Published in final edited form as:

Neurobiol Aging. 2019 April ; 76: 35–44. doi:10.1016/j.neurobiolaging.2018.12.003.

DNA methylation variability in Alzheimer's Disease

Zhiguang Huo¹, Yun Zhu², Lei Yu³, Jingyun Yang³, Philip De Jager⁴, David A. Bennett³, and Jinying Zhao²

¹Department of Biostatistics, University of Florida, Gainesville, FL, United States

²Department of Epidemiology, University of Florida, Gainesville, FL, United States

³Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, IL, United States

⁴Department of Neurology, College of Physicians & Surgeons, Columbia University, NY, United States

Abstract

DNA methylation plays a critical role in brain aging and AD. While prior studies have largely focused on testing mean DNA methylation, DNA methylation instability (quantified by DNA methylation variability) may also affect disease susceptibility. Using DNA methylation data collected by the Religious Orders Study and the Rush Memory and Aging Project (ROSMAP), we identified 249 and 115 variably methylated probes (VMPs) associated with amyloid- β and neurofibrillary tangles, respectively. These VMPs clustered into 133 and 14 regions, respectively. Notably, we found that majority of these VMPs did not overlap with differentially methylated probes (DMPs), indicating that VMPs and DMPs may capture different sets of genes associated with AD pathology. Overall, our results demonstrated that DNA methylation instability affects AD neuropathology, and highlights the importance of testing methylation variability in epigenetic research.

Corresponding Author: Jinying Zhao, MD, PhD, Professor, Department of Epidemiology, College of Public Health and Health Professions & College of Medicine, University of Florida, 2004 Mowry Road, Gainesville, Florida, 32610, Tel: 352-273-5933, Fax: 352-273-5365, jzhao66@ufl.edu.

Conflicts of interest: The authors have no conflicts of interest to declare.

1. All authors have disclosed that:

- (a) All authors do not have potential conflicts of interest, including any financial, personal or other relationships with other people or organizations within three years of beginning the work submitted that could inappropriately influence (bias) their work.
- (b) The participating institutions have no contracts relating to this research through which it or any other organization may stand to gain financially now or in the future.
- (c) There are no other agreements with all institutions that could be seen as involving a financial interest in this work.

3. The data presented in this manuscript have not been published or are being considered for publication elsewhere in whole or part, in any language, except as an abstract. We will not submit the manuscript to anywhere else while it is being considered by *Neurobiology of Aging*.

4. All authors have reviewed the contents of the manuscript being submitted, approve of its contents and validate the accuracy of the data.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Introduction

Alzheimer's dementia is a devastating neurodegenerative disorder affecting over 35 million people worldwide, and this number is expected to nearly triple by 2050 (Duthey, 2013). Accumulation of extracellular amyloid- β and intraneuronal neurofibrillary tangles (NFTs) are the two hallmarks of Alzheimer's disease (AD). DNA methylation plays an important role in regulating gene expression, and altered DNA methylation has been implicated in AD pathology (Irier and Jin, 2012). Several epigenome-wide association studies (EWAS) (De Jager et al., 2014; Lunnon et al., 2014; Smith et al., 2018; Watson et al., 2016) identified differentially methylated probes (DMPs) and differentially methylated regions (DMRs) associated with AD. While previous studies have largely focused on testing the difference in mean DNA methylation level between patients and controls. Recent evidence suggests that DNA methylation instability, may also affect disease susceptibility (Palumbo et al., 2018). For example, several recent studies found that increased DNA methylation variability in multiple genes was associated with cancers (Hansen et al., 2011), type 1 diabetes (Paul et al., 2016), aging (Jones et al., 2015), and rheumatoid arthritis (Webster et al., 2018). This evidence suggests that, in addition to the mean change in DNA methylation, altered DNA methylation variability or instability of DNA methylation may also play an important role in disease pathophysiology. To date, we are not aware of any study that examined the role of DNA methylation variability in AD neuropathology. Using existing DNA methylation and gene expression data generated in postmortem prefrontal cortex of older individuals participating in two community-based population cohorts of aging and dementia (the Religious Orders Study and the Rush Memory and Aging Project [ROSMAP]) (A Bennett et al., 2012b), we conducted analyses to examine whether altered DNA methylation stability contributes to AD pathology by identifying variably methylated probes (VMPs) and variably methylated regions (VMRs) in the postmortem brain tissue. For comparison purpose, we also examined the potential overlap between variably methylated genes (e.g., genes showing altered methylation variability) and differentially methylation genes (e.g., genes showing altered mean DNA methylation level) at a genome scale. Moreover, we assessed the impact of DNA methylation instability on gene expression profiled on the same brain cortex.

Materials and Methods

Study population.

This study included deceased participants from two ongoing, prospective studies of brain aging and dementia in older individuals, as described below. Detailed study design and assessment methods were described previously (A Bennett et al., 2012b, 2012a; Bennett et al., 2006). Both studies were approved by the Institutional Review Board of the Rush University Medical Center. Data are available for sharing at www.radc.rush.edu. Clinical characteristics of the study participants at the time of death are shown in Table 1. The mean (SD) age of brain donors was 88.0 (6.7) years at the time of death. Men account for 37% of the study participants, and AD accounts for 42%.

Religious Orders Study (ROS): Initiated in 1994, ROS enrolled older Catholic priests, nuns and brothers from across the USA free of known dementia at time of enrollment.

Participants agreed to annual clinical evaluations including standardized neurological examination and neuropsychological testing and signed both an informed consent and an Anatomic Gift Act donating their brains at time of death (A Bennett et al., 2012a; Bennett et al., 2006). Both the follow-up rate of survivors and the autopsy rate exceed 90% (autopsies of deaths).

Rush Memory and Aging Project (MAP): Established in 1997, MAP consists of older men and women from across the Chicagoland area, without known dementia at enrollment. Participants agreed to annual clinical evaluations and signed both an informed consent and an Anatomic Gift Act donating their brains, spinal cords and selected nerves and muscles at the time of death (A Bennett et al., 2012b; Bennett et al., 2006). The follow-up rate for survivors exceeds 90% and the autopsy rate for deceased subjects exceeds 80% (autopsies of deaths).

Assessment of neuropathological phenotypes.

We collected comprehensive neuropathological phenotypes for brain aging and AD in ROSMAP by the same investigators using a large common core of testing batteries. Thus, data can be efficiently merged for joint analyses. Detailed methods for clinical phenotyping have been described previously (Bennett et al., 2009, 2006). The quantitative AD pathologic burden (our primary outcomes) include the overall amyloid- β load based on percent area occupied with image analysis, and the density of PHF (paired helical filament) tau tangles assessed with stereology, which were identified by molecularly-specific immunohistochemistry and quantified by imaging analysis. These two composite summary measures were obtained by averaging the percent area occupied by amyloid- β or tangles across 8 regions, as previously described (Bennett et al., 2004). Neuropathologic examinations were made by board-certified neuropathologists blinded to the clinical data.

Assessment of cognitive performance.

During each annual clinical visit, participants received evaluations of their cognition functions via a battery of 21 cognitive performance tests across a range of cognitive abilities. Nineteen of them were used to construct a global composite measure of cognitive function, and summarized as 5 cognitive domains including episode memory, working memory, semantic memory, perceptual speed, and visuospatial ability. Detailed methods for the assessment of cognitive functions in ROSMAP cohorts have been described previously (A Bennett et al., 2012a; Wilson et al., 2015).

DNA methylation data pre-processing.

