


RESEARCH

Open Access



# Harnessing peripheral DNA methylation differences in the Alzheimer's Disease Neuroimaging Initiative (ADNI) to reveal novel biomarkers of disease

Aparna Vasanthakumar<sup>1\*†</sup> , Justin W. Davis<sup>1†</sup>, Kenneth Idler<sup>1†</sup>, Jeffrey F. Waring<sup>1†</sup>, Elizabeth Asque<sup>1</sup>, Bridget Riley-Gillis<sup>1</sup>, Shaun Grosskurth<sup>1</sup>, Gyan Srivastava<sup>2</sup>, Sungeun Kim<sup>3,8</sup>, Kwangsik Nho<sup>3</sup>, Kelly N. H. Nudelman<sup>3,4</sup>, Kelley Faber<sup>4</sup>, Yu Sun<sup>5,6</sup>, Tatiana M. Foroud<sup>4</sup>, Karol Estrada<sup>7,9</sup>, Liana G. Apostolova<sup>3</sup>, Qingqin S. Li<sup>5,6</sup>, Andrew J. Saykin<sup>3</sup> and for the Alzheimer's Disease Neuroimaging Initiative (ADNI)

## Abstract

**Background:** Alzheimer's disease (AD) is a chronic progressive neurodegenerative disease impacting an estimated 44 million adults worldwide. The causal pathology of AD (accumulation of amyloid-beta and tau), precedes hallmark symptoms of dementia by more than a decade, necessitating development of early diagnostic markers of disease onset, particularly for new drugs that aim to modify disease processes. To evaluate differentially methylated positions (DMPs) as novel blood-based biomarkers of AD, we used a subset of 653 individuals with peripheral blood (PB) samples in the Alzheimer's disease Neuroimaging Initiative (ADNI) consortium. The selected cohort of AD, mild cognitive impairment (MCI), and age-matched healthy controls (CN) all had imaging, genetics, transcriptomics, cerebrospinal protein markers, and comprehensive clinical records, providing a rich resource of concurrent multi-omics and phenotypic information on a well-phenotyped subset of ADNI participants.

(Continued on next page)

\* Correspondence: [aparna.vasanthakumar@abbvie.com](mailto:aparna.vasanthakumar@abbvie.com)

Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database ([adni.loni.usc.edu](http://adni.loni.usc.edu)). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: [http://adni.loni.usc.edu/wp-content/uploads/how\\_to\\_apply/ADNI\\_Acknowledgement\\_List.pdf](http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf)

† Aparna Vasanthakumar, Justin W. Davis, Kenneth Idler and Jeffrey F. Waring contributed equally to this work.

<sup>1</sup>Genomics Research Center, AbbVie, North Chicago, IL, USA

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

(Continued from previous page)

**Results:** In this manuscript, we report cross-diagnosis differential peripheral DNA methylation in a cohort of AD, MCI, and age-matched CN individuals with longitudinal DNA methylation measurements. Epigenome-wide association studies (EWAS) were performed using a mixed model with repeated measures over time with a  $P$  value cutoff of  $1 \times 10^{-5}$  to test contrasts of pairwise differential peripheral methylation in AD vs CN, AD vs MCI, and MCI vs CN. The most highly significant differentially methylated loci also tracked with Mini Mental State Examination (MMSE) scores. Differentially methylated loci were enriched near brain and neurodegeneration-related genes (e.g., *BDNF*, *BIN1*, *APOC1*) validated using the genotype tissue expression project portal (GTex).

**Conclusions:** Our work shows that peripheral differential methylation between age-matched subjects with AD relative to healthy controls will provide opportunities to further investigate and validate differential methylation as a surrogate of disease. Given the inaccessibility of brain tissue, the PB-associated methylation marks may help identify the stage of disease and progression phenotype, information that would be central to bringing forward successful drugs for AD.

**Keywords:** Alzheimer's disease, Peripheral blood, Biomarker, DNA methylation, ADNI

## Background

Nearly 44 million people worldwide have Alzheimer's disease (AD) or a related dementia, with global costs of the disease estimated to be approximately \$600 billion in 2016 and steadily increasing as the population ages, making it a major public health issue [1, 2]. The hallmark symptoms of AD include memory impairment and cognitive decline, both of which largely drive clinical diagnosis. Existing therapies do not treat the underlying cause of the disease, and only temporarily help relieve memory and cognitive problems. There are several drugs currently under development which aim to modify the disease process; however, there still exists a lack of understanding regarding the molecular mechanisms underlying the disease, thus making it challenging to identify new targets for therapy. Accurate diagnosis of prodromal AD is essential to starting treatments at the right time, and in treating the disease more effectively [3]. The identification of a robust, prodromal, and easily accessible biomarker has been of major interest in the field.

The Alzheimer's Disease Neuroimaging Initiative (ADNI) was launched in 2003 with the goal to establish an optimal panel of clinical assessments: imaging measures (MRI, PET) and biomarkers from blood and cerebrospinal fluid (CSF) to direct clinical trial design for AD drugs [4, 5]. We sought to use this resource to determine if epigenetic markers in PB could serve as biomarkers of AD.

Epigenetic modifications are inheritable and dynamic, and may lead to the regulation of gene expression via modifications to the cytosine residues and/or proteins associated with nucleosome assembly and function [6]. Methylation of the DNA cytosine bases has been studied for several decades and studies have associated methylation at promoter regions with repression of gene

expression [7]. DNA methylation changes as a result of mutations in the DNA methyltransferase-1-enzyme have been shown to be associated with several neuronal diseases including hereditary sensory and autonomic neuropathy-1, in which patients display disrupted methylation patterns potentially contributing to neurodegeneration [8]. De novo mutations of MeCP2, a methyl CpG-binding protein, are linked to Rett syndrome, a progressive neurodevelopmental disorder [9]. Other epigenetic mechanisms link exposures during the course of life such as nutrition, chemical and emotional environments, pregnancy conditions, drug intake, and social status to long-term health of the individual [10, 11]. These observations and others support the significance of DNA methylation and associated machinery in the temporal control of neural stem cell differentiation, neurodevelopment, and neurodegeneration.

