Lower DNA methylation levels in CpG island shores of <i>CR1</i> , <i>CLU</i> , and <i>PICALM</i> genes in the blood of Alzheimer's disease patients
Risa Mitsumori ¹ , Kazuya Sakaguchi ^{2, #a} , Shumpei Niida ¹ , and Nobuyoshi Shimoda ^{2*}
¹ Medical Genome Center, ² Department of Regenerative Medicine, National Center for
Geriatrics and Gerontology, Obu, Aichi, Japan **Current address: Axcelead Drug Discovery Partners, Inc., Fujisawa, Kanagwa, Japan
*Corresponding author
E-mail: shimoda@ncgg.go.jp (NS)

Abstract

The aim of the present study was to (1) investigate the relationship between late onset AD and DNA methylation levels in the top six Alzheimer's disease (AD)-related genes in blood and (2) examine its applicability to the diagnosis of AD. We examined methylation differences at CpG island shores in the top six genes using Sanger sequencing, and one of two groups of 48 AD patients and 48 elderly controls was used for a test or replication analysis. We found that methylation levels in three out of the six genes, *CR1*, *CLU*, and *PICALM*, were lower in AD subjects. The combination of *CLU* methylation levels and the *APOE* genotype classified AD patients with AUC=0.84 and 0.80 in the test and replication analyses, respectively. Our results implicate methylation differences at the CpG island shores of AD-related genes in the onset of AD and suggest their diagnostic value.

Introduction

Moderate concordance rates for Alzheimer's disease (AD) among genetically identical twins suggest the possible involvement of epigenetics in the etiology of AD [1, Among epigenetic components, DNA methylation and chromatin modifications are of great interest in AD because they have been shown to change with aging [3, 4], which is the main cause of the disease, and their unintended changes affect gene expression [5]. Since DNA methylation differences in genomes between AD patients and cognitive normal elderly individuals are expected to not only reveal AD-susceptible genes, but also provide biomarkers for clinical purposes, they have extensively examined in blood and brains using gene-specific and genome-wide approaches [6-10]. Sanger or pyrosequencing and DNA methylation arrays have generally been employed for region-specific and genome-wide analyses, respectively. Although large numbers of differentially methylated genes have been reported, only a few genes in the brain and blood have been independently confirmed [7, 11]. These findings suggest that methylation anomalies in AD samples are so marginal that depending on the sample size or sensitivity of the procedures employed, they cannot reproducibly be detected [12, 13]. Another possibility is that genomic regions with evident differences were not simply targeted by these studies. The whole-genome bisulfite sequencing of large samples of

cognitively normal and AD subjects with sufficient coverage may be the only approach to circumvent these limitations; however, this is prohibitive for large-scale studies due to the associated costs.

We are interested in identifying differentially methylated regions (DMRs) in blood DNA that may be used for the diagnosis of AD and its prognosis. We assumed that DMRs in AD-related genes, such as those identified by genome-wide association studies (GWASs) [14, 15], in the brain or any other tissues confer susceptibility to the onset of AD through subtle, but long-lasting expression changes in these genes, similar to single nucleotide polymorphisms (SNPs). Based on this assumption, we herein examined whether DMRs exist in AD-associated genes in the blood of late-onset AD (LOAD) patients. We focused on a genic region designated as CpG island shores in AD-associated genes because methylation levels at CpG island shores are vulnerable under various conditions, such as tissue differentiation, reprogramming, aging, and disease including AD [16-20]. CpG island shores are defined as 2-kb-long regions that lie on both sides of a CpG island [16]. CpG islands are, on average, 1000 base pairs (bp) in length, characterized by dense clusters of CpG dinucleotides, and located around the promoter regions of 70% of human genes [21]. CpG islands are exempt from DNA methylation irrespective of gene expression, and the further away CpG dinucleotides are from CpG islands, the higher the chance of methylation. Therefore, methylated cytosines appear to be symmetrically distributed with respect to promoter CpG islands, and the cytosines of CpGs in CpG island shores are often slightly and moderately methylated [16]. A previous study demonstrated that gene expression levels were negatively associated with methylation levels at CpG island shores [16].

We herein demonstrated that, in blood, AD-related DMRs existed in CpG island shores in three out of six AD-related genes examined, *CR1* (complement receptor 1), *CLU* (clusterin), and *PICALM* (phosphatidylinositol-binding clathrin assembly protein), and also that with the combination of *APOE* (apolipoprotein E) genotypes, the methylation levels of *CLU* in blood may be effectively used to differentiate AD patients from cognitive elderly with high sensitivity and specificity.

Materials and methods

Ethics statements

The present study was conducted with the informed consent of all individuals and with the approval of the Ethical Committee of the National Center for Geriatrics and Gerontology (NCGG). The design and performance of the present study involving human subjects were clearly described in a research proposal. All participants were voluntary and completed informed consent in writing before registering at the NCGG biobank, which collects human biomaterials and data for geriatrics research.

Patient sample collection

All 293 serum subjects and the associated clinical data were provided by the NCGG Biobank. Ninety-six subjects were patients with AD, 40 with VaD, 34 with DLB, and 27 with FTD, and 96 subjects were cognitively normal elderly controls (hereafter, controls). One of two groups of 48 AD patients and 48 age- and gender-matched controls was used as a test group and the other as a replication group. The cognitive status and severity of dementia were assessed by the Mini-Mental State Examination (MMSE). The status of the *APOE* e4 allele genotype (the main genetic risk factor for AD) and the MMSE score were obtained from the NCGG biobank. All

subjects were >60 years in age. All control subjects had a MMSE score of >25.

Methylation analyses

The methylation levels of target regions were assessed by Sanger sequencing, pyrosequencing, or by both methods. In either method, 200 ng of genomic DNA was initially treated with sodium bisulfite to convert non-methylated cytosines to uracil using the EZ DNA Methylation-Gold Kit (Zymo research) following the manufacturer's Target regions were PCR amplified by primer sets designed by the web tool, MethPrimer [22], for Sanger sequencing or by Pyrosequencing Assay Design Software ver.2.0 provided by Qiagen for pyrosequencing. The conditions of PCR, lengths of amplicons, and sequences of primers were shown in S1 Table. In Sanger sequencing, amplicons were cloned into the pGEM-T vector (Promega), which was then used to transform the E. coli strain, DH5α (Takara Bio). Plasmid DNA was amplified from each of 32 colonies by Templiphi (GE Healthcare) and used as a template for a dideoxy sequencing reaction (BigDye ver.3.1, Applied Biosystems). Sequences were elucidated on a Genetic Analyzer 3500 or 3130xL (Applied Biosystems), and data containing a minimum of twenty sequences were analyzed by QUMA [23] to quantify the mean methylation levels of PCR products. In pyrosequencing, the mean

methylation levels of PCR amplicons were directly quantified on PyroMark Q48 (Qiagen).