DNA methylation in the dorsolateral prefrontal cortex (DLPFC) was measured using Infinium HumanMethylation450 BeadChip as previously described (De Jager et al., 2014). Data pre-processing, QC and normalization were performed using the RnBeads software (Assenov et al., 2014) with default options unless otherwise stated. Of the 732 subjects and 474,813 autosomal CpG probes, we removed probes and/or subjects based on the following criteria: (1) samples with low bisulfite conversion rate (De Jager et al., 2014) ($n=21$); (2) outliers determined by PCA analysis (De Jager et al., 2014) ($n=5$); (3) 14,018 probes located within ± 5 bp of known SNPs ($MAF > 0.05$); (4) 5,473 probes using SNP criterion "3",

10,588 probes with high median detection p -values ($p > 0.05$), and 2,804 probes with non-acceptable context based on the *RnBeads* default options. Data normalization was performed using the “*swan*” method implemented in the *RnBeads* software. Our final data analysis included a total of 441,930 probes and 706 subjects. Missing values were computed by the *KNN* algorithm with $k = 100$ using R package *impute* (Hastie et al., 2001) from Bioconductor.

Identifying VMPs/VMRs associated with AD neuropathology.

To identify VMPs associated with AD pathology, we employed the *DiffVar* function (Phipson and Oshlack, 2014) implemented in the *missMethyl* package (Phipson et al., 2015). In this analysis, we adjusted for age at death, sex, batch, study (ROS or MAP), and education level. Given the high correlation of DNA methylation between adjacent CpG sites, disease-related CpGs can be clustered in genomic regions, we thus also performed region-based analysis to identify variably methylated regions (VMRs) associated with amyloid- β tangles, separately, using a 5kb sliding window approach. Statistical significance of a VMR was obtained by combining p -values (from *DiffVar*) of all CpG probes within each 5kb window using the *Brown's* method (Brown, 1975) implemented in the R package *empiricalBrownMethod* (Poole et al., 2016), which takes into account the correlation between adjacent probes. Multiple testing was controlled by the false discovery rate (FDR) (Benjamini and Hochberg, 1995) and a q -value < 0.05 was used to determine statistical significance for both VMPs and VMRs analyses. A more stringent genome-wide estimate of significance level $= 3.6 \times 10^{-8}$ (Saffari et al., 2018) was adopted in the Manhattan plot in addition to the q -value < 0.05 criteria.

VMP/VMRs associated with cognitive performance.

To identify VMPs/VMRs associated with cognitive performance, we employed the same statistical methods with adjustments for same covariates. The cognitive performance was quantified by the composite measure of cognitive functions proximate to death, including global cognitive function, episode memory, working memory, semantic memory, perceptual speed, or visuospatial ability.

Sensitivity analyses.

To examine whether and how non-linear effect of age affects methylation variation, we additionally adjusted for age^2 and age^3 (age at death). To examine whether neuronal cell proportions influence our results, we first estimated the NeuN+ cells (primary neurons) proportion using the *projectCellType* function in the *minfi* R package (Aryee et al., 2014) based on the DLPFC cell epigenotype reference database (Guintivano et al., 2013), and then additionally adjusted for neurons proportions in the statistical models. To examine whether population substructure affects our results, we calculated the genomic inflation factors using the R package *QQperm* (Yang et al., 2011).

Functional annotation and validation.

To explore the potential functions of the identified VMPs/VMRs, we first annotated them to genomic features using the *GenomicFeatures* package in Bioconductor (Lawrence et al., 2013). We then tested the association between gene expression (RNA-seq in same brain

cortex) of VMPs and AD neuropathology using the *limma* function in R (Smyth, 2005). We focused on *cis*-regulation (± 5 kb of a tested VMP) and adjusted for age at death, sex, batch, study (ROS or MAP) and education level. Functional enrichment analysis was conducted using the Kolmogorov Smirnov test (*KS*-test) (Stephens, 1974), a non-parametric approach that doesn't require an arbitrary cutoff to claim significance. In this analysis, we considered two pathway databases: (1) curated genes based on the Molecular Signatures Database v6.1 (<http://software.broadinstitute.org/gsea/msigdb/index.jsp>), including Biocarta, KEGG, Reactome, and GO gene sets; and (2) previously published GWAS loci (a.k.a. GWAS catalog (Welter et al., 2013) and genes/ proteins known to be involved in AD pathology (Campion et al., 2016). Pathways with less than 10 or greater than 400 genes were excluded from the enrichment analysis. Multiple testing was corrected for total number of pathways using FDR, and $q < 0.05$ was considered statistical significance.

Co-methylation networks.

To examine whether genes that are differentially variably methylated in relation to AD neuropathology are co-methylated, we conducted the Weighted Gene Correlation Network Analysis (WGCNA) (Langfelder and Horvath, 2008). This analysis included a total of 1,145 or 967 VMPs showing nominal associations ($p < 0.001$) with amyloid- β or tangles, respectively. Co-methylation modules were constructed separately among subjects with high (upper tertile) vs low (bottom tertile) pathological burden for each measure. Differential networks – modules vary by pathological burden, were identified by comparing the two groups via preservation analysis in the WGCNA. Hub genes within each co-methylation modules were detected using the *ARACNE* algorithm (Margolin et al., 2006) in the R package *minet* (Meyer et al., 2008). Network visualization was done using the R package *igraph* (Csardi and Nepusz, 2006).

Overlapping between VMPs and DMPs.

To examine whether and to what extent the identified VMPs overlapped with DMPs in relation to AD pathology, we performed differential methylation analysis to identify DMPs associated with AD neuropathology using R package *limma* (Smyth, 2005), adjusting for same covariates as what used in the VMP analysis.

Replication.

To replicate our findings in the ROSMAP, we downloaded DNA methylation data from the publicly available dataset GSE80970 ($n=142$, mean age 85.7, 38% men). The data contains DNA methylation data (450K array) in dorsolateral prefrontal cortex (same brain region as that in ROSMAP) and clinical information for Braak (a semi-quantitative measure for tangle), age, and gender. Detailed information for this dataset has been described previously (Smith et al., 2018). As the data did not include phenotype information for amyloid, we focused on replication of tangle-related VMPs only.

Results

VMPs/VMRs associated with AD neuropathology.

At the level of $q < 0.05$, we identified 249 VMPs (223 hypervariable (i.e., greater variability with more neuropathology), 26 hypovariable (i.e., less variability with more neuropathology) associated with amyloid- β load (Figure 1A). By contrast, 115 VMPs (48 hypervariable, 67 hypovariable) were associated with tangles (Figure 1B). Tables 2 and 3 list the top 50 most significant VMPs associated with amyloid- β and tangles, respectively. The identified VMPs are clustered into 133 and 14 VMRs for amyloid- β and tangles, respectively. A full list of the VMPs/VMRs associated with amyloid- β and tangles is shown in the online Tables S1 and S2, respectively.

VMPs associated with cognitive performance.

We identified 1 VMP associated with global cognitive function, 6 VMPs associated with semantic memory, and 14 VMPs with working memory at FDR 5% (Tables S3).

Sensitivity analyses.

After additional adjustment for non-linear age effect (Table S4) or cell proportions (Tables S5), the identified VMPs remain to be significant, indicating that the observed associations of VMPs with AD pathology are unlikely to be confounded by these factors. In addition, the genomic inflation factors were 1.02 for tangles and 1.06 for amyloid- β , indicating that population stratification did not confound our results either.

Genomic distribution of the identified VMPs.

