion in feature selection of statistical pattern recognition. The F-ratio,  $f_j$ , for probe,  $j$ , is defined by:

$$f_j = \frac{(\mu_{1j} - \mu_{2j})^2}{\sigma_{1j}^2 + \sigma_{2j}^2}$$

where  $\mu_{ij}$  and  $\sigma_{ij}^2$  are the sample mean and sample variance, respectively, estimated from probe,  $j$ , using training samples from class,  $i$  ( $i = 1, 2$ ). The F-ratio measures the difference between two sample means normalized by the averaged sample variances. In feature selection, the F-ratio for each probe is first computed and probes are then ranked in order of decreasing F-ratio magnitude. The best probe set is given by some of the highest F-ratios.

The normal distribution of the methylation array data was checked using the Kolmogorov–Smirnov test. Pearson's correlation coefficients were used to measure correlations between participant age, onset age, number of episodes, total duration of all MDD episodes, SIGH-D scores, number of WBCs, percentages of each WBC type, and methylation level of each site. A false discovery rate method (Benjamini–Hochberg adjustment) was used to correct p-values.

A ROC analysis was carried out using EZR. We determined the optimal cut-off points by maximizing the Youden index (sensitivity + specificity - 1)<sup>50</sup>.

The discrimination performance was assessed as follows: two or three target markers were selected from the six identified probes (cg15794987, cg07763047, cg21347377, cg17458347, cg14567489, and cg16579770). Using the methylation levels of pyrosequencing or amplicon bisulfite sequencing as distinguishing features, we had a two- or three-dimensional vector  $x$ , whose components were the methylation levels. First, the sample mean vector  $\mu_i$  and the sample covariance matrix  $\sum_i$  of class  $\omega_i$  were estimated using 10 training samples from AOD and HC participants in Group 1b. The resulting Bayes classifier was assigned to a test sample  $x$  being classified to class  $\omega_1$ , i.e., HC, if

$$d_1(x) < d_2(x)$$

where  $d_i = \frac{1}{2}(x - \mu_i)^T \sum_i^{-1} (x - \mu_i)$ ,  $i = 1, 2$ .

For the methylation distribution analysis, we calculated the values of determinants using the methylation levels of the probes cg16579770 and cg07763047. In these calculations, 10 AOD samples were fixed, and 10 samples were 100 times randomly extracted from either the 30 HC samples or 25 LOD samples to calculate the average determinant of HC or LOD. The one-tailed t-test was performed with the null hypothesis that the ratio to HC is 1.

A summary of the statistical strategy used in this study is shown in Supplementary Fig. S12.

## Data availability

All data in this published article (and its Supplementary Information files) are available. The other data analyzed in the current study are not publicly available for ethical reasons.