Identification of CpG island shores

To identify CpG island shores in AD-associated genes, we plotted CpG dinucleotide sites along the genes using a program designated CyGnusPlotter, which was written for the present study and is freely available via the internet (https://github.com/kzyskgch/CyGnusPlotter). These diagrams showed that, as reported in many human genes, there was a cluster of CpGs at their promoter regions, named the CpG island. Since CpG island shores are defined as regions that immediately flank CpG islands, the exact boundary was only identified after the edges of the hypomethylated region were confirmed.

Genotypes of SNPs

The genotypes of AD-related SNPs in *CR1*, *CLU*, and *PICALM* loci in our samples were provided by the NCGG Biobank, which supported a genome-wide association study (GWAS) in Japan to identify SNPs linked to AD. Since most samples used in the present study were part of larger samples utilized for GWAS, we were able to obtain the genotypes of SNPs in *CR1*, *CLU*, and *PICALM* loci in

substantial numbers of our samples after due formalities.

Statistical analyses

Comparisons between two and more groups were made using the Student's *t*-test and ANOVA, respectively, by PRISM ver.5 (GraphPad Software, Inc., San Diego, CA, USA). The sensitivity and specificity of the measured variables for the diagnosis of AD were evaluated using a receiver operating characteristic (ROC) curve. ROC plots were drawn by Minitab 17 (Minitab Inc.). Regarding multiple biomarkers, logistic regression analyses were conducted to derive analytical expression for the risk of AD using methylation levels as continuous variables and the APOE genotype as a nominal variable. The classification performance of *CLU* methylation, the *APOE* genotype, and the combination of both, were assessed using the area under the ROC curve (AUC). ROC is a plot of the probability of correctly classifying positive samples against the rate of incorrectly classifying true negative samples. Therefore, the AUC measure of an ROC plot is a measure of predictive accuracy. A DeLong test was used to compare AUC between groups [24]. All tests were two-tailed, and a p value <0.05 was considered to be significant.

9

Results

Selection of target regions

We selected target regions in the CpG island shores of six known AD-related genes in the AlzGene database ([25], http://www.alzgene.org/) for methylation analyses: APOE, CLU, CR1, PICALM, BIN1, and ABCA7 as follows: we initially referred to the distribution of CpG dinucleotides along the genes (Fig 1 and S1-S5 Figs). Promoter CpG islands were easily recognized as a cluster of CpG dinucleotides around the 5' ends of the genes, and regions juxtaposed to the promoter CpG islands were putative CpG island shores. We PCR-amplified portions of the putative CpG island shores at which at least several CpGs existed and quantified their methylation levels. Slightly or moderately methylated portions that met the criteria for CpG island shores were used in subsequent analyses. These portions were only found in a limited space in the case of the APOE gene (region II in Fig 1).

Fig 1. Relative positions of CpG dinucleotides selected for the methylation analysis in the APOE gene.

The schematic shows the distribution of CpG dinucleotides along the *APOE* gene. The position of the transcription start site is defined as +1; hence, "-2,000" in the parenthesis

indicates that the diagram includes 2 kb upstream of the APOE gene, while "5,582" in the parenthesis is the position 2 kb downstream from the end of the last exon. Vertical lines indicate the positions of CpG dinucleotides. Open rectangles depict exons and exon numbers are shown above the rectangles. Four arrows indicate Alu elements in the locus and their orientations relative to the coding strand of the APOE gene [26]. The fundamentals of these diagrams were automatically drawn by the web-based tool, CyGnusPlotter. It collects the genomic structure of the most representative isoform of a requested gene from the Ensembl database with upstream and downstream regions of the designated lengths. Black horizontal bars with Greek letters (I-V) indicate the regions for bisulfite sequencing. Filled and open circles show methylated and unmethylated cytosines, respectively. Each column represents a unique CpG site in the examined amplicon, and each line shows an individual DNA clone. indicate positions at which CpG is absent due to DNA polymorphisms. percentages of the mean methylation levels derived from four subjects were shown at the bottom. There was a conventional CpG island (region III) that covers exon 1. Moderate DNA methylation was observed at a limited region (II) juxtaposed to the CpG island, a region called the "CpG island shore". Abbreviations: APOE, apolipoprotein E.

DNA methylation differences

We quantified the methylation levels of target regions in the six genes in 48 control and AD blood samples by cloning and sequencing PCR products from bisulfite-treated genomic DNAs. Demographic data for the samples included in the present study are shown in Table 1. When stratified by the disease status, we found that the mean methylation levels in three out of the six genes, PICALM, CR1, and CLU, were significantly lower in AD than in control samples (Fig 2). These results suggested that the development of AD was related to methylation changes in the CpG island shores of AD-related genes. AD-related hypomethylation in blood samples was also reported for TREM2 (Triggering receptor expressed on myeloid cells 2), another AD-related gene [27]. We judged that the region surveyed in their study corresponded to the CpG island shore of TREM2 based on its location (S6 Fig) and methylation level [27]. We compared the methylation levels of the same region in our control and AD samples by following the procedure reported previously. As shown in Fig 2, the mean methylation level was significantly lower in AD than in control samples, which was consistent with previous findings [27].

Table 1. Characteristics of subjects

	Test (n=96)		Replication (n=96)		non-AD dementia (n=95)		
	Controls	AD	Controls	AD	DLB	VaD ^a	$FTLD^b$
	(n=48)	(n=48)	(n=48)	(n=48)	(n=34)	(n=40)	(n=27)
Age, mean (SD)	71.9 (3.7)	72.7 (5.4)	70.2 (1.5)	70.5 (1.8)	79 (5.4)	79 (6.9)	70.2 (8.8)
Woman, No (%)	22 (46)	24 (50)	26 (54)	24 (50)	21 (62)	15 (38)	17 (63)
MMSE score, mean(SD)	28.5 (1.5)	13.3 (5.1)	29.4 (0.8)	17.7 (3.8)	16.7 (4.7)	17.4 (4.8)	16.6 (6.9)
APOE & homozygote, No. (%)	1 (2.1)	9 (19)	0 (0)	4 (8.3)	3 (8.8)	1 (2.6)	0 (0)
APOE & heterozygote, No. (%)	9 (19)	22 (46)	6 (13)	28 (58)	9 (26)	7 (18)	10 (37)
APOE & non-carriers, No. (%)	38 (79)	17 (35)	42 (88)	16 (33)	22 (65)	31 (79)	17 (63)

^aThe MMSE score of one subject and the *APOE* genotype of another subject were not available. ^bThe MMSE score of one subject was not available. Abbreviations: *APOE* £4, the £4 allele of the human apolipoprotein E gene; ROC, receiver operating characteristic curve; AD, Alzheimer's disease; DLB, dementia with Lewy bodies; VaD, vascular dementia; FTD, frontotemporal dementia.