To explore the potential functional impact of the VMPs on transcriptional activities, we annotated the identified VMPs to pre-determined genomic features (Figure 2). Compared to the null distribution of CpG probes included in the Illumina HumanMethylation450 array, VMPs associated with amyloid- β were significantly enriched in introns (15% vs 21%, $p < 0.001$) and intergenic regions (57% vs 66%, $p < 0.001$), but depleted in promoter regions (23% vs 8.4%, $p < 0.001$). By contrast, VMPs associated with tangles were over-represented in promoter regions (23% vs 31%, $p = 0.001$), but depleted in intergenic regions (57% vs 51%, $p = 0.001$). Figure S3 shows the distributions of methylation variability status (hyper- vs hypo-variable) for the identified VMPs with respect to genomic locations (promoters, intergenic regions, introns, and exons) or/and neuropathological features (tangle vs amyloid). It shows that: (1) a large proportion of tangle-related VMPs are hypo-variably methylated (58% hypo vs 42% hyper), and these VMPs are more likely to be located in promoters (36% vs. 23%), and less likely to be located in exons (1% vs. 5%), introns (10% vs. 15%), and intergenic regions (52% vs. 57%), as compared to the null distribution of the CpG probes in the Illumina HumanMethylation450 array. (2) Majority of the amyloid-related VMPs are hyper-variable (90% hyper vs 10% hypo), and these VMPs are more likely to be located in intergenic regions (70% vs. 57%), followed by introns (21% vs. 15%), exons (1% vs. 5%), and promoters (9% vs. 23%), after adjusting for the null distribution of the CpG probes on the chip. To further examine whether this differential association (i.e., tangle-related VMPs are more likely to be hypovariable whereas amyloid-related VMPs are

more likely to be hypervariable) is confounded by genomic locations, we fitted a logistic regression model, in which methylation variability status (hyper- or hypo-variable, y/n) was the dependent variable, and a neuropathological trait of interest (tangle or amyloid) was the independent variable, adjusting for genomic location. We found that the observed differential association of DNA methylation variability status with AD neuropathology is independent of genomic location of the probes (p -value = 1.32×10^{-16}). These findings support a role of DNA methylation instability in AD pathology, and demonstrated that DNA methylation variability is trait-specific. The differential genomic distributions for amyloid-related and tangle-related VMPs also suggest that these two hallmarks of AD are etiologically heterogeneous.

Functional validation by RNA-seq.

Using RNA-seq data profiled from the same brain neocortical regions, we examined whether the identified VMPs/VMRs affected gene expression. We found that, of the 115 VMPs (annotated to 70 unique genes) associated with tangles, 35 genes were also differentially expressed ($q < 0.05$), whereas of the 249 VMPs (annotated to 113 unique genes) associated with amyloid- β , only 2 genes showed differential expression ($q < 0.05$). Genes showing both differential variability and differential expression are listed in the Table S6. While these results demonstrated that the identified VMPs might affect gene expression, the seemingly larger effect of tangle-related VMPs, as compared to amyloid- β -related VMPs, on gene expression deserves further investigation.

Enrichment analysis.

Pathway enrichment analysis revealed that the identified VMPs/VMRs associated with AD neuropathology were over-represented in biological processes related to neuron differentiation, neuron projection development, calcium ion transmembrane transport, positive regulation of nervous system development, synaptic membrane, and neuronal postsynaptic density. Moreover, the putative VMR genes were enriched in GWAS loci that were previously implicated in AD (Beecham et al., 2014; Campion et al., 2016; Jun et al., 2016; Sherva et al., 2014). Table 4 shows the top enriched pathways. A full list of the enriched pathways is shown in Table S7. GWAS loci enrichment is shown in Table S8.

Co-methylation networks.

Our network analysis identified four co-methylated modules for VMPs/VMRs associated with amyloid- β (Figure S1A) and four modules for those associated with tangles (Figure S1B). The network structure of two modules associated with amyloid- β (brown and yellow) was not preserved, suggesting potential differential networks between patients with high (upper tertile) vs low (bottom tertile) burden for AD pathology (Figure S2).

Overlapping between VMPs and DMPs.

To compare whether and to what extent the identified VMPs overlap with DMPs, we conducted differential methylation analyses and identified 415 and 3,859 DMPs associated with amyloid- β and tangles, respectively (Table S9). Of these, 17.7% of the amyloid- β -related VMPs were also DMPs, whereas 26.1% of the tangle-related VMPs overlapped with

DMPs. These results demonstrated that majority of the AD-related VMPs were not overlapped with DMPs, implying that VMPs and DMPs may capture different sets of genes, and thus may reflect different aspects of AD pathology. In addition, in order to test whether the overlapping between VMPs and DMPs could be more or less than would expect by chance, we conducted the Fisher's exact test. We obtained highly significant p -values for both neuropathological phenotypes (2.17×10^{-80} for amyloid- β , 7.52×10^{-32} for tangles), suggesting that the observed overlapping between VMPs and DMPs should most likely represent true biology rather than chance alone.

External replication of tangle-related VMPs.

Of the 115 tangle-related VMPs identified in ROSMAP, 31 probes were nominally associated with Braak score at a raw p -value < 0.05 after adjusting for age at death and gender. Of these, 29 probes (94%) had the same direction (hyper-/hypo-variably methylated) as that in the ROSMAP. However, only 4 probes passed multiple testing at $FDR < 0.05$. The lack of replication could be due to multiple reasons. First, the sample size used in the external validation was rather small compared to that used in ROSMAP. Second, the Braak score included in the publicly available dataset is different from the quantitative measure of tangles (quantified by molecularly specific immunohistochemistry across 8 brain regions) used in our analysis. Third, the publicly available dataset didn't contain information for confounding variables such as education. We were unable to replicate the association of amyloid-related VMPs due to lack of the amyloid phenotype in the external dataset.

Discussion

In two community-based population cohort studies of aging and dementia, we found that the variability in DNA methylation was related to AD neuropathology. Specifically, we identified 249 VMPs (clustered into 133 VMRs) and 115 VMPs (clustered into 14 VMRs) significantly associated with amyloid- β and tangles, respectively. The identified VMR genes were enriched in biological processes related to excitatory synapse, neuron differentiation, calcium ion transmembrane transport, positive regulation of nervous system development, and neuronal postsynaptic density. Our results support a role of DNA methylation variability in AD pathology, and highlight the importance of testing DNA methylation variability in epigenetic research.

Of the identified VMR genes, the *TUBGCP3* gene showed the strongest association with amyloid- β . This gene encodes the tubulin gamma complex associated protein 3, and is expressed in multiple brain regions. It plays an important role in microtubule nucleation, reduction of which may contribute to neuronal degeneration (Cash et al., 2003; Jean and Baas, 2013). Another VMR gene is *NGFR*, which acts as a proapoptotic receptor in neuron cell death via binding to A β in the AD brain (Lee et al., 2001). Genetic polymorphisms of this gene were associated with AD (Cozza et al., 2008). Other top-ranked VMR genes include *CCR3* (Leung et al., 2013; Villeda et al., 2011; Zhu et al., 2017), *C10orf54* (De Jager et al., 2014; Lord and Cruchaga, 2014), and *WDR81* (De Jager et al., 2014; Kauwe et al., 2014), many of which have been previously associated with cognitive impairment and/or AD pathology. The most significant VMR gene associated with tangles was *FAT1*, which is

a *fas*-binding protein that plays an important role in neurodevelopment (Menges et al., 2009), neurodegeneration (Sul et al., 2013), and tumorigenesis (Menges et al., 2009). The second gene showing strong variably methylation in relation to tangles was *LRPAP1*. This gene encodes a protein that interacts with low density lipoprotein receptor-related protein (LRP) that plays a critical role in brain A β clearance (Kanekiyo et al., 2013). Moreover, genetic polymorphisms in *LRPAP1* was associated with AD (Sánchez et al., 2001). Another VMR gene associated with tangles was *MYC*, a proto-oncogene that encodes a nuclear phosphoprotein that plays a role in cell cycle progression, apoptosis and cellular transformation (Dang, 1999). In a mouse model, neuronal expression of *MYC* causes a neurodegenerative phenotype (Lee et al., 2009). Moreover, there is evidence demonstrating that *MYC* might be a key regulator of cell cycle-mediated neuronal cell death (Lee et al., 2009). Notably, some of the VMR genes may contribute to AD pathology through changing variability, but not mean level of DNA methylation, and thus would otherwise be missed by testing DMPs only.