Several studies have observed widespread alterations in DNA cytosine methylation patterns both at the global level as well as at the individual loci in AD brains (reviewed in [12–15]). In 2014, two seminal papers identified DNA methylation patterns that characterize AD brains and correlate with progression as defined by their Braak stages [16, 17]. Given that observed differences in DNA methylation levels across tissues are stable in a healthy individual, and may be exploited to determine early changes associated with disease processes [18, 19], we sought to understand patterns of peripheral blood DNA methylation in the ADNI cohort. Our objectives from this study were to (1) generate a public resource for peripheral DNA methylation marks in a cohort of cognitively normal, MCI, and AD patients; (2) to identify cross-sectional differences in peripheral blood DNA methylation associated with mild cognitive impairment (MCI) and AD patients relative to cognitively normal controls (CN); and (3) identify novel non-invasive

disease biomarkers. This information would also help identify subjects who are more susceptible to disease progression. Our goal is to gain a broader understanding of how peripheral DNA methylation differences correlate with the diagnosis of and progression of Alzheimer's disease and to enable the research and clinical community to leverage these results to assess the potential for use of methylation changes as pharmacodynamic or disease modifying biomarkers.

## Results

### Making available a robust resource for DNA methylation differences in the peripheral blood of Alzheimer's disease patients

A total of 1920 samples from 653 individual subjects (CN, MCI, AD) were analyzed using the Illumina EPIC arrays (Table 1). Two experimental factors were considered for patient selection: (1) time- our ability to capture the longitudinal aspect of the study (patients with samples at two or more visits), and (2) diagnosis and its time-varying nature (patients converting from CN to MCI, CN to AD, or MCI to AD). Details of patient selection are included in the "Methods" section. The current study focuses on differential methylation analysis of subjects based on diagnosis. One hundred and ninety-nine duplicates and a single triplicate were included amongst the samples that were run on the EPIC arrays for technical replication but are not used in the final analysis here.

### Distribution of differentially methylated positions (DMPs) is consistent across each cross-diagnosis comparison

After extensive quality control evaluation to filter poor probes and low-quality samples, the data were normalized and M-values (i.e., the logit of the beta values) were used for all further analyses. We analyzed differential DNA methylation across diagnosis groups using a mixed model with a random effect to account for within-subject dependency as detailed in the methods section.

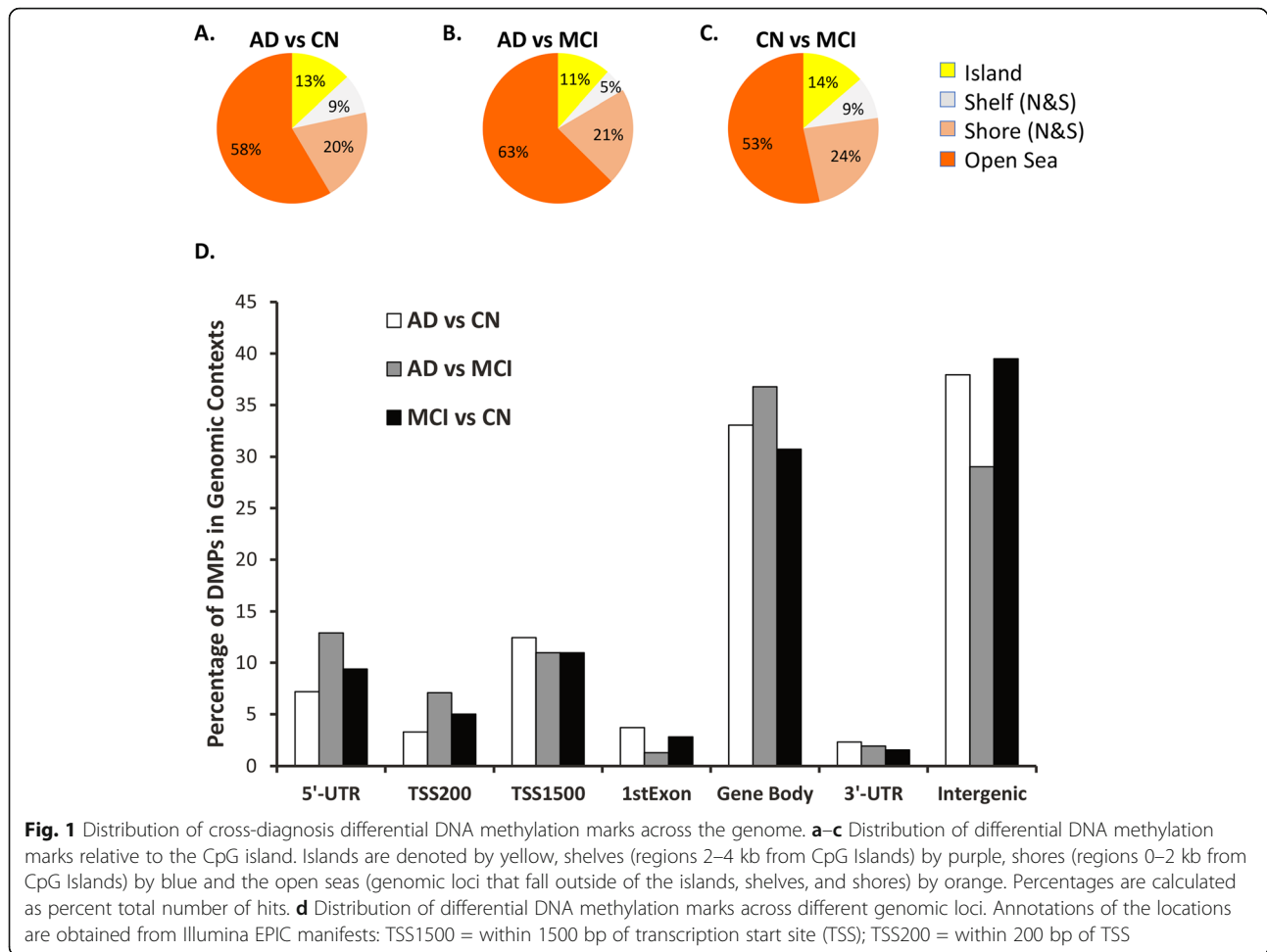
This allowed us to include all available time points for all subjects. The model included covariates to adjust for age at diagnosis, sex, educational attainment, and peripheral blood cell composition, and this yielded 260, 91, and 137 DMPs, respectively, for the three clinical phenotypic comparisons: AD vs. CN, AD vs. MCI, and MCI vs. CN, with a  $p$  value threshold of  $1 \times 10^{-5}$  (Table S2). The majority of the DMPs were clustered within the open seas (genomic loci that fall outside of the CpG islands), and the adjacent shores (regions 0–2 kb from CpG islands), and shelves (regions 2–4 kb from CpG islands) (Fig. 1a–c). The relative levels of enrichment of specific genomic regions (e.g., gene body, 5'-UTR) within the DMP list from three comparisons were similar and did not show significant differences (Fig. 1d).