Fig 2. Methylation rates at CpG island shores in seven AD-associated genes in the peripheral blood of control and AD subjects.

Each dot represents the mean methylation rate at the CpG island shores of the indicated genes in a subject. Subjects were composed of a group of 48 cognitive normal elderly (Ctrl) and 48 AD patients, named the test group, except for TREM2, for which another set of 96 subjects, named the replication group, was also examined. Methylation rates were quantified by Sanger sequencing, except for TREM2, for which pyrosequencing was employed [27]. The relative position of the CpG island shore investigated in APOE was region II in Fig 1, while those in the other six genes are shown in S1-S6 Figs. The horizontal lines and error bars show the mean \pm standard deviation. p-values were obtained using the Student's t-test. Abbreviations: Ctrl, control; AD, Alzheimer's disease

Replication and verification of methylation differences

To assess the reproducibility of the results obtained above, we compared the methylation levels of *APOE* and *CLU* in the replication set of 48 blood samples of control and AD using the same method as for the test set, and confirmed the absence and presence of methylation differences in *APOE* and *CLU*, respectively (Fig 3A, B).

We then verified the methylation levels of *CLU* in two sets of blood samples using pyrosequencing, which directly measures the methylation levels of PCR products (Fig 3C, D). Lower methylation levels in AD samples were confirmed in either set of blood samples. These results indicate that methylation differences between controls and AD patients were reproducible.

Fig 3. Confirmation of the absence and presence of methylation differences at CpG

island shores of APOE and CLU, respectively, between control and AD subjects.

Methylation rates in the replication group were quantified by Sanger sequencing for the

CpG island shores of APOE (A) and CLU (B). Pyrosequencing was also employed for

CLU to examine methylation differences in the test (C) and replication groups (D).

The positions surveyed by Sanger sequencing were the same as those in Figure 2. The

position of the genomic region in *CLU* quantified by pyrosequencing is shown in S2 Fig.

Abbreviations: Ctrl, control subjects; AD, Alzheimer's disease subjects.

No effect of SNPs, MMSE, or age on methylation

SNPs that correlate with the methylation degree of nearby CpG sites in cis are called methylation quantitative trait loci (mQTLs). Previous studies suggested that

15

DNA methylation acts as an intermediary of genetic risk [28-30] and mQTLs associated with AD have been reported [31, 32]. Since the six AD-associated genes examined above were all discovered through GWAS, AD-related SNPs exist that are linked to *CLU*, *PICALM*, and *CR1*, in which lower methylated regions in AD blood were identified. Therefore, these SNPs may cause methylation differences in the three genes. To examine this possibility, we assessed the relationship between AD-related SNPs and the methylation degree, but found no relationship in any of the genes (Fig 4A-C). These results demonstrated that these SNPs and methylation are independently associated with the onset of AD. The independence of AD pathology-related methylation changes from AD risk variants has also been reported [10, 30].

We investigated whether the degree of methylation in the three genes was related to the scores of the Mini-Mental State Examination (MMSE) and age (Fig 4D-I) based on Pearson's correlation coefficient. However, no relationships were observed in any of the three genes. These results indicate that methylation levels in the genes do not reflect the severity of disease and they are unrelated to aging in the elderly.

Fig 4. Methylation rates in AD-related DMRs are independent of known AD-related SNPs, MMSE, and age.

Methylation rates in AD-related DMR in CR1, CLU, and PICALM were plotted against the genotypes of AD-related SNPs in *CR1* (rs3818361) (A) and *CLU* (rs11136000) (B) and upstream of *PICALM* (rs3851179) (C). No significant differences were detected in any combinations of the genotypes in any regions with the Kruskal-Wallis test [non-parametric one-way ANOVA] with Dunn's multiple comparison post-test. Methylation rates in DMR in CR1, CLU, and PICALM are also plotted against MMSE scores (D, E, F) and age (G, H, I), respectively. Relationships were analyzed based on Pearson's correlation coefficient. The numbers of samples examined are shown in parentheses. Methylation rates are derived from AD and control subjects in the test group for (A), (C), (G), and (I), and from those in the test and replication groups for (B) and (H). Data from AD subjects in the test group or in the test and replication groups were used to plot methylation rates against MMSE for (D) and (F) or for €, respectively. All methylation rates were quantified at the same regions as in Fig 2 by Sanger sequencing. Relative positions of SNPs in the three genes and information on the SNPs genotyped in the test and replication groups are shown in S1-S3 Figs and S2 Table, respectively. Abbreviations: Ctrl subjects, control; AD, Alzheimer's disease subjects; DMRs, differentially methylated regions; SNPs, single nucleotide

polymorphisms; MMSE, Mini-Mental State Examination; ANOVA, analysis of variance.

Methylation in non-AD dementia

To clarify how specific the methylation changes detected may be in dementia, we examined the methylation levels of *APOE* and *CLU* in blood samples of control, AD, dementia with Lewy bodies (DLB), vascular dementia (VaD), and frontotemporal dementia (FTD). Although no significant differences were observed in the methylation levels of *APOE* between the control and any dementia (Fig 5A), *CLU* methylation levels were lower in DLB than in the control, similar to AD samples (Fig 5B). This result indicates that *CLU* hypomethylation is related to limited types of dementia.

Fig 5. Methylation rates at CpG island shores of APOE and CLU in non-AD dementia.

Methylation rates at the CpG island shores of *APOE* and *CLU* in DLB, VaD, and FTD were quantified by Sanger sequencing and differences relative to control subjects in the test group were assessed using one-way ANOVA with Dunnett's multiple comparison

18

post-test. The positions of the CpG island shores examined were the same as those in Fig 2. Abbreviations: Ctrl, control; AD, Alzheimer's disease; DLB, dementia with Lewy bodies; VaD, vascular dementia; FTD, frontotemporal dementia.