Our network analysis identified two modules (the brown module and yellow module). The hub genes of these networks highlight the potential role of *HOXA3* and *HOXA5* in AD pathology. This is corroborated by previous evidence demonstrating that expression of the *HOX* family genes plays critical roles in the formation and function of the central nervous system (Lizen et al., 2017; Rezsóhazy et al., 2015), and is also in line with a recent study demonstrating that elevated DNA methylation level of the *HOXA* genes was associated with AD neuropathology (Smith et al., 2018). While the current analysis and previous studies (Hansen et al., 2011; Jones et al., 2015; Paul et al., 2016; Webster et al., 2018) suggest a potential important role of altered DNA methylation variability in human diseases, the biological mechanisms behind the observed associations remain to be determined. It is possible that altered DNA methylation stability may reflect the exposure of disease risk factors, and thus differential methylation variability could represent the different exposure profiles of these factors between cases and controls (Teschendorff et al., 2012). It is also possible that the altered DNA methylation variability could represent beneficial adaptations in response of disease process (Feinberg and Irizarry, 2010; Hansen et al., 2011). Further research is needed to test these hypotheses.

Several aspects of our findings deserve discussions. First, while we found that some genes showed differential methylation at both the mean and variance levels, e.g., about 18% and 26% of the VMPs associated with amyloid- β and tangles, respectively, overlapped with the DMPs. However, many genes showed either variable methylation or differential mean methylation, but not both. For instance, majority of the identified VMPs (more than 70%) did not overlap with those DMPs detected based on mean DNA methylation analysis. These observations indicate that VMPs and DMPs analyses capture different sets of genes associated with AD pathology, and thus both mean and variance should be tested in epigenetic analysis. Statistical analysis that tests only the mean DNA methylation level without considering variability will miss important disease-related genes. Second, the VMPs/VMRs identified in the current study most likely represented true biological variability instead of technical artifacts, because: (1) many VMR genes, especially those associated with tangles, also showed differential expression; (2) the identified VMR genes were significantly over-represented in excitatory synapse, and neuron differentiation,

calcium ion transmembrane transport, etc. Moreover, these VMR genes were enriched in previous GWAS loci for AD (Beecham et al., 2014; Jun et al., 2016; Sherva et al., 2014). Third, as part of the QC procedures, we compared DMPs identified in this study with those reported in previous EWAS in the same brain cortex (De Jager et al., 2014). We found that 65.4% of the DMPs identified in the current analysis overlapped with previously reported probes associated with neuritic plaque (De Jager et al., 2014), demonstrating the credibility of our QC procedures used in the current analysis. However, the DMPs identified in the current analysis are not identical to those reported previously by De Jager et al. (2014) due to the following reasons: (1) different preprocessing pipeline/QC criteria and thus different number and sets of initial probes included in the final data analysis; (2) slightly different phenotypes used in the two studies. In the work by De Jager et al. (2014), the authors used neuritic plaque quantified by microscopic examination of silver-stained slides from 5 brain regions (midfrontal cortex, midtemporal cortex, inferior parietal cortex, entorhinal cortex, and hippocampus), whereas in our analysis, we used a composite measure of neuropathological burden quantified by molecularly-specific immunohistochemistry across 8 brain regions (hippocampus, entorhinal cortex, midfrontal gyrus, inferior temporal, anterior gyrus, calcarine cortex, cingulate region, and superior frontal gyrus); and (3) different statistical models along with different covariates adjustments. Fourth, our results demonstrate that the two AD hallmarks – amyloid- β and tangles – are etiologically highly heterogeneous. For instance, of the identified VMPs associated with either amyloid- β or tangles, only one probe (*cg01869224*) associates with both traits. Moreover, we observed differential genomic distribution of the VMPs associated with amyloid- β or tangles, that is, amyloid- β -associated VMPs are largely enriched in introns and intergenic regions, but depleted in promoter regions, whereas those associated with neurofibrillary tangles are over-represented in promoters but depleted in intergenic regions. While the mechanisms behind the observed trait-specific distribution of these VMPs are unclear, it is possible that DNA methylation variability of probes located in the promoter regions may have a larger impact on gene expression than those located in the introns/intergenic regions (Lee and Wiemels, 2015; Suzuki and Bird, 2008). In line with this hypothesis, majority (62.5%) of the VMR genes associated with tangles are differentially expressed, whereas only a small proportion (1.8%) of the amyloid- β -associated VMR genes showed differential expression. The observed trait-specific genomic distributions of the identified VMPs also suggest that the two hallmarks of AD (i.e., amyloid plaque and tangle) may be etiologically heterogeneous. Fifth, nearly 90% of the amyloid- β -associated VMPs are hypervariable probes, whereas majority (~58%) of the tangle-related VMPs are hypovariable ones. While these observations support a role of DNA methylation variability in gene regulation and AD pathogenesis, our findings unravel the etiological heterogeneity of the two hallmarks, and highlight the importance of studying both traits in future research. Last, because the relationship between gene expression and DNA methylation is not one-on-one, it is possible that DNA methylation variability affects the two phenotypes through different biological pathways. Together, our findings suggest that the relationship between DNA methylation variability and gene expression in human AD brain may be trait-specific.

Our study has several limitations. First, the present study only examined VMRs in one brain region (prefrontal cortex), but it is possible that DNA methylation variability varies across

different brain regions. Future research should investigate the region-specific effect of DNA methylation variability on AD pathology. Second, participants included in the current analysis are highly educated European Caucasians, and our results may not be generalized to other ethnic groups or population settings. Third, the identified association between DNA methylation variability and AD pathology do not imply causality. However, our study has several strengths. To our knowledge, this is the first study examining the association between variably methylated probes/regions and AD neuropathology in a large collection of postmortem human brains. Moreover, we conducted statistical analyses including differentially variably methylation analysis, functional annotation analysis, pathway enrichment analysis and co-methylation network analysis. Further, we studied different neuropathological phenotypes, e.g., amyloid- β and tangles, and identified trait-specific association of variably methylated genes with AD neuropathology.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors would like to thank the participants of the Religious Orders Study and Memory and Aging Project studies and the staff of the Rush Alzheimer's Disease Center.

Funding sources: This work was supported by the National Institutes of Health grants RF1AG52476, P30AG10161, RF1AG15819, R01AG17917, R01AG16042, U01AG46152, RF1AG36042, and R01AG36836.

References

- A Bennett D, A Schneider J, Arvanitakis Z, S Wilson R, 2012a Overview and findings from the religious orders study. *Curr. Alzheimer Res.* 9, 628–645. [PubMed: 22471860]
- A Bennett D, A Schneider J, S Buchman A, L Barnes L, A Boyle P, S Wilson R, 2012b Overview and findings from the rush Memory and Aging Project. *Curr. Alzheimer Res.* 9, 646–663. [PubMed: 22471867]
- Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, Irizarry RA, 2014 Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* 30, 1363–1369. [PubMed: 24478339]
- Assenov Y, Müller F, Lutsik P, Walter J, Lengauer T, Bock C, 2014 Comprehensive analysis of DNA methylation data with RnBeads. *Nat. Methods* 11, 1138. [PubMed: 25262207]
- Beecham GW, Hamilton K, Naj AC, Martin ER, Huentelman M, Myers AJ, Corneveaux JJ, Hardy J, Vonsattel J-P, Younkin SG, others, 2014 Genome-wide association meta-analysis of neuropathologic features of Alzheimer's disease and related dementias. *PLoS Genet.* 10, e1004606. [PubMed: 25188341]
- Benjamini Y, Hochberg Y, 1995 Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser B* 289–300.
- Bennett DA, De Jager PL, Leurgans SE, Schneider JA, 2009 Neuropathologic intermediate phenotypes enhance association to Alzheimer susceptibility alleles. *Neurology* 72, 1495–1503. [PubMed: 19398704]
- Bennett DA, Schneider JA, Arvanitakis Z, Kelly JF, Aggarwal NT, Shah RC, Wilson RS, 2006 Neuropathology of older persons without cognitive impairment from two community-based studies. *Neurology* 66, 1837–1844. [PubMed: 16801647]
- Bennett DA, Schneider JA, Wilson RS, Bienias JL, Arnold SE, 2004 Neurofibrillary tangles mediate the association of amyloid load with clinical Alzheimer disease and level of cognitive function. *Arch. Neurol* 61, 378–384. [PubMed: 15023815]