### DMPs from each pairwise comparisons are enriched for brain-related pathways

There were 42 DMPs that cleared the  $p$ -value of  $1 \times 10^{-5}$  in the AD vs CN comparison (Fig. 2a). The DMP that was most significantly associated with AD relative to CN was annotated to *FAM8AI*, which encodes a protein that is associated with endoplasmic reticulum-associated degradation of proteins with roles in Alzheimer's disease pathogenesis (Fig. 2b). Additionally, when we interrogated the genes located closest to the top DMPs using Tissue Specific Expression Analysis (TSEA), a web-based tool designed to look for tissue-specific expression patterns across 25 different tissue types via GTex Data [20, 21], we observed enrichment for brain-specific genes ( $P_{\text{adj-val}} = 9 \times 10^{-4}$ ) (Figure S3A, Table S3). Other tissues that showed enrichment for the AD vs. CN comparison included: pituitary ( $P_{\text{adj.-value}} = 0.016$ ) and uterus ( $P_{\text{adj.-value}} = 9 \times 10^{-4}$ ). We measured the correlation of observed differential DNA methylation with a cognitive score, MMSE (the mini-mental status examination) and found a significant ( $p$  value =  $3.8 \times 10^{-5}$ ) correlation of MMSE, with DNA methylation differences at this locus (Fig. 2c). We tested the enrichment

**Table 1** ADNI patient cohort selected for DNA methylation analysis and used for final analysis after normalization and quality control

Starting Patient Cohort (N=653; 1720 DNA samples)					
Diagnosis Groups	Age in years Mean (SD)		Number of Males (%)		Number APOE e4 positive (%)
Cognitively Normal (n=223)	76.23 (6.7)		112 (50.2%)		58 (26%)
Mild Cognitive Impairment (n=336)	72.58 (7.82)		189 (56.3%)		153 (46%)
Alzheimer's Disease (n=94)	77.19 (7.69)		60 (63.8%)		64 (68%)
After Normalization, Quality Control and Removal of Replicates (total 1707)					
Diagnosis at 1 <sup>st</sup> Visit	Number at Visit 1	Number at Visit 2	Number at Visit 3	Number at Visit 4	Number at Visit 5
Cognitively Normal	220	200	162	15	4
Mild Cognitive Impairment	333	312	235	23	4
Alzheimer's Disease	94	53	51	-	-