Blood-brain methylation discordance

In view of etiology, it is important to establish whether methylation changes in blood also occur in the brains of the same individuals. Since we had no brain samples available from the same individuals that provided the blood samples used in this study, we needed to clarify whether the degree of DNA methylation in the brain reflects that in blood in the same individuals at the three AD-related hypomethylated regions. To achieve this, we utilized two online searchable databases, designated BECon (Blood-Brain Epigenetic Concordance) [33] and the Blood Brain DNA methylation Comparison Tool [34], both of which show the extent to which DNA methylation at a given CpG site in blood is related to that in three or four different regions, respectively, in the brains of matched individuals (S7 and S8 Figs). Both studies employed the Illumina 450K HumanMethylation array (HM450K), which has five probes against the regions we surveyed: three probes for the CLU region and one each for CR1 and PICALM. We found that the CpG sites targeted by the probes in the brain showed either low or moderate methylation levels, suggesting that they were in the CpG island shores. However, interindividual variations in the methylation levels of these CpGs were less in the brain than in blood, and no correlation was noted between blood and any brain regions, except for a CpG site in *PICALM* in the entorhinal cortex (S8C Fig). These results indicated that overall methylation levels in the CpG island shores of the three genes in blood do not reflect those in brain.

Classification utility

The identification of subjects at high risk of developing AD is important for early interventions and clinical trials. Multiple logistic regression (MLR) analyses were used to derive linear classifier models that differentiate control and AD subjects using the data of test and replication groups (S3 Table). Methylation levels of *CLU* in the test and replication groups were employed for the analyses with the number of *APOE* £4 alleles as covariates since the *APOE* epsilon4 (£4) allele is the strongest genetic risk factor for AD. The performance of each model was assessed by ROC, which establishes classification values, and AUC, which assesses multimarker classification performance. Stepwise feature selection selected the top-performing linear combination of *CLU* methylation and the *APOE* genotype as MLR variables in

both groups, which attain AUCs of 0.83 and 0.85 by the data of the test and replication groups, respectively (Fig 6A). Since slightly better classification performance was obtained with the data of the replication group, we applied the model to the data of the test group, which still yielded high AUC of 0.80 (Fig 6B). In both cases, the differences in AUCs between APOE $\varepsilon 4$ data alone and the combination of methylation and APOE $\varepsilon 4$ data were significant (DeLong's test; p=0.0083 for A, p= 0.0064 for B).

Fig 6. ROC results for classifying control and AD subjects.

ROC analyses using either methylation rates at the CpG island shore of *CLU* (long dashed line), the number of the *APOE* & allele (short dashed line), or the combination of both (solid line). A plot of ROC results from the replication group (A). The AUCs obtained were 0.66 for the methylation rate only (95% CI: 0.55-0.77), 0.78 for the *APOE* & allele number only (95% CI: 0.68-0.87), and 0.85 for both (95% CI: 0.77-0.93). The model developed from the replication group was applied to the test group to draw another ROC plot (B). The AUCs obtained were 0.69 for the methylation rate only (95% CI: 0.56-0.80), 0.73 for the *APOE* & allele number only (95% CI: 0.63-0.83), and 0.80 for both (95% CI: 0.71-0.89). ROC plots represent sensitivity (i.e., true positive rate) versus 1-specificity (i.e., false positive rate).

Abbreviations: ROC, receiver operating characteristic curve; AD, Alzheimer's disease;

AUC; area under curve; CI, confidence interval.

Discussion

In the present study, Sanger sequencing of more than twenty thousand clones from 96 blood samples quantified the methylation levels of the CpG island shores of the top six AD-associated genes in the AlzGene database [25], which provides a comprehensive field synopsis of genetic association research conducted on AD. Although it was a very low throughput, we detected decreases in DNA methylation in three out of the six genes. Furthermore, we replicated AD-related DNA hypomethylation in the CpG island shore of TREM2 in our blood samples using pyrosequencing. While significant DNA methylation changes were not observed in APOE, BIN1, or ABCA7, the regions surveyed were limited fractions of the CpG island shores in these genes, and, thus, we cannot deny the possibility of AD-related methylation changes in other regions in these genes. Alternatively, AD-associated DNA methylation changes might occur in a tissue-specific manner since CpG dinucleotides with AD pathology-related methylation were found in the ABCA7 and BIN1 loci in the brain's DNA [10, 20, 35].

Genome-wide surveys that aimed to detect AD-related methylation changes in blood have been conducted using Illumina HM450K and 850K (MethylationEPIC) arrays; however, AD-associated hypomethylation was not detected in *CR1*, *CLU*, or *PICALM* [8, 36, 37]. This discrepancy may be explained by the modest mean differences detected in *CR1*, *CLU*, and *PICALM*, which were all less than 10%, because depending on the probes, these

arrays are not sufficiently sensitive to detect small differences, particularly when sample numbers are limited [12, 13].

It is reasonable to assume that methylation changes relevant to AD, if any, occur in the brain, a prominently affected tissue, and that blood methylation changes most likely reflect peripheral responses to the disorder rather than causally related variations. Thus, methylation changes in blood have been searched for as surrogate markers for brain methylation changes in most cases. Contrary to expectations, co-methylation changes that occur in both the blood and brains of AD patients were shown to be limited to a subset of DNA methylation sites [8, 36]. Methylation changes in the brains with neuropathology were not replicated in CD4+ lymphocytes [38] or in the whole blood in AD [9, 39] from the same individuals. Similarly, in silico analyses in the present study showed no relationships between methylation in blood and the brain at CpG sites in AD-associated DMRs, and the DMRs found in blood were not detected in previous studies that searched for DMRs in AD brains [9, 10]. These results were consistent with the finding that for the majority of DNA methylation sites, interindividual variations in whole blood are not a strong predictor of those Therefore, the methylation changes detected in blood in the present study in the brain [34]. appear not to occur in AD brains. On the other hand, the chance of AD-related methylation changes in three out of the six AD-linked genes in blood appeared to be more than coincidence. A previous study reported that age-related cognitive impairment was improved

by the introduction of young blood in old mice (heterochronic parabiosis) [40, 41]; therefore, epigenetic changes that confer susceptibility to AD do not necessarily appear to be confined to neuronal cells and may also occur in blood.

A possible change that concomitantly occurs with the hypomethylation of CpG island shores is an increase in the expression of genes because CpG island shore methylation is strongly and negatively related to gene expression [16]. Although it currently remains unclear whether CpG island shore hypomethylation in *CLU*, *CR1*, and *PICALM* is related to their up-regulation, the inverse correlation is true for *TREM2* in blood [27], and it may also be the case for *CLU* because the transcript levels of *CLU* were shown to be elevated in the blood of AD patients [42, 43]. Based on the long duration of this disease, even small changes in DNA methylation in AD-related genes, such as those observed in the present study, that accompany small changes in their expression may confer susceptibility to the progression of AD.