- Brown MB, 1975 400: A method for combining non-independent, one-sided tests of significance. *Biometrics* 987–992.
- Campion D, Pottier C, Nicolas G, Le Guennec K, Rovelet-Lecrux A, 2016 Alzheimer disease: modeling an A β -centered biological network. *Mol. Psychiatry* 21, 861. [PubMed: 27021818]
- Cash AD, Aliev G, Siedlak SL, Nunomura A, Fujioka H, Zhu X, Raina AK, Vinters HV, Tabaton M, Johnson AB, others, 2003 Microtubule reduction in Alzheimer's disease and aging is independent of τ filament formation. *Am. J. Pathol* 162, 1623–1627. [PubMed: 12707046]
- Cozza A, Melissari E, Iacopetti P, Mariotti V, Tedde A, Nacmias B, Conte A, Sorbi S, Pellegrini S, 2008 SNPs in neurotrophin system genes and Alzheimer's disease in an Italian population. *J. Alzheimer's Dis* 15, 61–70. [PubMed: 18780967]
- Csardi G, Nepusz T, 2006 The igraph software package for complex network research. *InterJournal, Complex Syst.* 1695, 1–9.
- Dang C V, 1999 c-Myc target genes involved in cell growth, apoptosis, and metabolism. *Mol. Cell. Biol.* 19, 1–11.
- De Jager PL, Srivastava G, Lunnon K, Burgess J, Schalkwyk LC, Yu L, Eaton ML, Keenan BT, Ernst J, McCabe C, others, 2014 Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. *Nat. Neurosci* 17, 1156. [PubMed: 25129075]
- Duthey B, 2013 Background paper 6.11: Alzheimer disease and other dementias. *A Public Heal. Approach to Innov* 1–74.
- Feinberg AP, Irizarry RA, 2010 Stochastic epigenetic variation as a driving force of development, evolutionary adaptation, and disease. *Proc. Natl. Acad. Sci* 107, 1757–1764. [PubMed: 20080672]
- Guintivano J, Aryee MJ, Kaminsky ZA, 2013 A cell epigenotype specific model for the correction of brain cellular heterogeneity bias and its application to age, brain region and major depression. *Epigenetics* 8, 290–302. [PubMed: 23426267]
- Hansen KD, Timp W, Bravo HC, Sabunciyan S, Langmead B, McDonald OG, Wen B, Wu H, Liu Y, Diep D, others, 2011 Increased methylation variation in epigenetic domains across cancer types. *Nat. Genet* 43, 768. [PubMed: 21706001]
- Hastie T, Tibshirani R, Narasimhan B, Chu G, 2001 impute: Imputation for microarray data. *Bioinformatics* 17, 520–525. [PubMed: 11395428]
- Irier HA, Jin P, 2012 Dynamics of DNA methylation in aging and Alzheimer's disease. *DNA Cell Biol.* 31, S--42.
- Jean DC, Baas PW, 2013 It cuts two ways: microtubule loss during Alzheimer disease. *EMBO J.* 32, 2900–2902. [PubMed: 24076651]
- Jones MJ, Goodman SJ, Kobor MS, 2015 DNA methylation and healthy human aging. *Aging Cell* 14, 924–932. [PubMed: 25913071]
- Jun G, Ibrahim-Verbaas CA, Vronskaya M, Lambert J-C, Chung J, Naj AC, Kunkle BW, Wang L-S, Bis JC, Bellenguez C, others, 2016 A novel Alzheimer disease locus located near the gene encoding tau protein. *Mol. Psychiatry* 21, 108. [PubMed: 25778476]
- Kanekiyo T, Cirrito JR, Liu C-C, Shinohara M, Li J, Schuler DR, Shinohara M, Holtzman DM, Bu G, 2013 Neuronal clearance of amyloid- β by endocytic receptor LRP1. *J. Neurosci* 33, 19276–19283. [PubMed: 24305823]
- Kauwe JSK, Bailey MH, Ridge PG, Perry R, Wadsworth ME, Hoyt KL, Staley LA, Karch CM, Harari O, Cruchaga C, others, 2014 Genome-wide association study of CSF levels of 59 Alzheimer's disease candidate proteins: significant associations with proteins involved in amyloid processing and inflammation. *PLoS Genet.* 10, e1004758. [PubMed: 25340798]
- Langfelder P, Horvath S, 2008 WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9, 559. [PubMed: 19114008]
- Lawrence M, Huber W, Pages H, Aboyoun P, Carlson M, Gentleman R, Morgan MT, Carey VJ, 2013 Software for computing and annotating genomic ranges. *PLoS Comput. Biol* 9, e1003118. [PubMed: 23950696]
- Lee H, Casadesus G, Nunomura A, Zhu X, Castellani RJ, Richardson SL, Perry G, Felsher DW, Petersen RB, Smith MA, 2009 The neuronal expression of MYC causes a neurodegenerative phenotype in a novel transgenic mouse. *Am. J. Pathol* 174, 891–897. [PubMed: 19164506]

- Lee R, Kermani P, Teng KK, Hempstead BL, 2001 Regulation of cell survival by secreted proneurotrophins. *Science* (80-.). 294, 1945–1948.
- Lee S-T, Wiemels JL, 2015 Genome-wide CpG island methylation and intergenic demethylation propensities vary among different tumor sites. *Nucleic Acids Res.* 44, 1105–1117. [PubMed: 26464434]
- Leung R, Proitsi P, Simmons A, Lunnon K, Güntert A, Kronenberg D, Pritchard M, Tsolaki M, Mecocci P, Kloszewska I, others, 2013 Inflammatory proteins in plasma are associated with severity of Alzheimer's disease. *PLoS One* 8, e64971. [PubMed: 23762274]
- Lizen B, Moens C, Mouheiche J, Sacré T, Ahn M-T, Jeannotte L, Salti A, Gofflot F, 2017 Conditional Loss of Hoxa5 Function Early after Birth Impacts on Expression of Genes with Synaptic Function. *Front. Mol. Neurosci* 10, 369. [PubMed: 29187810]
- Lord J, Cruchaga C, 2014 The epigenetic landscape of Alzheimer's disease. *Nat. Neurosci.* 17, 1138. [PubMed: 25157507]
- Lunnon K, Smith R, Hannon E, De Jager PL, Srivastava G, Volta M, Troakes C, Al-Sarraj S, Burrage J, Macdonald R, others, 2014 Methylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease. *Nat. Neurosci* 17, 1164. [PubMed: 25129077]
- Margolin AA, Nemenman I, Basso K, Wiggins C, Stolovitzky G, Dalla Favera R, Califano A, 2006 ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context, in: *BMC Bioinformatics*. p. S7.
- Menges CW, Altomare DA, Testa JR, 2009 FAS-associated factor 1 (FAF1): diverse functions and implications for oncogenesis. *Cell Cycle* 8, 2528–2534. [PubMed: 19597341]
- Meyer PE, Lafitte F, Bontempi G, 2008 Minet: an open source R/Bioconductor package for mutual information based network inference. *BMC Bioinformatics* 9.
- Palumbo D, Affinito O, Monticelli A, Coccozza S, 2018 DNA Methylation variability among individuals is related to CpGs cluster density and evolutionary signatures. *BMC Genomics* 19, 229. [PubMed: 29606093]
- Paul DS, Teschendorff AE, Dang MAN, Lowe R, Hawa MI, Ecker S, Beyan H, Cunningham S, Fouts AR, Ramelius A, others, 2016 Increased DNA methylation variability in type 1 diabetes across three immune effector cell types. *Nat. Commun.* 7, 13555. [PubMed: 27898055]
- Phipson B, Maksimovic J, Oshlack A, 2015 missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. *Bioinformatics* 32, 286–288. [PubMed: 26424855]
- Phipson B, Oshlack A, 2014 DiffVar: a new method for detecting differential variability with application to methylation in cancer and aging. *Genome Biol.* 15, 465. [PubMed: 25245051]
- Poole W, Gibbs DL, Shmulevich I, Bernard B, Knijnenburg TA, 2016 Combining dependent P-values with an empirical adaptation of Brown's method. *Bioinformatics* 32, i430–i436. [PubMed: 27587659]
- Rezsohazy R, Saurin AJ, Maurel-Zaffran C, Graba Y, 2015 Cellular and molecular insights into Hox protein action. *Development* 142, 1212–1227. [PubMed: 25804734]
- Saffari A, Silver MJ, Zavattari P, Moi L, Columbano A, Meaburn EL, Dudbridge F, 2018 Estimation of a significance threshold for epigenome-wide association studies. *Genet. Epidemiol.* 42, 20–33. [PubMed: 29034560]
- Sánchez L, Alvarez V, González P, González I, Alvarez R, Coto E, 2001 Variation in the LRP-associated protein gene (LRPAP1) is associated with late-onset Alzheimer disease. *Am. J. Med. Genet* 105, 76–78. [PubMed: 11425005]
- Sherva R, Tripodis Y, Bennett DA, Chibnik LB, Crane PK, De Jager PL, Farrer LA, Saykin AJ, Shulman JM, Naj A, others, 2014 Genome-wide association study of the rate of cognitive decline in Alzheimer's disease. *Alzheimer's Dement.* 10, 45–52. [PubMed: 23535033]
- Smith RG, Hannon E, De Jager PL, Chibnik L, Lott SJ, Condliffe D, Smith AR, Haroutunian V, Troakes C, Al-Sarraj S, others, 2018 Elevated DNA methylation across a 48-kb region spanning the HOXA gene cluster is associated with Alzheimer's disease neuropathology. *Alzheimer's Dement.*
- Smyth GK, 2005 Limma: linear models for microarray data, in: *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Springer, pp. 397–420.