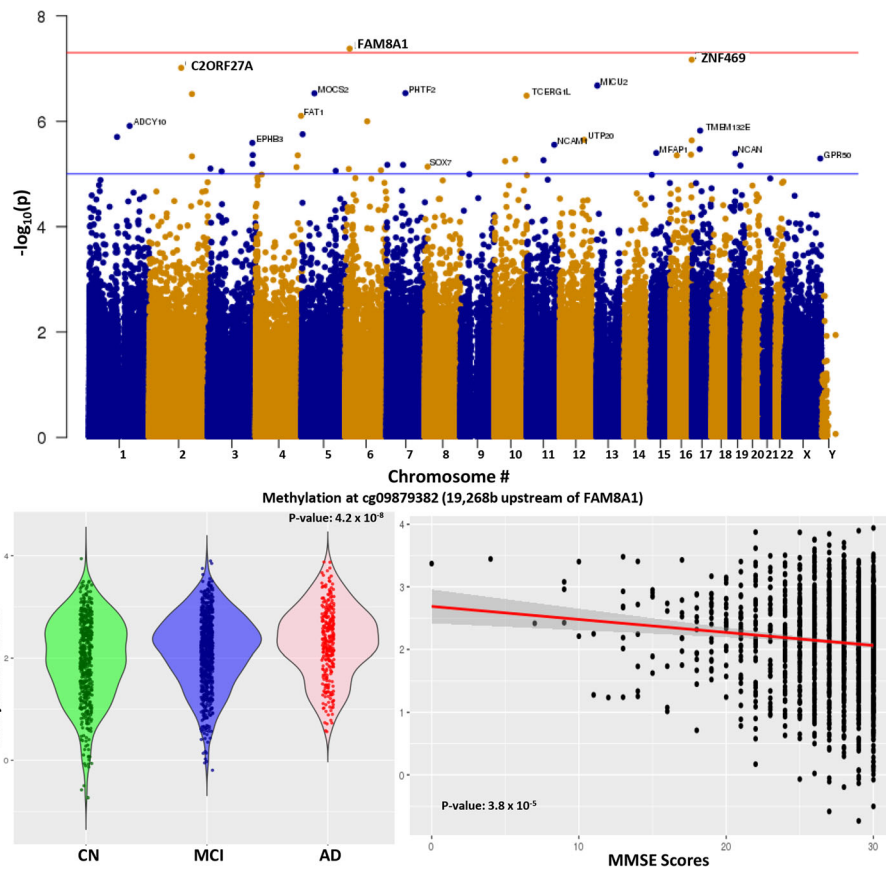


of neural gene expression in parallel using gene ontology analysis, which identified neurogenesis and neuronal differentiation as some of the most highly enriched pathways in the AD vs. CN annotated DMPs (Table 2).

In a similar way, we identified differential methylation from the MCI vs CN comparison, which yielded 25 DMPs at a  $p$  value threshold of  $1 \times 10^{-5}$  (Fig. 3a). The DMP that had the strongest association with MCI vs CN was annotated to *CLIP4* (Fig. 3b). The clustering of the methylation signal correlates with the presence of a SNP at the CpG or within the probe that appears to differentially correlate with disease status. *CLIP4* is a member of the CAP-Gly Domain Containing Linker Protein Family, an important paralog of which, *CLIP3*, is associated with microtubule binding. Again, TSEA analysis identified enrichment of the brain-specific signals ( $P_{\text{adj.}}\text{-value} = 0.0007$ ). There was also a significant ( $p$  value =  $2.0 \times 10^{-5}$ ) correlation of MMSE score with DNA methylation differences at this locus (Fig. 3c). We also found neurogenesis, cell projection, and brain-specific high CpG-rich

promoters as some of the most highly enriched pathways/components when the MCI vs. CN DMPs were annotated (Table 2).

Differential methylation analysis of the AD vs MCI comparison yielded 13 DMPs that were significant (Fig. 4a). The strongest associated DMP was annotated to *NUCB2* (nucleobindin 2), a calcium ion binding protein that regulates intracellular calcium levels. Given the small number of hits, TSEA showed no enrichment of brain-specific pathways, but a slight enrichment of lung-related pathways (Figure S3C, Table S3). There was a significant ( $p$  value =  $4 \times 10^{-4}$ ) correlation of MMSE score with DNA methylation differences at this locus (Fig. 4c). Interestingly, parallel testing in gene ontology analysis showed enrichment of genes that are downregulated in Alzheimer's disease as well as cell projections, and neuronal pathways (Table 2). In addition, *BINI*, *BDNF*, and *APOC1* while not the top most differentially methylated hits, were among the significant DMP hits (Figure S4A–C).



**Fig. 2** Comparison of DNA methylation in AD (Alzheimer's disease) vs CN (cognitively normal). **a** Manhattan plot showing the top hits in the AD vs CN comparison. The blue line indicates  $p$  value threshold of  $1 \times 10^{-5}$  and the red line indicates  $p$  value threshold of  $1 \times 10^{-7}$ . **b** Distribution of unadjusted *M* values in *FAM8A1*, the top DMP across CN (green), MCI (blue), and AD (red). Violin plots outline the spread of the data. **c** Correlation of MMSE scores with differential methylation at the *FAM8A1* locus

### Deriving genetic information from differential DNA methylation signals

Several studies have found an association of genetic variants with the DNA methylation signals at specific probes [22, 23]. To further evaluate the likelihood of DMPs correlating with AD, we also queried all the DMP-associated genes within the GWAS catalog for AD, (<https://ebi.ac.uk/gwas/>) which includes 72 individual GWAS studies and found overlaps between the GWAS hits and the DMPs from AD vs. CN, AD vs. MCI, and MCI vs. CN comparisons (Figure S5A). Some of the overlaps included *BIN1* (Figure S5C), *KCNN2* (Figure S5B), *DIP2C* (Figure S5C), *PAK2* (Figure S5D), *C3orf67* (Figure S5E), and *WNT3* (Figure S5F). We were also able to utilize the methylation data to identify disease-specific associations with some novel SNPs previously linked to neurodevelopmental and neuropsychiatric disorders. For example, *ANK3* has been associated with mental retardation, *SLC45A1* with intellectual developmental disorder, and *CHI3L1* with schizophrenia (Table S4), suggesting

that differential methylation data may help reveal novel genetic variations that associate with AD. These associations could be interesting hypotheses requiring further testing.

### Replication of differential methylation signals across multiple datasets

Finally, we queried a second dataset for differential methylation at the loci identified in our study. A comprehensive study of about 1628 samples assessed human samples across several different types of tissues, including leukocytes, brain regions, and several cancer tissues [24]. Comparison of differential methylation in leukocytes from 65 healthy control subjects that were age 65 years or older with 35 AD subjects within the aforementioned study identified several DMPs. In an effort to replicate our findings from ADNI peripheral blood, we tested for overlaps across our study and the output from the above and observed overlaps across 11 CpGs (Table 3).