Although further studies are needed to clarify whether hypomethylation in AD-related genes in blood is causal or consequential, these DNA methylation signatures have potential as clinical biomarkers for AD. We herein demonstrated that in combination with the *APOE* genotype, *CLU* methylation in CpG island shores offered good prediction performance for the diagnosis of AD in our subjects. Since the methylation levels of *CR1*, *CLU*, and *PICALM* are not related to MMSE scores, methylation levels in the CpG island

shores of these genes in the blood of AD patients may have already been lower than those in controls before the onset of the disease. This is worth testing with samples in longitudinal studies. Furthermore, we discovered that the hypomethylation of *CLU* occurred in the blood of DLB, but not in VaD or FTD. This result suggests that hypomethylation is a disease-specific phenomenon and may reflect some shared underlying pathophysiological mechanisms between the two diseases, as suggested in a previous study [44], while these results may also be explained, in part, by possible diagnostic misclassification or the comorbidity of DLB with AD. Neurodegenerative disease-specific differential DNA methylation has also been reported for the *ANKI* gene [45].

Ideal biomarkers for AD were proposed to have sensitivity and specificity of >80% [46]; however, the diagnostic ability obtained by the combinatorial use of *CLU* methylation and the *APOE* genotype did not meet these criteria. Therefore, better DMRs for the clinical use of blood DNA in AD need to be identified. Further efforts to identify DMRs in other AD-associated genes will be a rational approach of choice. The success of combined uses of methylation rates and another variable regarding AD, such as the neuritic plaque burden [9, 10] and cognitive measure, [8, 47], may be referred to for unbiased screening of AD-related methylation markers in blood.

Acknowledgments

We thank the NCGG Biobank for providing the study materials, clinical information, and technical support.

References

- 1. Lunnon K, Mill J. Epigenetic studies in Alzheimer's disease: current findings, caveats, and considerations for future studies. American journal of medical genetics Part B, Neuropsychiatric genetics: the official publication of the International Society of Psychiatric Genetics. 2013;162b(8):789-99. Epub 2013/09/17. doi: 10.1002/ajmg.b.32201. PubMed PMID: 24038819; PubMed Central PMCID: PMCPMC3947441.
- 2. Bennett DA, Yu L, Yang J, Srivastava GP, Aubin C, De Jager PL. Epigenomics of Alzheimer's disease. Translational research: the journal of laboratory and clinical medicine. 2015;165(1):200-20. Epub 2014/06/07. doi: 10.1016/j.trsl.2014.05.006. PubMed PMID: 24905038; PubMed Central PMCID: PMCPMC4233194.
- 3. Wang Y, Karlsson R, Lampa E, Zhang Q, Hedman AK, Almgren M, et al. Epigenetic influences on aging: a longitudinal genome-wide methylation study in old Swedish twins. Epigenetics. 2018;13(9):975-87. Epub 2018/09/29. doi: 10.1080/15592294.2018.1526028. PubMed PMID: 30264654; PubMed Central PMCID: PMCPMC6284777.
- 4. Cheung P, Vallania F, Warsinske HC, Donato M, Schaffert S, Chang SE, et al. Single-Cell Chromatin Modification Profiling Reveals Increased Epigenetic Variations with Aging. Cell. 2018;173(6):1385-97.e14. Epub 2018/05/01. doi: 10.1016/j.cell.2018.03.079. PubMed PMID: 29706550; PubMed Central PMCID: PMCPMC5984186.

- Pal S, Tyler JK. Epigenetics and aging. Science advances. 2016;2(7):e1600584.
 Epub 2016/08/03. doi: 10.1126/sciadv.1600584. PubMed PMID: 27482540; PubMed Central
 PMCID: PMCPMC4966880.
- 6. Roubroeks JAY, Smith RG, van den Hove DLA, Lunnon K. Epigenetics and DNA methylomic profiling in Alzheimer's disease and other neurodegenerative diseases. Journal of neurochemistry. 2017;143(2):158-70. Epub 2017/08/15. doi: 10.1111/jnc.14148. PubMed PMID: 28805248.
- 7. Fransquet PD, Ryan J. The current status of blood epigenetic biomarkers for dementia. Critical reviews in clinical laboratory sciences. 2019;56(7):435-57. Epub 2019/07/23. doi: 10.1080/10408363.2019.1639129. PubMed PMID: 31328605.
- 8. Chouliaras L, Pishva E, Haapakoski R, Zsoldos E, Mahmood A, Filippini N, et al. Peripheral DNA methylation, cognitive decline and brain aging: pilot findings from the Whitehall II imaging study. Epigenomics. 2018;10(5):585-95. Epub 2018/04/26. doi: 10.2217/epi-2017-0132. PubMed PMID: 29692214; PubMed Central PMCID: PMCPMC6021930.
- 9. Lunnon K, Smith R, Hannon E, De Jager PL, Srivastava G, Volta M, et al. Methylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease. Nat Neurosci. 2014;17(9):1164-70. Epub 2014/08/19. doi: 10.1038/nn.3782. PubMed PMID: 25129077; PubMed Central PMCID: PMCPMC4410018.

- 10. De Jager PL, Srivastava G, Lunnon K, Burgess J, Schalkwyk LC, Yu L, et al. Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. Nat Neurosci. 2014;17(9):1156-63. Epub 2014/08/19. doi: 10.1038/nn.3786. PubMed PMID: 25129075; PubMed Central PMCID: PMCPMC4292795.
- Lord J, Cruchaga C. The epigenetic landscape of Alzheimer's disease. Nat Neurosci.
 2014;17(9):1138-40. Epub 2014/08/27. doi: 10.1038/nn.3792. PubMed PMID: 25157507;
 PubMed Central PMCID: PMCPMC5472058.
- 12. Tsai PC, Bell JT. Power and sample size estimation for epigenome-wide association scans to detect differential DNA methylation. International journal of epidemiology.