- Stephens MA, 1974 EDF statistics for goodness of fit and some comparisons. *J. Am. Stat. Assoc* 69, 730–737.
- Sul J-W, Park M-Y, Shin J, Kim Y-R, Yoo S-E, Kong Y-Y, Kwon K-S, Lee YH, Kim E, 2013 Accumulation of the parkin substrate, FAF1, plays a key role in the dopaminergic neurodegeneration. *Hum. Mol. Genet.* 22, 1558–1573. [PubMed: 23307929]
- Suzuki MM, Bird A, 2008 DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet* 9, 465. [PubMed: 18463664]
- Teschendorff AE, Jones A, Fiegl H, Sargent A, Zhuang JJ, Kitchener HC, Widschwendter M, 2012 Epigenetic variability in cells of normal cytology is associated with the risk of future morphological transformation. *Genome Med.* 4, 24. [PubMed: 22453031]
- Villeda SA, Luo J, Mosher KI, Zou B, Britschgi M, Bieri G, Stan TM, Fainberg N, Ding Z, Eggel A, others, 2011 The ageing systemic milieu negatively regulates neurogenesis and cognitive function. *Nature* 477, 90. [PubMed: 21886162]
- Watson CT, Roussos P, Garg P, Ho DJ, Azam N, Katsel PL, Haroutunian V, Sharp AJ, 2016 Genome-wide DNA methylation profiling in the superior temporal gyrus reveals epigenetic signatures associated with Alzheimer's disease. *Genome Med.* 8, 5. [PubMed: 26803900]
- Webster AP, Plant D, Ecker S, Zufferey F, Bell JT, Feber A, Paul DS, Beck S, Barton A, Williams FMK, others, 2018 Increased DNA methylation variability in rheumatoid arthritis discordant monozygotic twins. *bioRxiv* 314963.
- Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, Klemm A, Flicek P, Manolio T, Hindorf L, others, 2013 The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.* 42, D1001–D1006. [PubMed: 24316577]
- Wilson RS, Boyle PA, Yu L, Barnes LL, Sytsma J, Buchman AS, Bennett DA, Schneider JA, 2015 Temporal course and pathologic basis of unawareness of memory loss in dementia. *Neurology* 85, 984–991. [PubMed: 26311746]
- Yang J, Weedon MN, Purcell S, Lettre G, Estrada K, Willer CJ, Smith AV, Ingelsson E, O'connell JR, Mangino M, others, 2011 Genomic inflation factors under polygenic inheritance. *Eur. J. Hum. Genet* 19, 807. [PubMed: 21407268]
- Zhu C, Xu B, Sun X, Zhu Q, Sui Y, 2017 Targeting CCR3 to Reduce Amyloid β Production, Tau Hyperphosphorylation, and Synaptic Loss in a Mouse Model of Alzheimer's Disease. *Mol. Neurobiol* 54, 7964–7978. [PubMed: 27878757]

Highlights

- Conventional epigenetic analysis focuses on testing the mean difference in DNA methylation between cases and controls, but DNA methylation instability (e.g., increased variability) may also affect disease susceptibility.
- No study has examined the association of methylation variability with AD.
- We identified genomic regions harboring variably methylated probes associated with A β plaques and neurofibrillary tangles (NFTs).
- These regions are largely non-overlapped with regions detected by mean methylation analysis, highlighting the importance of testing methylation variability in future research.

Figure 1.

Volcano plots showing the identified VMPs associated with amyloid plaques (**1A**) and tangles (**1B**). The p -values in $-\log_{10}$ scale (Y-axis) are plotted against the $\log_2$ fold change ($\log_2FC$) of variability with respect to unit increase of the pathological burden (X-axis). A positive $\log_2FC$ represents hypervariable methylation and a negative $\log_2FC$ represents hypovariable methylation. The VMPs with $q < 0.05$ are marked with the blue color and the VMPs with $q > 0.05$ are marked with red color. The dashed horizontal line represents genome wide significance level $\alpha = 3.6 \times 10^{-8}$, and the VMPs with $< 3.6 \times 10^{-8}$ are annotated by genes within $\pm 5\text{kb}$ of the probe.

Figure 2. Genomic distribution of the identified VMPs associated with amyloid- β and tangles. The “null” represents all probes included in the Illumina HumanMethylation450 array.

Table 1.

Clinical characteristics of the brain donors (N = 706)

Characteristics	% or Mean $\pm$ SD
Age (years)	88.0 $\pm$ 6.7
Males (%)	261 (37.0)
Education level (years)	16.4 $\pm$ 3.6
Study affiliation (MAP, %)	326 (46.2)
MMSE score	15.8 $\pm$ 10.2
PMI (hours)	7.5 $\pm$ 5.9
AD	298 (42.2)
PHF tau tangles [*]	6.4 $\pm$ 8.1
Amyloid- β plaques [†]	1.4 $\pm$ 1.1

^{*} Mean of PHF tau tangles score in 8 brain regions[†] Mean of amyloid- β score (square root transformed) in 8 brain regions

Table 2.

Top 50 most significant variably methylated probes (VMPs) associated with amyloid- β ($q < 0.05$).