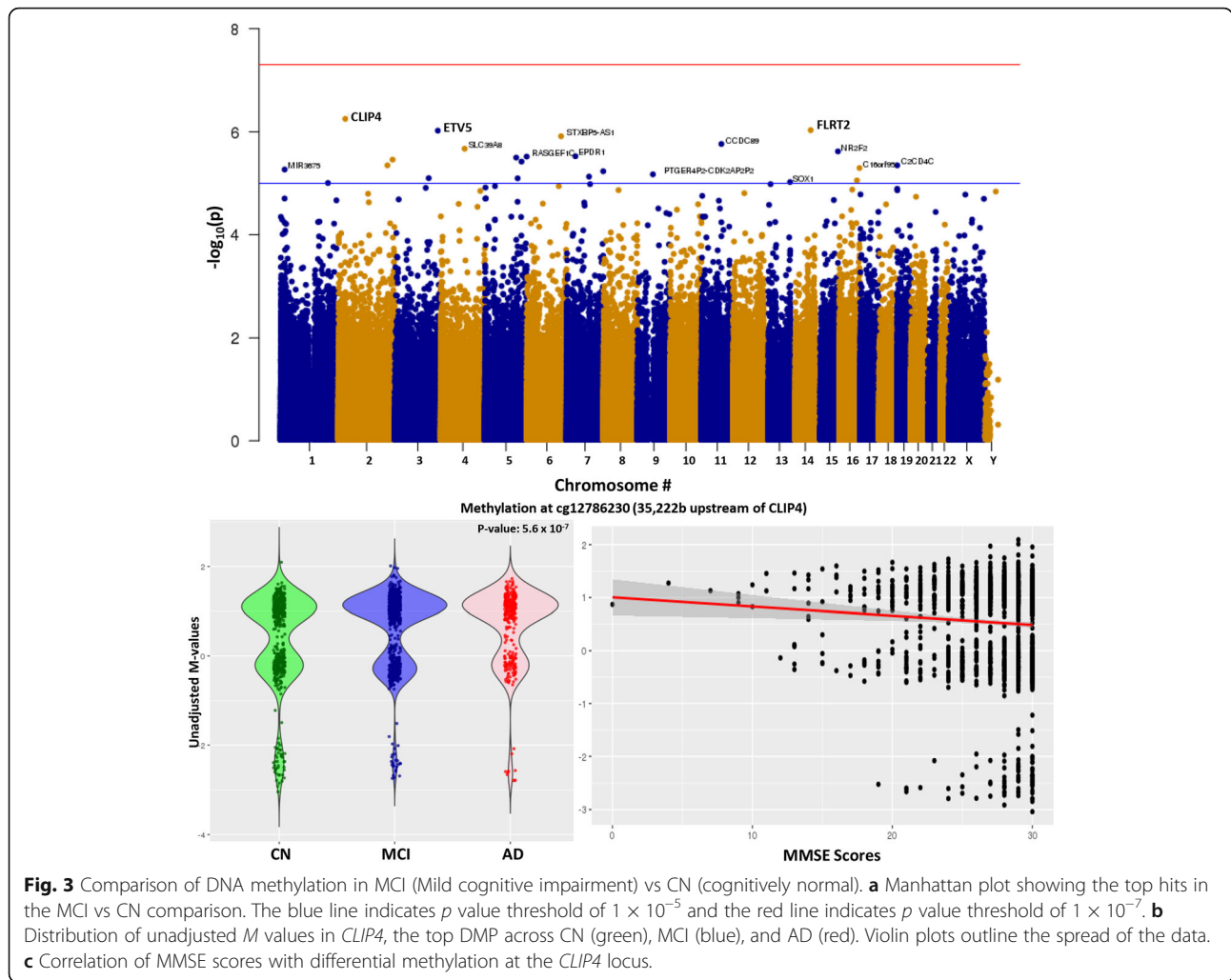
**Table 2** Gene ontology analysis of genes within 50 kb of differentially methylated positions from each cross-diagnosis comparison

Gene Set Name	# Genes in Gene Set (K)	# Genes in overlap (k)	k/K	FDR q-value
Cross-Diagnosis Comparison of AD vs CN at All Visits				
BENPORATH_SUZ12_TARGETS	1038	37	0.0356	1.98E-22
BENPORATH_ES_WITH_H3K27ME3	1118	38	0.034	2.60E-22
BENPORATH_EED_TARGETS	1062	37	0.0348	4.29E-22
<b>GO_NEURON_DIFFERENTIATION</b>	<b>874</b>	<b>29</b>	<b>0.0332</b>	<b>6.78E-17</b>
<b>GO_NEUROGENESIS</b>	<b>1402</b>	<b>35</b>	<b>0.025</b>	<b>2.00E-16</b>
BENPORATH_PRC2_TARGETS	652	25	0.0383	4.04E-16
GO_CELL_DEVELOPMENT	1426	33	0.0231	1.36E-14
MIKKELSEN_NPC_HCP_WITH_H3K27ME3	341	18	0.0528	2.97E-14
MEISSNER_BRAIN_HCP_WITH_H3K4ME3_AND_H3K27ME3	1069	28	0.0262	8.47E-14
MIKKELSEN_MEF_HCP_WITH_H3K27ME3	590	20	0.0339	3.87E-12
Cross-Diagnosis Comparison of MCI vs CN at All Visits				
<b>GO_NEUROGENESIS</b>	<b>1402</b>	<b>18</b>	<b>0.0128</b>	<b>1.35E-09</b>
<b>GO_CELL_PROJECTION</b>	<b>1786</b>	<b>20</b>	<b>0.0112</b>	<b>1.52E-09</b>
MEISSNER_NPC_HCP_WITH_H3_UNMETHYLATED	536	12	0.0224	2.32E-09
BENPORATH_SUZ12_TARGETS	1038	15	0.0145	7.80E-09
<b>MEISSNER_BRAIN_HCP_WITH_H3K4ME3_AND_H3K27ME3</b>	<b>1069</b>	<b>15</b>	<b>0.014</b>	<b>1.15E-08</b>
BENPORATH_ES_WITH_H3K27ME3	1118	15	0.0134	2.08E-08
GO_CELL_DEVELOPMENT	1426	16	0.0112	7.88E-08
BENPORATH_EED_TARGETS	1062	14	0.0132	7.96E-08
<b>GO_CELL_PROJECTION_PART</b>	<b>946</b>	<b>13</b>	<b>0.0137</b>	<b>1.50E-07</b>
GSE40443_INDUCED_VS_TOTAL_TREG_UP	200	7	0.035	3.03E-07
Cross-Diagnosis Comparison of AD vs MCI at All Visits				
GO_INTRINSIC_COMPONENT_OF_PLASMA_MEMBRANE	1649	15	0.0091	7.16E-08
GO_CELL_SURFACE	757	10	0.0132	4.97E-07
GRYDER_PAX3FOXO1_ENHANCERS_IN_TADS	975	11	0.0113	6.04E-07
<b>GO_CELL_PROJECTION</b>	<b>1786</b>	<b>14</b>	<b>0.0078</b>	<b>1.22E-06</b>
BIOCARTA_LAIR_PATHWAY	17	3	0.1765	2.91E-06
GO_ALPHA_ACTININ_BINDING	21	3	0.1429	5.66E-06
<b>BLALOCK_ALZHEIMERS_DISEASE_DN</b>	<b>1237</b>	<b>11</b>	<b>0.0089</b>	<b>5.94E-06</b>
<b>GO_NEURON_PART</b>	<b>1265</b>	<b>11</b>	<b>0.0087</b>	<b>7.34E-06</b>
GO_REGULATION_OF_TRANSPORT	1804	13	0.0072	7.65E-06