 2015;44(4):1429-41. Epub 2015/05/15. doi: 10.1093/ije/dyv041. PubMed PMID: 25972603;

 PubMed Central PMCID: PMCPMC4588864.
- 13. Mansell G, Gorrie-Stone TJ, Bao Y, Kumari M, Schalkwyk LS, Mill J, et al. Guidance for DNA methylation studies: statistical insights from the Illumina EPIC array. BMC genomics. 2019;20(1):366. Epub 2019/05/16. doi: 10.1186/s12864-019-5761-7. PubMed PMID: 31088362; PubMed Central PMCID: PMCPMC6518823.
- 14. Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML, et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nat Genet. 2009;41(10):1088-93. Epub 2009/09/08. doi: 10.1038/ng.440. PubMed PMID: 19734902; PubMed Central PMCID: PMCPMC2845877.

- 15. Lambert JC, Heath S, Even G, Campion D, Sleegers K, Hiltunen M, et al.

 Genome-wide association study identifies variants at CLU and CR1 associated with

 Alzheimer's disease. Nat Genet. 2009;41(10):1094-9. Epub 2009/09/08. doi: 10.1038/ng.439.

 PubMed PMID: 19734903.
- 16. Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, et al. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nature Genetics. 2009;41(2). doi: 10.1038/ng.298. PubMed PMID: WOS:000263091300017.
- 17. Doi A, Park IH, Wen B, Murakami P, Aryee MJ, Irizarry R, et al. Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. Nat Genet. 2009;41(12):1350-3. Epub 2009/11/03. doi: 10.1038/ng.471. PubMed PMID: 19881528; PubMed Central PMCID: PMCPMC2958040.
- 18. Shimoda N, Izawa T, Yoshizawa A, Yokoi H, Kikuchi Y, Hashimoto N. Decrease in cytosine methylation at CpG island shores and increase in DNA fragmentation during zebrafish aging. Age (Dordrecht, Netherlands). 2014;36(1):103-15. Epub 2013/06/06. doi: 10.1007/s11357-013-9548-5. PubMed PMID: 23736955; PubMed Central PMCID: PMCPMC3889898.
- 19. Huynh JL, Garg P, Thin TH, Yoo S, Dutta R, Trapp BD, et al. Epigenome-wide

differences in pathology-free regions of multiple sclerosis-affected brains. Nat Neurosci. 2014;17(1):121-30. Epub 2013/11/26. doi: 10.1038/nn.3588. PubMed PMID: 24270187; PubMed Central PMCID: PMCPMC3934491.

- 20. Humphries C, Kohli MA, Whitehead P, Mash DC, Pericak-Vance MA, Gilbert J. Alzheimer disease (AD) specific transcription, DNA methylation and splicing in twenty AD associated loci. Molecular and cellular neurosciences. 2015;67:37-45. Epub 2015/05/26. doi: 10.1016/j.mcn.2015.05.003. PubMed PMID: 26004081; PubMed Central PMCID: PMCPMC4540636.
- 21. Saxonov S, Berg P, Brutlag DL. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proc Natl Acad Sci U S A. 2006;103(5):1412-7. Epub 2006/01/25. doi: 10.1073/pnas.0510310103. PubMed PMID: 16432200; PubMed Central PMCID: PMCPMC1345710.
- 22. Li LC, Dahiya R. MethPrimer: designing primers for methylation PCRs. Bioinformatics (Oxford, England). 2002;18(11):1427-31. Epub 2002/11/09. doi: 10.1093/bioinformatics/18.11.1427. PubMed PMID: 12424112.
- 23. Kumaki Y, Oda M, Okano M. QUMA: quantification tool for methylation analysis.
 Nucleic acids research. 2008;36(Web Server issue):W170-5. Epub 2008/05/20. doi:
 10.1093/nar/gkn294. PubMed PMID: 18487274; PubMed Central PMCID:
 PMCPMC2447804.

- 24. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach.

 Biometrics. 1988;44(3):837-45. Epub 1988/09/01. PubMed PMID: 3203132.
- 25. Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE. Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. Nat Genet. 2007;39(1):17-23. Epub 2006/12/29. doi: 10.1038/ng1934. PubMed PMID: 17192785.
- 26. Paik YK, Chang DJ, Reardon CA, Davies GE, Mahley RW, Taylor JM. Nucleotide sequence and structure of the human apolipoprotein E gene. Proc Natl Acad Sci U S A. 1985;82(10):3445-9. Epub 1985/05/01. doi: 10.1073/pnas.82.10.3445. PubMed PMID: 2987927; PubMed Central PMCID: PMCPMC397792.
- 27. Ozaki Y, Yoshino Y, Yamazaki K, Sao T, Mori Y, Ochi S, et al. DNA methylation changes at TREM2 intron 1 and TREM2 mRNA expression in patients with Alzheimer's disease. Journal of psychiatric research. 2017;92:74-80. Epub 2017/04/17. doi: 10.1016/j.jpsychires.2017.04.003. PubMed PMID: 28412600.
- 28. Kerkel K, Spadola A, Yuan E, Kosek J, Jiang L, Hod E, et al. Genomic surveys by methylation-sensitive SNP analysis identify sequence-dependent allele-specific DNA methylation. Nat Genet. 2008;40(7):904-8. Epub 2008/06/24. doi: 10.1038/ng.174. PubMed PMID: 18568024.
- 29. Liu Y, Aryee MJ, Padyukov L, Fallin MD, Hesselberg E, Runarsson A, et al.

Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. Nature biotechnology. 2013;31(2):142-7. Epub 2013/01/22. doi: 10.1038/nbt.2487. PubMed PMID: 23334450; PubMed Central PMCID: PMCPMC3598632.

- 30. Chibnik LB, Yu L, Eaton ML, Srivastava G, Schneider JA, Kellis M, et al. Alzheimer's loci: epigenetic associations and interaction with genetic factors. Annals of clinical and translational neurology. 2015;2(6):636-47. Epub 2015/07/01. doi: 10.1002/acn3.201. PubMed PMID: 26125039; PubMed Central PMCID: PMCPMC4479524.
- 31. Sanchez-Mut JV, Heyn H, Silva BA, Dixsaut L, Garcia-Esparcia P, Vidal E, et al. PM20D1 is a quantitative trait locus associated with Alzheimer's disease. Nature medicine. 2018;24(5):598-603. Epub 2018/05/08. doi: 10.1038/s41591-018-0013-y. PubMed PMID: 29736028.
- 32. Watson CT, Roussos P, Garg P, Ho DJ, Azam N, Katsel PL, et al. Genome-wide DNA methylation profiling in the superior temporal gyrus reveals epigenetic signatures associated with Alzheimer's disease. Genome Med. 2016;8(1):5. Epub 2016/01/26. doi: 10.1186/s13073-015-0258-8. PubMed PMID: 26803900; PubMed Central PMCID: PMCPMC4719699.
- 33. Edgar RD, Jones MJ, Meaney MJ, Turecki G, Kobor MS. BECon: a tool for interpreting DNA methylation findings from blood in the context of brain. Transl Psychiatry. 2017;7(8):e1187. Epub 2017/08/02. doi: 10.1038/tp.2017.171. PubMed PMID: 28763057;

PubMed Central PMCID: PMCPMC5611738.