Received: 2 May 2020; Accepted: 11 January 2021

Published online: 27 January 2021

## References

- Lopizzo, N. *et al.* Gene-environment interaction in major depression: focus on experience-dependent biological systems. *Front. Psychiatry* <https://doi.org/10.3389/fpsy.2015.00068> (2015).
- Patterson, C. C., Dahlquist, G. G., Gyurus, E., Green, A. & Soltesz, G. Incidence trends for childhood type 1 diabetes in Europe during 1989–2003 and predicted new cases 2005–20: a multicentre prospective registration study. *Lancet* **373**, 2027–2033. [https://doi.org/10.1016/S0140-6736\(09\)60568-7](https://doi.org/10.1016/S0140-6736(09)60568-7) (2009).
- Zuo, L., Pannell, B. K. & Liu, Z. Characterization and redox mechanism of asthma in the elderly. *Oncotarget* **7**, 25010–25021. <https://doi.org/10.18632/oncotarget.7075> (2016).
- Grayson, L. & Thomas, A. A systematic review comparing clinical features in early age at onset and late age at onset late-life depression. *J. Affect. Disord.* **150**, 161–170. <https://doi.org/10.1016/j.jad.2013.03.021> (2013).
- Herrmann, L. L., Goodwin, G. M. & Ebmeier, K. P. The cognitive neuropsychology of depression in the elderly. *Psychol. Med.* **37**, 1693–1702. <https://doi.org/10.1017/S0033291707001134> (2007).
- Herrmann, L. L., Le Masurier, M. & Ebmeier, K. P. White matter hyperintensities in late life depression: a systematic review. *J. Neurol. Neurosurg. Psychiatry* **79**, 619–624. <https://doi.org/10.1136/jnnp.2007.124651> (2008).
- Yamagata, H. *et al.* Identification of commonly altered genes between in major depressive disorder and a mouse model of depression. *Sci. Rep.* **7**, 3044. <https://doi.org/10.1038/s41598-017-03291-x> (2017).
- Miyata, S. *et al.* Blood transcriptomic markers in patients with late-onset major depressive disorder. *PLoS ONE* **11**, e0150262. <https://doi.org/10.1371/journal.pone.0150262> (2016).
- Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6–21. <https://doi.org/10.1101/gad.947102> (2002).
- Ziller, M. J. *et al.* Charting a dynamic DNA methylation landscape of the human genome. *Nature* **500**, 477–481. <https://doi.org/10.1038/nature12433> (2013).
- Uchida, S., Yamagata, H., Seki, T. & Watanabe, Y. Epigenetic mechanisms of major depression: targeting neuronal plasticity. *Psychiatry Clin. Neurosci.* **72**, 212–227. <https://doi.org/10.1111/pcn.12621> (2018).
- Story Jovanova, O. *et al.* DNA methylation signatures of depressive symptoms in middle-aged and elderly persons: meta-analysis of multiethnic epigenome-wide studies. *JAMA Psychiatry* **75**, 949–959. <https://doi.org/10.1001/jamapsychiatry.2018.1725> (2018).
- Li, M. *et al.* What do DNA methylation studies tell us about depression? A systematic review. *Transl. Psychiatry* **9**, 68. <https://doi.org/10.1038/s41398-019-0412-y> (2019).
- Higuchi, F. *et al.* State-dependent changes in the expression of DNA methyltransferases in mood disorder patients. *J. Psychiatr. Res.* **45**, 1295–1300. <https://doi.org/10.1016/j.jpsychires.2011.04.008> (2011).
- Mikeska, T., Candiloro, I. L. & Dobrovic, A. The implications of heterogeneous DNA methylation for the accurate quantification of methylation. *Epigenomics* **2**, 561–573. <https://doi.org/10.2217/epi.10.32> (2010).
- Jeremian, R. *et al.* Investigation of correlations between DNA methylation, suicidal behavior and aging. *Bipolar Disord.* **19**, 32–40. <https://doi.org/10.1111/bdi.12466> (2017).
- Bishop, J. R. *et al.* Methylation of FKBP5 and SLC6A4 in relation to treatment response to mindfulness based stress reduction for posttraumatic stress disorder. *Front. Psychiatry* **9**, 418. <https://doi.org/10.3389/fpsy.2018.00418> (2018).
- Naismith, S. L., Norrie, L. M., Mowszowski, L. & Hickie, I. B. The neurobiology of depression in later-life: clinical, neuropsychological, neuroimaging and pathophysiological features. *Prog. Neurobiol.* **98**, 99–143. <https://doi.org/10.1016/j.pneurobio.2012.05.009> (2012).
- Crawford, B. *et al.* DNA methylation and inflammation marker profiles associated with a history of depression. *Hum. Mol. Genet.* **27**, 2840–2850. <https://doi.org/10.1093/hmg/ddy199> (2018).
- Bostrom, A. E. *et al.* A MIR4646 associated methylation locus is hypomethylated in adolescent depression. *J. Affect. Disord.* **220**, 117–128. <https://doi.org/10.1016/j.jad.2017.05.017> (2017).
- Walker, R. M. *et al.* DNA methylation in a Scottish family multiply affected by bipolar disorder and major depressive disorder. *Clin. Epigenetics* **8**, 5. <https://doi.org/10.1186/s13148-016-0171-z> (2016).
- Numata, S. *et al.* Blood diagnostic biomarkers for major depressive disorder using multiplex DNA methylation profiles: discovery and validation. *Epigenetics* **10**, 135–141. <https://doi.org/10.1080/15592294.2014.1003743> (2015).
- Cordova-Palamera, A. *et al.* Genome-wide methylation study on depression: differential methylation and variable methylation in monozygotic twins. *Transl. Psychiatry* **5**, e557. <https://doi.org/10.1038/tp.2015.49> (2015).
- Davies, M. N. *et al.* Hypermethylation in the ZBTB20 gene is associated with major depressive disorder. *Genome Biol.* **15**, R56. <https://doi.org/10.1186/gb-2014-15-4-r56> (2014).
- Byrne, E. M. *et al.* Monozygotic twins affected with major depressive disorder have greater variance in methylation than their unaffected co-twin. *Transl. Psychiatry* **3**, e269. <https://doi.org/10.1038/tp.2013.45> (2013).