## Discussion

We have successfully assayed peripheral blood samples from ADNI to investigate differential DNA methylation in mild cognitively impaired and Alzheimer's disease patients across serial visits using the Illumina EPIC chip. The success rate for the experiment was 99.7%, with only 15 samples out of the total 1920 that failed the run and/or quality control thresholds. Our work establishes the robustness of DNA methylation as a peripheral marker and demonstrates the consistency and reproducibility of its detection at > 99% concordance across replicates.

The cross-diagnosis analysis demonstrates that a common set genomic loci in the periphery are differentially methylated in individuals with AD compared to normal healthy individuals. Several of these differential methylation marks were also replicated in a second peripheral DNA methylation dataset. Additionally, PB DNA methylation differences were found to be enriched near or within genes previously shown to associate with brain-associated pathways. The differential methylation at these sites correlates with cognitive scores, suggesting a relationship between the differential methylation with endophenotypes of disease progression.

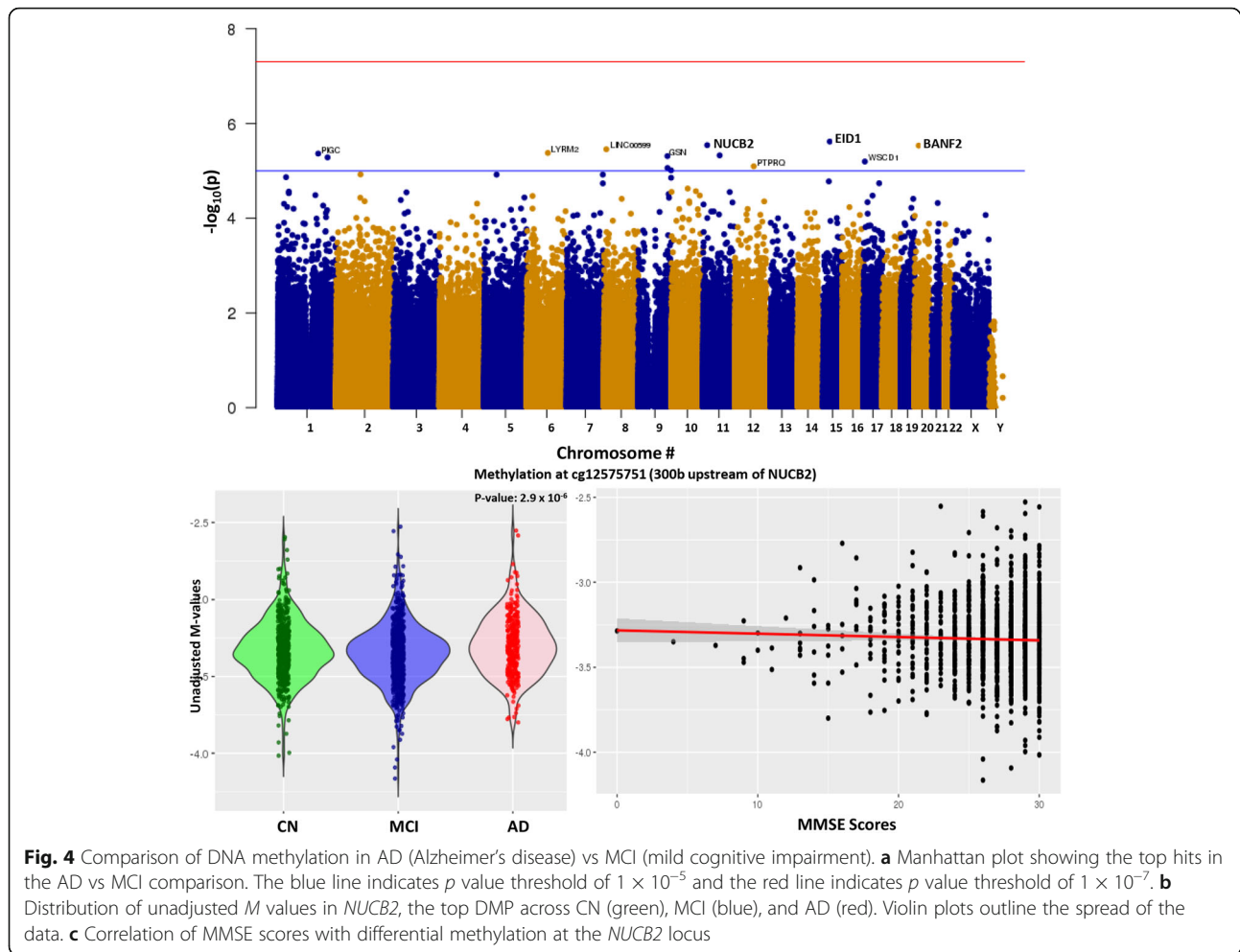


When assessing the overlap in DNA methylation patterns in the periphery and the brain, previous studies have demonstrated that genome-wide DNA methylation profiles are specific to the tissue being studied [16, 25–27]. These studies have suggested that even though many of the DMPs were associated with differentially expressed transcripts, blood-based epigenome-wide association studies from methylation arrays may not correlate with disease etiology [25]. In contrast, some other studies have shown conservation of DNA methylation patterns across blood and brain [18, 28, 29], specifically at promoter regions [18], or via co-expression modules that correlate the brain and the blood to age [29]. Our study picks up some signals in the periphery that are enriched for brain-specific loci; however, this warrants additional studies to detect the blood-brain overlap in DNA methylation. Interestingly, a recent article based on the ENIGMA studies (MRI readouts from 3337 individuals) demonstrated an association of blood DNA

methylation with volumes of the hippocampus, thalamus, and nucleus accumbens (NAcc) [30].

The ADNI participant cohort has previously been used to identify novel biomarkers of disease development and progression [31–33], and is uniquely suited to measure and validate these changes. Ongoing work includes the integration of the methylation data with the rich phenotypic (e.g., cognitive, memory, neuroimaging) and multi-omic data (e.g., genotypic, expression, metabolomics) from the ADNI dataset. This will allow for the use of peripheral DNA methylation marks to function as a dynamic biomarker of disease progression and response to drug treatment.

Peripheral differential methylation has been used as a biomarker of disease occurrence and progression across several therapeutic areas including autoimmune diseases, cancers, and heart disease [34–36]. Previous methylation studies undertaken with PB or peripheral blood mononuclear cell (PBMC) samples mostly