- 34. Hannon E, Lunnon K, Schalkwyk L, Mill J. Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes. Epigenetics. 2015;10(11):1024-32. Epub 2015/10/13. doi: 10.1080/15592294.2015.1100786. PubMed PMID: 26457534; PubMed Central PMCID: PMCPMC4844197.
- 35. Yu L, Chibnik LB, Srivastava GP, Pochet N, Yang J, Xu J, et al. Association of Brain DNA methylation in SORL1, ABCA7, HLA-DRB5, SLC24A4, and BIN1 with pathological diagnosis of Alzheimer disease. JAMA neurology. 2015;72(1):15-24. Epub 2014/11/05. doi: 10.1001/jamaneurol.2014.3049. PubMed PMID: 25365775; PubMed Central PMCID: PMCPMC4344367.
- 36. Lunnon K, Smith RG, Cooper I, Greenbaum L, Mill J, Beeri MS. Blood methylomic signatures of presymptomatic dementia in elderly subjects with type 2 diabetes mellitus. Neurobiol Aging. 2015;36(3):1600.e1-4. Epub 2015/01/15. doi: 10.1016/j.neurobiolaging.2014.12.023. PubMed PMID: 25585531; PubMed Central PMCID: PMCPMC5747361.
- 37. Madrid A, Hogan KJ, Papale LA, Clark LR, Asthana S, Johnson SC, et al. DNA Hypomethylation in Blood Links B3GALT4 and ZADH2 to Alzheimer's Disease. Journal of Alzheimer's disease: JAD. 2018;66(3):927-34. Epub 2018/10/30. doi: 10.3233/jad-180592.

PubMed PMID: 30372681; PubMed Central PMCID: PMCPMC6478383.

- 38. Yu L, Chibnik LB, Yang J, McCabe C, Xu J, Schneider JA, et al. Methylation profiles in peripheral blood CD4+ lymphocytes versus brain: The relation to Alzheimer's disease pathology. Alzheimer's & dementia: the journal of the Alzheimer's Association. 2016;12(9):942-51. Epub 2016/03/27. doi: 10.1016/j.jalz.2016.02.009. PubMed PMID: 27016692; PubMed Central PMCID: PMCPMC5014706.
- 39. Smith RG, Hannon E, De Jager PL, Chibnik L, Lott SJ, Condliffe D, et al. Elevated DNA methylation across a 48-kb region spanning the HOXA gene cluster is associated with Alzheimer's disease neuropathology. Alzheimer's & dementia: the journal of the Alzheimer's Association. 2018;14(12):1580-8. Epub 2018/03/20. doi: 10.1016/j.jalz.2018.01.017. PubMed PMID: 29550519; PubMed Central PMCID: PMCPMC6438205.
- 40. Katsimpardi L, Litterman NK, Schein PA, Miller CM, Loffredo FS, Wojtkiewicz GR, et al. Vascular and neurogenic rejuvenation of the aging mouse brain by young systemic factors. Science (New York, NY). 2014;344(6184):630-4. Epub 2014/05/07. doi: 10.1126/science.1251141. PubMed PMID: 24797482; PubMed Central PMCID: PMCPMC4123747.
- 41. Villeda SA, Plambeck KE, Middeldorp J, Castellano JM, Mosher KI, Luo J, et al. Young blood reverses age-related impairments in cognitive function and synaptic plasticity in mice. Nature medicine. 2014;20(6):659-63. Epub 2014/05/06. doi: 10.1038/nm.3569.

PubMed PMID: 24793238; PubMed Central PMCID: PMCPMC4224436.

- 42. Thambisetty M, Simmons A, Velayudhan L, Hye A, Campbell J, Zhang Y, et al. Association of plasma clusterin concentration with severity, pathology, and progression in Alzheimer disease. Archives of general psychiatry. 2010;67(7):739-48. Epub 2010/07/07. doi: 10.1001/archgenpsychiatry.2010.78. PubMed PMID: 20603455; PubMed Central PMCID: PMCPMC3111021.
- 43. Xing YY, Yu JT, Cui WZ, Zhong XL, Wu ZC, Zhang Q, et al. Blood clusterin levels, rs9331888 polymorphism, and the risk of Alzheimer's disease. Journal of Alzheimer's disease: JAD. 2012;29(3):515-9. Epub 2012/01/20. doi: 10.3233/jad-2011-111844. PubMed PMID: 22258514.
- Chung EJ, Babulal GM, Monsell SE, Cairns NJ, Roe CM, Morris JC. Clinical
 Features of Alzheimer Disease With and Without Lewy Bodies. JAMA neurology.
 2015;72(7):789-96. Epub 2015/05/20. doi: 10.1001/jamaneurol.2015.0606. PubMed PMID:
 25985321; PubMed Central PMCID: PMCPMC4501861.
- 45. Smith AR, Smith RG, Burrage J, Troakes C, Al-Sarraj S, Kalaria RN, et al. A cross-brain regions study of ANK1 DNA methylation in different neurodegenerative diseases.

 Neurobiol Aging. 2019;74:70-6. Epub 2018/11/16. doi: 10.1016/j.neurobiolaging.2018.09.024.

 PubMed PMID: 30439595.
- 46. The Ronald and Nancy Reagan Research Institute of the Alzheimer's Association

and the National Institute on Aging Working Group. Consensus report of the Working Group on: "Molecular and Biochemical Markers of Alzheimer's Disease". . Neurobiol Aging. 1998;19(2):109-16. Epub 1998/04/29. PubMed PMID: 9558143.

47. Starnawska A, Tan Q, McGue M, Mors O, Borglum AD, Christensen K, et al.

Epigenome-Wide Association Study of Cognitive Functioning in Middle-Aged Monozygotic

Twins. Frontiers in aging neuroscience. 2017;9:413. Epub 2018/01/10. doi:

10.3389/fnagi.2017.00413. PubMed PMID: 29311901; PubMed Central PMCID:

PMCPMC5733014.

Supporting information

S1 Fig. Relative positions of CpG dinucleotides selected for the methylation analysis of

the CR1 gene.