26. Zhu, Y. *et al.* Genome-wide profiling of DNA methylome and transcriptome in peripheral blood monocytes for major depression: a monozygotic discordant twin study. *Transl Psychiatry* **9**, 215. <https://doi.org/10.1038/s41398-019-0550-2> (2019).
27. Starnawska, A. *et al.* Epigenome-wide association study of depression symptomatology in elderly monozygotic twins. *Transl Psychiatry* **9**, 214. <https://doi.org/10.1038/s41398-019-0548-9> (2019).
28. Clark, S. L. *et al.* A methylation study of long-term depression risk. *Mol. Psychiatry* <https://doi.org/10.1038/s41380-019-0516-z> (2019).
29. Allen, N. C. *et al.* Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: the SzGene database. *Nat. Genet.* **40**, 827–834. <https://doi.org/10.1038/ng.171> (2008).
30. Shi, J., Gershon, E. S. & Liu, C. Genetic associations with schizophrenia: meta-analyses of 12 candidate genes. *Schizophr. Res.* **104**, 96–107. <https://doi.org/10.1016/j.schres.2008.06.016> (2008).
31. Perlis, R. H. *et al.* Family-based association study of lithium-related and other candidate genes in bipolar disorder. *Arch. Gen. Psychiatry* **65**, 53–61. <https://doi.org/10.1001/archgenpsychiatry.2007.15> (2008).
32. Zhao, J. *et al.* Gene expression of GABA and glutamate pathway markers in the prefrontal cortex of non-suicidal elderly depressed patients. *J. Affect. Disord.* **138**, 494–502. <https://doi.org/10.1016/j.jad.2012.01.013> (2012).
33. Yin, H. *et al.* A pilot integrative genomics study of GABA and glutamate neurotransmitter systems in suicide, suicidal behavior, and major depressive disorder. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **171B**, 414–426. <https://doi.org/10.1002/ajmg.b.32423> (2016).
34. Zong, L. *et al.* Genetic and epigenetic regulation on the transcription of GABRB2: genotype-dependent hydroxymethylation and methylation alterations in schizophrenia. *J. Psychiatr. Res.* **88**, 9–17. <https://doi.org/10.1016/j.jpsychires.2016.12.019> (2017).
35. Madak, J. T., Bankhead, A. 3rd., Cuthbertson, C. R., Showalter, H. D. & Neamati, N. Revisiting the role of dihydroorotate dehydrogenase as a therapeutic target for cancer. *Pharmacol. Ther.* **195**, 111–131. <https://doi.org/10.1016/j.pharmthera.2018.10.012> (2019).
36. Reis, R. A. G., Calil, F. A., Feliciano, P. R., Pinheiro, M. P. & Nonato, M. C. The dihydroorotate dehydrogenases: past and present. *Arch. Biochem. Biophys.* **632**, 175–191. <https://doi.org/10.1016/j.abb.2017.06.019> (2017).
37. Shen, W. *et al.* Discovery of a new structural class of competitive hDHODH inhibitors with in vitro and in vivo anti-inflammatory, immunosuppressive effects. *Eur. J. Pharmacol.* **791**, 205–212. <https://doi.org/10.1016/j.ejphar.2016.09.004> (2016).
38. Barnes, J., Mondelli, V. & Pariante, C. M. Genetic contributions of inflammation to depression. *Neuropsychopharmacology* **42**, 81–98. <https://doi.org/10.1038/npp.2016.169> (2017).
39. Xu, J. *et al.* Identification of a novel human DDX40 gene, a new member of the DEAH-box protein family. *J. Hum. Genet.* **47**, 681–683. <https://doi.org/10.1007/s100380200104> (2002).
40. Martin, R., Straub, A. U., Doebele, C. & Bohnsack, M. T. DEXD/H-box RNA helicases in ribosome biogenesis. *RNA Biol.* **10**, 4–18. <https://doi.org/10.4161/rna.21879> (2013).
41. Hori, H. *et al.* Integrated profiling of phenotype and blood transcriptome for stress vulnerability and depression. *J. Psychiatr. Res.* **104**, 202–210. <https://doi.org/10.1016/j.jpsychires.2018.08.010> (2018).
42. Otsubo, T. *et al.* Reliability and validity of Japanese version of the Mini-International Neuropsychiatric Interview. *Psychiatry Clin. Neurosci.* **59**, 517–526. <https://doi.org/10.1111/j.1440-1819.2005.01408.x> (2005).
43. Williams, J. B. A structured interview guide for the Hamilton Depression Rating Scale. *Arch. Gen. Psychiatry* **45**, 742–747 (1988).
44. Harada, K. *et al.* Disrupted orbitomedial prefrontal limbic network in individuals with later-life depression. *J. Affect. Disord.* **204**, 112–119. <https://doi.org/10.1016/j.jad.2016.06.031> (2016).
45. Teschendorff, A. E. *et al.* A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics* **29**, 189–196. <https://doi.org/10.1093/bioinformatics/bts680> (2013).
46. Kim, M. S. *et al.* Genome-wide identification of OTP gene as a novel methylation marker of breast cancer. *Oncol. Rep.* **27**, 1681–1688. <https://doi.org/10.3892/or.2012.1691> (2012).
47. Krueger, F. & Andrews, S. R. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**, 1571–1572. <https://doi.org/10.1093/bioinformatics/btr167> (2011).
48. Kanda, Y. Investigation of the freely available easy-to-use software ‘EZ’ for medical statistics. *Bone Marrow Transplant.* **48**, 452–458. <https://doi.org/10.1038/bmt.2012.244> (2012).
49. Iizuka, N. *et al.* Oligonucleotide microarray for prediction of early intrahepatic recurrence of hepatocellular carcinoma after curative resection. *The Lancet* **361**, 923–929. [https://doi.org/10.1016/s0140-6736\(03\)12775-4](https://doi.org/10.1016/s0140-6736(03)12775-4) (2003).
50. Akobeng, A. K. Understanding diagnostic tests 3: Receiver operating characteristic curves. *Acta Paediatr* **96**, 644–647. <https://doi.org/10.1111/j.1651-2227.2006.00178.x> (2007).