provided a snapshot of DNA methylation changes in the periphery that associated with disease status. A recent study described the identification of PB DNA methylation changes that associated with normal brain aging and cognitive decline in the Whitehall imaging study [37]. For most biomarkers being studied,

longitudinal measures appear to more sensitively predict cognitive decline [38, 39]. Our study design includes longitudinal DNA samples and further analysis will measure dynamic changes in DNA methylation that associate with disease progression. The potential value of DMPs as a surrogate for disease is critically

**Table 3** Replication of differential methylation across datasets (using Fernandez et al. Leukocyte data)

Comparison	Gene	Discovery CpG	Discovery P-value	Replication CpG locus	Replication P-value
AD vs CN	ADCYAP1	cg16288125	$3.02 \times 10^{-5}$	chr18:905,549-905,550	3.78E-09
	EPHB3	cg22462726	$2.57 \times 10^{-6}$	chr3:184,561,230-184,561,231	1.78E-05
	GDF10	cg00414835	$2.5 \times 10^{-5}$	chr10:47,300,094-47,300,095	7.89E-04
	PARP1	cg27113848	$2.74 \times 10^{-5}$	chr1:226,408,700-226,408,701	1.25E-01
MCI vs AD	DAB2IP	cg20416296	$8.69 \times 10^{-6}$	chr9:121,698,699-121,698,700	4.49E-01
	MET	cg04432493	$8.6 \times 10^{-5}$	chr7:116,672,738-116,672,739	7.05E-03
	APOC1	cg07773593	$6.06 \times 10^{-5}$	chr19:44,914,258-44,914,259	9.97E-03
	ITGB1	cg05376034	$6.42 \times 10^{-5}$	chr10:32,958,721-32,958,722	6.20E-02
MCI vs CN	SOX1	cg07911664	$9.5 \times 10^{-6}$	chr13:112,066,581-112,066,582	1.12E-01
	WNT1	cg22376688	$1.56 \times 10^{-5}$	chr12:48,972,583-48,972,584	3.90E-06
	DAPK1	cg10240127	$3.1 \times 10^{-5}$	chr9:87,497,927-87,497,928	8.85E-04



important and can change our approach to clinical studies. Presentation of these results gives the field an opportunity to further investigate and validate the DMPs as surrogates of disease.

## Methods

### Subjects

ADNI is a longitudinal study with approximately 50 sites across the USA and Canada that was launched in 2003 with a major goal being to track the AD progression using clinical and cognitive tests, magnetic resonance imaging (MRI), fludeoxyglucose PET, amyloid PET, cerebrospinal fluid, and blood biomarkers. The institutional review boards of all participating sites reviewed and approved the data collection protocol provided by ADNI. Clinical descriptions of the ADNI cohort have been published [40]. Six hundred and fifty-three individuals from two phases of ADNI (ADNI2 and ADNIGO) were selected for performing DNA methylation analysis (Table 1) based on the completeness of their other datasets (genotyping [APOE, TOMM40], genome wide array, whole genome sequencing, proteomic and imaging data). A total of 1720 samples were obtained, and randomized using a modified incomplete balanced block design, whereby all samples from a subject were on the same chip, with remaining chip space occupied by age-matched samples from a subject of the opposite sex with a different diagnosis. Unused chip space was leveraged for technical reproducibility assessment via replicated DNA samples. A total of 200 samples were replicated across all the chips (Figure S1), for a total 1920 samples processed. Among these replicates, we found consistent DNA methylation signals both within plates and across plates. The correlation coefficient was 99.63% when the replicates were on the same plate with the same scan date, and 99.25% when the replicates were on different plates with different scan dates (Table S1).