The schematic shows the distribution of CpG dinucleotides along the CRI gene. The position of the transcription start site is defined as +1; hence, "-2,000" in the parenthesis indicates that the diagram includes 2 kb upstream of the CRI gene, while "146,501" in the parenthesis is the position 2 kb downstream from the end of the last exon. Vertical lines indicate the positions of CpG dinucleotides. Open rectangles depict exons, and a region spanning the first exon is enlarged. The fundamentals of these diagrams were automatically drawn by the web-based tool, CyGnusPlotter. It collects the genomic structure of the most representative isoform of a requested gene from the Ensembl database with upstream and downstream regions of the designated lengths. Underlined was the region for which the methylation level was analyzed by bisulfite sequencing. It included a CpG to which Illumina designed a probe for the 450k array and gave the ID "cg14726637". The position of the AD-related SNP, rs3818361, was also shown. The thick horizontal line represents the

S2 Fig. Relative positions of CpG dinucleotides selected for the methylation analysis of the CLU gene.

39

position of amplicons for bisulfite sequencing.

The figure is drawn similar to that in S1 Fig, except for the dashed line, which indicates the region for which methylation levels were quantified by pyrosequencing.

S3 Fig. Relative positions of CpG dinucleotides selected for the methylation analysis of the *PICALM* gene.

The figure is drawn similar to that in S1 Fig. The AD-related risk SNP, rs3851179, which is linked to the gene, is not shown in the figure because it is located ~80 kb upstream of the gene.

S4 Fig. Relative positions of CpG dinucleotides selected for the methylation analysis of the ABCA7 gene.

The figure is drawn similar to that in S1 Fig.

S5 Fig. Relative positions of CpG dinucleotides selected for the methylation analysis of the *BIN1* gene.

The figure is drawn similar to that in S1 Fig. In this gene, there are two regions (I and II) for which methylation levels were quantified by bisulfite sequencing.

40

S6 Fig. Relative positions of CpG dinucleotides selected for the methylation analysis of

the TREM2 gene.

The figure is drawn similar to that in S1 Fig, except that the underlined region was the region

for which methylation levels were analyzed by pyrosequencing.

S7 Fig. In silico examination of blood-brain methylation concordance by the online tool,

BECon.

Each plot shows the inter-individual variability of the methylation level at a CpG site across

16 subjects. The five CpGs examined were derived from three AD-related DMRs identified

in the present study: one CpG in CRI (A), three CpG sites in CLU (B), and one CpG site in

PICALM (C). The numbers in "Correlation" columns in the tables indicate Spearman's rank

correlation coefficients (r_S or ρ), which were obtained by comparisons of methylation levels

between blood and either one of three different cortical regions (Broadmann area 10 (BA10),

prefrontal cortex; Broadmann area 7 (BA7), parietal cortex; and Broadmann area 20 (BA20),

temporal cortex).

S8 Fig. In silico examination of blood-brain methylation concordance by the Blood

Brain DNA Methylation Comparison Tool.

Methylation levels at the five CpGs in S7 Fig, a CpG site in CR1 (A), three CpG sites in CLU

41

(B), and a CpG site in PICALM (C), in blood and four brain regions (PFC, prefrontal cortex;

EC, entorhinal cortex; STG, superior temporal gyrus; CER, cerebellum) from the same

individual donors were plotted in the rectangles, and the correlation of DNA methylation in

blood with the four brain regions are plotted in square boxes. Methylation data were

generated by Hannon et al [34].

S1 Table. Primers and PCR conditions for bisulfite genomic sequencing.

S2 Table. Association of known AD-related SNPs with AD in NCGG samples (test and

replication groups).

S3 Table. Results of binominal regression analyses

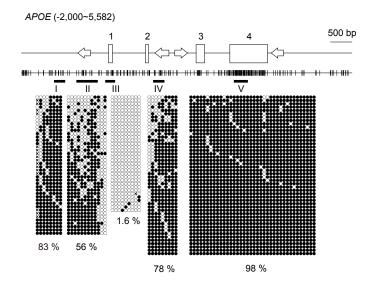


Fig. 1 Mitsumori et al.

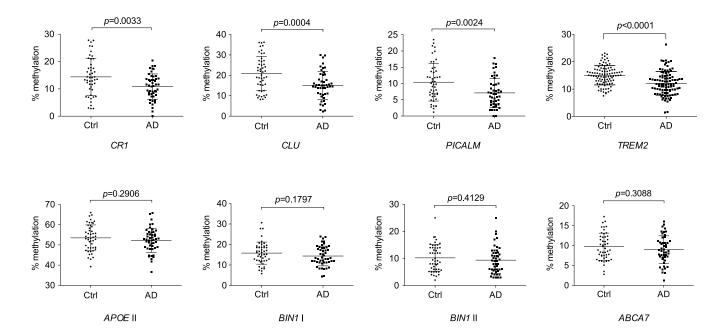


Fig. 2 Mitsumori et al.

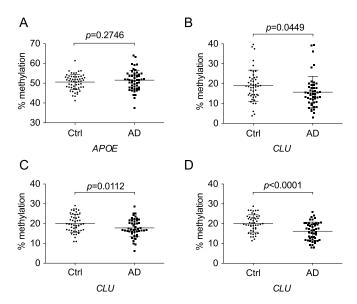


Fig. 3 Mitsumori et al.

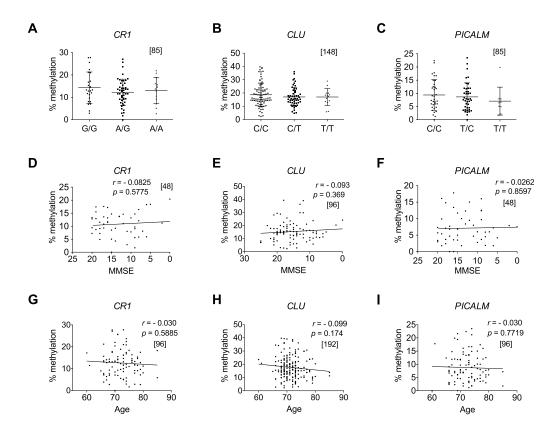


Fig. 4 Mitsumori et al.

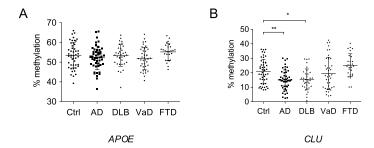


Fig. 5 Mitsumori et al.

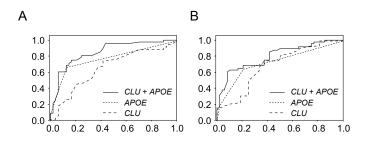


Fig. 6 Mitsumori et al.