## Acknowledgements

This study was supported by the “Integrated Research on Neuropsychiatric Disorders” project conducted under the Strategic Research Program for Brain Sciences from the MEXT and AMED, the JSPS KAKENHI (16K10189 and 20K07946), the Core Research for Evolutional Science and Technology (CREST), the Industrial Strategic Research and Development from the Yamaguchi Prefecture, SENSHIN Medical Research Foundation, and contributions by Pfizer.

The authors would like to thank TAKARA Bio Inc. for assistance with the genome-wide DNA methylation array. We also thank Mr. Tatsuhiko Adachi (ALLIANCE Biosystems, Inc.) for assistance with the pyrosequencing analysis.

## Author contributions

H.Y., K.M., S.U., S.M., M.F., M.M., Y.H., Y.W., and S.N. designed the experiments and interpreted the data. H.Y., K.M., K.H., and Y.W. performed psychiatric diagnosis and/or blood sampling for the human studies. A.K. performed blood sample preparation and DNA isolation from blood samples. H.Y., H.O., T.S., M.K., C.C., Y.H., and S.N. analyzed the results. The manuscript was written by H.Y. All authors have reviewed the manuscript.

## Competing interests

The authors report here all financial or other relationships that are relevant to the subject of this article. H.Y. has received a research grant from Pfizer, Eisai, and MSD. K.M. has received honoraria from Meiji Seika Pharma, Kyowa Pharmaceutical Industry, Mochida Pharmaceutical, Eli Lilly, Janssen Pharmaceutical K.K., and Pfizer. S.N. has received honoraria and/or research grant support from Otsuka Pharmaceutical, Meiji Seika Pharma, Sumitomo Dainippon Pharma, Kyowa Pharmaceutical Industry, Shionogi, Mitsubishi Tanabe Pharma, Mochida

Pharmaceutical, Eisai, Tsumura, Eli Lilly, MSD, Astellas, and Pfizer. The other authors declare no actual or potential conflict of interest.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-81758-8>.

**Correspondence** and requests for materials should be addressed to H.Y.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021