CpG probe	Chr	Position (bps)	Genomic feature	p-value	q-value	Gene *	Direction †
cg01624068	13	113,137,799	intergenic	1.35×10^{-9}	2.68×10^{-4}	<i>TUBGCP3</i>	↑
cg10506070	5	114,502,492	intergenic	1.60×10^{-9}	2.68×10^{-4}		↑
cg00391025	3	100,427,239	intergenic	2.13×10^{-9}	2.68×10^{-4}		↑
cg19734937	2	63,855,048	intergenic	2.58×10^{-9}	2.68×10^{-4}		↑
cg25871696	13	81,914,087	intergenic	3.03×10^{-9}	2.68×10^{-4}		↑
cg26102082	17	47,590,272	exons	7.70×10^{-9}	5.67×10^{-4}	<i>NGFR</i>	↓
cg06060962	6	32,515,150	introns	1.01×10^{-8}	5.75×10^{-4}	<i>HLA-DRB5</i>	↑
cg02226953	7	150,525,449	intergenic	1.04×10^{-8}	5.75×10^{-4}		↑
cg03660952	11	2,644,418	exons	1.21×10^{-8}	5.93×10^{-4}	<i>KCNQ1OT1</i>	↑
cg23968456	10	73,521,631	exons	1.87×10^{-8}	8.25×10^{-4}	<i>C10orf54</i>	↓
cg14313916	8	8,640,017	intergenic	2.61×10^{-8}	1.00×10^{-3}		↑
cg15948245	7	104,581,746	promoters	2.72×10^{-8}	1.00×10^{-3}		↑
cg21110873	12	131,712,414	intergenic	3.66×10^{-8}	1.24×10^{-3}		↑
cg19550439	5	64,470,134	introns	4.64×10^{-8}	1.47×10^{-3}	<i>ADAMTS6</i>	↑
cg23493751	3	46,204,787	promoters	5.86×10^{-8}	1.65×10^{-3}	<i>CCR3</i>	↑
cg25814432	19	54,620,431	introns	5.97×10^{-8}	1.65×10^{-3}	<i>PRPF31</i>	↑
cg17917920	7	93,220,788	intergenic	7.00×10^{-8}	1.82×10^{-3}		↑
cg13501912	3	87,424,518	intergenic	7.72×10^{-8}	1.90×10^{-3}		↑
cg06890950	4	81,382,711	introns	8.28×10^{-8}	1.93×10^{-3}	<i>C4orf22</i>	↑
cg08041573	21	33,651,723	intergenic	8.72×10^{-8}	1.93×10^{-3}		↑
cg03043928	11	2,085,203	intergenic	9.28×10^{-8}	1.95×10^{-3}		↑
cg21002735	11	77,350,306	intergenic	9.85×10^{-8}	1.98×10^{-3}	<i>CLNS1A</i>	↑
cg07725198	1	27,229,130	intergenic	1.13×10^{-7}	2.00×10^{-3}	<i>GPATCH3</i>	↑
cg09609314	2	174,147,631	intergenic	1.20×10^{-7}	2.00×10^{-3}		↑
cg18677996	5	17,031,869	intergenic	1.25×10^{-7}	2.00×10^{-3}		↑
cg10655371	7	91,749,682	introns	1.29×10^{-7}	2.00×10^{-3}	<i>CYP51A1</i>	↑
cg20996620	7	20,623,878	intergenic	1.30×10^{-7}	2.00×10^{-3}		↑
cg06931905	8	42,036,940	intergenic	1.35×10^{-7}	2.00×10^{-3}		↑
cg00689014	17	39,869,138	introns	1.38×10^{-7}	2.00×10^{-3}	<i>JUP</i>	↑
cg13290719	4	58,063,449	introns	1.43×10^{-7}	2.00×10^{-3}	<i>IGFBP7-AS1</i>	↑
cg10575940	17	75,365,389	intergenic	1.44×10^{-7}	2.00×10^{-3}		↑
cg07561890	2	133,036,548	intergenic	1.45×10^{-7}	2.00×10^{-3}		↑
cg16608672	5	30,346,809	intergenic	1.86×10^{-7}	2.49×10^{-3}		↑
cg14446166	7	148,934,349	intergenic	2.05×10^{-7}	2.60×10^{-3}		↑
cg27340958	4	7,755,730	promoters	2.19×10^{-7}	2.60×10^{-3}	<i>AFAP1-AS1</i>	↑
cg26529712	19	35,399,253	intergenic	2.20×10^{-7}	2.60×10^{-3}		↑
cg26879538	6	30,014,916	introns	2.21×10^{-7}	2.60×10^{-3}	<i>ZNRD1ASP</i>	↑

CpG probe	Chr	Position (bps)	Genomic feature	<i>p</i> -value	<i>q</i> -value	Gene *	Direction [†]
cg03318222	19	56,847,740	introns	2.24×10 ⁻⁷	2.60×10 ⁻³	<i>ZSCAN5A</i>	↑
cg21512370	13	48,990,327	introns	2.34×10 ⁻⁷	2.65×10 ⁻³	<i>LPAR6</i>	↑
cg12447346	1	161,055,663	introns	2.45×10 ⁻⁷	2.70×10 ⁻³	<i>NECTIN4</i>	↑
cg27616595	7	146,484,518	intergenic	2.77×10 ⁻⁷	2.98×10 ⁻³		↑
cg01481116	11	93,867,992	introns	3.11×10 ⁻⁷	3.27×10 ⁻³	<i>PANX1</i>	↑
cg14519534	7	102,082,149	introns	3.32×10 ⁻⁷	3.36×10 ⁻³	<i>ORAI2</i>	↑
cg24620463	7	95,917,902	intergenic	3.35×10 ⁻⁷	3.36×10 ⁻³		↑
cg20185525	17	41,278,712	introns	3.54×10 ⁻⁷	3.48×10 ⁻³	<i>BRCA1</i>	↑
cg21775279	1	28,285,385	exons	3.71×10 ⁻⁷	3.56×10 ⁻³	<i>SMPDL3B</i>	↓
cg03169557	16	89,598,950	exons	4.05×10 ⁻⁷	3.80×10 ⁻³	<i>SPG7</i>	↓
cg02180296	14	24,404,306	intergenic	4.24×10 ⁻⁷	3.90×10 ⁻³	<i>DHRS4-AS1</i>	↑
cg02312525	22	50,525,778	intergenic	4.48×10 ⁻⁷	4.04×10 ⁻³	<i>MLC1</i>	↑
cg19988798	5	175,511,300	intergenic	4.58×10 ⁻⁷	4.05×10 ⁻³		↑

* The nearest annotated gene within ±5kb of the CpG probe

[†] ↑ indicates hyper-variability (increased variability in subjects with higher neuropathological burden as compared to those with lower burden), and
 ↓ indicates hypo-variability (decreased variability in subjects with higher neuropathological burden as compared to those with lower burden).

Table 3.Top 50 most significant variably methylated probes (VMPs) associated with tangles ($q < 0.05$)

CpG probe	Chr	Position (bps)	Genomic features	p-value	q-value	Gene *	Direction †
cg21406811	1	51,271,043	introns	4.08×10^{-12}	1.63×10^{-6}	<i>FAF1</i>	↑
cg23067228	15	70,881,206	intergenic	7.37×10^{-12}	1.63×10^{-6}		↑
cg22895202	14	21,238,166	promoters	3.11×10^{-11}	4.58×10^{-6}	<i>EDDM3B</i>	↑
cg09183598	19	57,679,015	promoters	5.26×10^{-11}	5.81×10^{-6}	<i>DUXA</i>	↑
cg12308275	15	43,558,855	promoters	1.54×10^{-10}	1.37×10^{-5}	<i>TGM5</i>	↑
cg16008918	4	78,740,080	intergenic	2.03×10^{-10}	1.49×10^{-5}		↑
cg13689497	19	3,201,528	intergenic	2.82×10^{-9}	1.78×10^{-4}		↑
cg17678740	7	53,254,947	intergenic	4.36×10^{-9}	2.41×10^{-4}		↓
cg23524184	17	48,129,754	intergenic	2.75×10^{-8}	1.25×10^{-3}	<i>ITGA3</i>	↑
cg22661129	8	35,580,715	promoters	3.11×10^{-8}	1.25×10^{-3}	<i>UNC5D</i>	↑
cg25868285	7	134,455,275	intergenic	3.12×10^{-8}	1.25×10^{-3}		↑
cg04559604	8	22,031,763	intergenic	5.34×10^{-8}	1.97×10^{-3}		↑
cg01363574	3	40,354,920	intergenic	1.03×10^{-7}	3.31×10^{-3}	<i>EIF1B-AS1</i>	↑
cg16353628	2	85,239,916	introns	1.05×10^{-7}	3.31×10^{-3}	<i>KCMF1</i>	↑
cg09352908	3	37,017,749	intergenic	1.23×10^{-7}	3.63×10^{-3}		↓
cg14115740	9	98,054,883	intergenic	2.16×10^{-7}	5.97×10^{-3}		↓
cg04212651	6	112,671,523	promoters	2.30×10^{-7}	5.98×10^{-3}	<i>RFPL4B</i>	↑
cg02745683	4	3,513,990	exons	2.60×10^{-7}	6.28×10^{-3}	<i>LRPAP1</i>	↑
cg20376421	12	56,546,193	promoters	2.70×10^{-7}	6.28×10^{-3}	<i>MYL6B</i>	↓
cg26227935	18	77,682,412	intergenic	3.13×10^{-7}	6.91×10^{-3}		↑
cg19729949	5	142,514,108	intergenic	4.11×10^{-7}	8.66×10^{-3}		↑
cg01869224	12	53,356,063	intergenic	4.79×10^{-7}	9.13×10^{-3}		↑
cg05929056	21	43,274,690	intergenic	4.81×10^{-7}	9.13×10^{-3}		↑
cg12538597	11	85,906,334	intergenic	5.13×10^{-7}	9.13×10^{-3}		↓
cg20660860	2	122,095,855	exons	5.27×10^{-7}	9.13×10^{-3}	<i>CLASPI</i>	↑
cg24663683	5	39,721,720	intergenic	5.26×10^{-7}	9.13×10^{-3}		↓
cg09537031	8	19,171,706	intergenic	5.87×10^{-7}	9.13×10^{-3}		↑
cg09828625	20	23,331,259	promoters	5.97×10^{-7}	9.13×10^{-3}	<i>NXT1</i>	↑
cg08866589	5	141,303,260	promoters	6.02×10^{-7}	9.13×10^{-3}	<i>KIAA0141</i>	↑
cg20208633	4	55,405,888	intergenic	6.19×10^{-7}	9.13×10^{-3}		↑
cg07252851	5	74,063,056	promoters	6.55×10^{-7}	9.29×10^{-3}	<i>GFM2</i>	↓
cg10690677	1	87,019,175	introns	6.73×10^{-7}	9.29×10^{-3}	<i>CLCA4</i>	↓
cg12550816	19	47,163,979	promoters	7.07×10^{-7}	9.46×10^{-3}	<i>DACT3-AS1</i>	↓
cg06159435	11	118,966,351	promoters	7.31×10^{-7}	9.46×10^{-3}	<i>H2AFX</i>	↓
cg13838599	9	128,989,097	intergenic	7.49×10^{-7}	9.46×10^{-3}		↑
cg06459913	16	11,361,823	introns	7.93×10^{-7}	9.59×10^{-3}	<i>RMI2</i>	↑
cg18803079	1	64,014,643	promoters	8.03×10^{-7}	9.59×10^{-3}	<i>EFCAB7</i>	↓

CpG probe	Chr	Position (bps)	Genomic features	<i>p</i> -value	<i>q</i> -value	Gene *	Direction †
cg14559409	10	65,930,703	intergenic	8.87×10^{-7}	1.03×10^{-2}		↓
cg21962918	5	134,094,454	intergenic	1.00×10^{-6}	1.11×10^{-2}		↓
cg02738677	3	133,265,129	intergenic	1.01×10^{-6}	1.11×10^{-2}		↓
cg10269358	19	17,728,585	exons	1.03×10^{-6}	1.11×10^{-2}	<i>UNC13A</i>	↑
cg22331096	8	128,749,328	introns	1.16×10^{-6}	1.21×10^{-2}	<i>MYC</i>	↓
cg01778384	13	27,236,961	introns	1.20×10^{-6}	1.21×10^{-2}	<i>WASF3</i>	↑
cg02746684	3	197,347,276	introns	1.23×10^{-6}	1.21×10^{-2}	<i>LOC220729</i>	↑
cg08450091	3	82,857,215	intergenic	1.24×10^{-6}	1.21×10^{-2}		↓
cg04005938	7	148,334,417	intergenic	1.27×10^{-6}	1.21×10^{-2}		↓
cg27618483	2	155,433,424	intergenic	1.31×10^{-6}	1.21×10^{-2}		↓
cg17095167	13	21,833,918	intergenic	1.31×10^{-6}	1.21×10^{-2}		↓
cg01757206	17	7,183,913	intergenic	1.46×10^{-6}	1.32×10^{-2}	<i>SLC2A4</i>	↓
cg26342398	1	85,156,326	promoters	1.61×10^{-6}	1.42×10^{-2}	<i>SSX2IP</i>	↓

* The nearest annotated gene within ± 5 kb of the CpG probe

† ↑ indicates hypervariable (increased variability in subjects with higher neuropathological burden as compared to those with lower burden), and ↓ indicates hypovariable (decreased variability in subjects with higher neuropathological burden as compared to those with lower burden).

Table 4.

Top 30 enriched pathways related to AD neuropathology

Pathway/GO terms	<i>q</i> -value [*]	
	amyloid- β	tangles
Regulation of neuron projection development	3.88×10^{-12}	3.43×10^{-12}
Calcium ion transmembrane transporter activity	2.10×10^{-9}	2.42×10^{-5}
Postsynapse	6.94×10^{-9}	8.00×10^{-11}
Calcium ion transmembrane transport	1.04×10^{-8}	1.16×10^{-4}
Positive regulation of cell projection organization	1.73×10^{-8}	4.03×10^{-7}
Positive regulation of neuron projection development	2.48×10^{-8}	4.07×10^{-8}
Cell leading edge	2.94×10^{-8}	4.22×10^{-8}
Positive regulation of neuron differentiation	8.55×10^{-8}	6.84×10^{-8}
Divalent inorganic cation transmembrane transporter activity	2.71×10^{-7}	1.69×10^{-3}
Calcium ion transport	1.29×10^{-6}	8.72×10^{-5}
Actin binding	1.51×10^{-6}	4.07×10^{-8}
REACTOME: Axon guidance	2.15×10^{-6}	1.19×10^{-5}
Positive regulation of nervous system development	3.77×10^{-6}	4.36×10^{-7}
Cell projection membrane	5.90×10^{-6}	2.04×10^{-6}
Divalent inorganic cation transport	5.90×10^{-6}	6.08×10^{-4}
Regulation of dendrite development	8.98×10^{-6}	1.60×10^{-4}
Gtpase binding	9.91×10^{-6}	6.31×10^{-5}
Synaptic membrane	1.02×10^{-5}	4.07×10^{-8}
Guanyl nucleotide exchange factor activity	1.21×10^{-5}	6.84×10^{-8}
Main axon	1.55×10^{-5}	8.56×10^{-4}
Excitatory synapse	2.39×10^{-5}	2.93×10^{-9}
Regulation of small gtpase mediated signal transduction	2.39×10^{-5}	1.08×10^{-6}
Regulation of cell morphogenesis involved in differentiation	8.82×10^{-5}	4.36×10^{-7}
Cation channel activity	1.17×10^{-4}	1.31×10^{-6}
Neuron porjection morphonensis	1.31×10^{-4}	4.03×10^{-7}
Axon	1.33×10^{-4}	9.97×10^{-8}
Receptor complex	2.27×10^{-4}	1.50×10^{-6}
<i>Ras</i> guanyl nucleotide exchange factor activity	1.42×10^{-3}	4.36×10^{-7}
Gated channel activity	3.47×10^{-3}	1.06×10^{-7}
Neuronal postsynaptic density	3.99×10^{-2}	1.11×10^{-6}

* Adjusted for multiple testing by the Benjamini-Hochberg method