### EPIC chip runs

Illumina EPIC chips (Illumina, Inc., San Diego, CA, USA) were used to assay for DNA methylation levels according to published Illumina protocols. Genomic DNA samples obtained from NCRAD (National Centralized Repository for Alzheimer's Disease and Related Dementias) were bisulfite converted using the EZ-DNA Methylation kits (Zymo Research, Irvine, CA, USA) and subsequently analyzed using the Illumina Infinium HD methylation protocol on the HiScan (Illumina).

### Normalization and quality control methods

The derived beta values were transformed to  $M$  values and used for further analysis. The scan output was run through Genome Studio software (Illumina) to assay for

initial QC metrics. One sample out of the total 1920 failed the run and had no CpG calls. The remaining samples had an average of CpG call of 864,640. Four additional samples failed quality control since  $\geq 1\%$  of CpG sites had a detection  $p$  value  $> 0.05$  using watermelon [41]. All 1915 samples were normalized using the dasen method in wateRmelon [41].

### Sample identity checks

Sample sex was examined by computing the ratio of the X and Y probe intensities for each subject compared to their expected value, with  $> 99\%$  of subjects mapping to the given sex (Figure S2A). The following R packages were used to check sample quality and possible sample mix-ups via sex-mismatches: Cham p[42], minfin [43], methylumi [44], and watermelon [41]. Additionally, we used the 59 tracking cpGs on the Illumina EPIC chips which are proxies for SNP fingerprinting (i.e., probe contains C allele that is a common variant), and compared those to the ADNI GWAS genotyping array data at the same positions (Figure S2B) using a clustering algorithm ( $k = 3$ ) to convert cpG signal to genotype based on Hardy-Weinberg equilibrium. The GWAS data were procured from LONI (<http://www.loni.usc.edu/>). After normalization, quality control, and removal of duplicates, 1707 samples were analyzed for differences in DNA methylation.

### Statistical analysis

Since we wanted to include all the samples available for each subject in our initial analysis to compare across diagnoses, we fitted a mixed effects model on the  $M$  values to account for repeated measures of DNA methylation for the patients. This was done using the limma package [45–48] using dupcor estimated at the subject-level. We evaluated the association between DNA methylation level and diagnosis in multivariate models adjusted for age, sex, education, cell composition changes, and DNA storage/source in the model as shown in supplementary material. As it is known that peripheral blood cell composition can substantially affect methylation differences [41] between individuals, differential methylation analysis requires that any change in cell composition be adjusted for. Cell composition estimates were obtained using estimateCellCounts [43] at default settings such that estimates are made for CD8T, CD4T, NK, Bcell, Mono, and Gran. Because they lie in [0,1] and are constrained to sum to 1 within a sample, including all 6 values as covariates would induce multicollinearity. Therefore, only 5 cell type values are used as covariates. Furthermore, the difference in the storage of the sample used for DNA isolation (whole blood vs. buffy coat) had an impact on the cell composition,

prompting its use as an additional covariate, as detailed in Supplementary Material.

### Functional analysis of top differentially methylated positions (DMPs)

Tissue specific analysis of differentially methylated marks was performed using Tissue Specific Expression Analysis (TSEA) at <http://genetics.wustl.edu/jdlab/tsea/> [20]. Gene ontology analysis was performed using the molecular signature database (MSigDB) at <http://software.broadinstitute.org/gsea/index.jsp> [49, 50]. Curated gene sets (Biocarta, KEGG, and Reactome), Gene ontology gene sets (GO biological process, GO cellular component, and GO molecular function), and Immunologic signatures were included in the pathway analysis, and an FDR  $q$  value of 0.05 was set as the threshold.

Additional details regarding statistical analyses are included in supplemental information.

### Supplementary information

**Supplementary information** accompanies this paper at <https://doi.org/10.1186/s13148-020-00864-y>.

**Additional file 1: Fig. S1.** Plate layout after randomization. **Fig. S2.** Quality Control and Confirmation of Sample Identity. **Fig. S3.** Tissue-Specific Expression Analysis identifies Brain-Specific Enrichment. **Fig. S4.** DNA Methylation Differences in Genes Previously Associated with AD Pathogenesis. **Fig. S5.** GWAS Hits Show Concordant DNA Methylation differences. **Table S1.** Correlation of Replicate Samples (200 total technical replicates) within non-normalized data. **Table S2.** Table of all DMPs. **Table S3.** Interpretation of TSEA results.

### Abbreviations

ADNI: Alzheimer's Disease Neuroimaging Initiative; AD: Alzheimer's disease; CN: Cognitively normal; CSF: Cerebrospinal fluid; DMP: Differentially methylated position; EWAS: Epigenome-wide association study; GWAS: Genome-wide association study; MCI: Mild cognitive impairment; MMSE: Mini Mental Status Examination; MRI: Magnetic resonance imaging; NCRAD: National Centralized Repository for Alzheimer's Disease and Related Dementias; NFL: Neurofilament light; OMIM: Online Mendelian Inheritance in Man; PB: Peripheral blood; PBMC: Peripheral blood mononuclear cells; SNP: Single nucleotide polymorphism; TSEA: Tissue Specific Expression Analysis; TSS: Transcriptional Start Site; UTR: Untranslated regions

### Acknowledgements

Samples from the National Cell Repository for Alzheimer's Disease (NCRAD), which receives government support under a cooperative agreement grant (U24 AG024904) awarded by the National Institute on Aging (NIA), were used in this study. We thank contributors who collected samples used in this study, as well as patients and their families, whose help and participation made this work possible.

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.;

NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health ([www.fnih.org](http://www.fnih.org)). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

### Authors' contributions

Conception and design: AV, JWD, KI, JFW, AS, KN, QSL, and KE; Sample acquisition: SK, KEF and TMF; Acquisition of data: KI, EA, and GS; analysis and interpretation: AV, JWD, BRG, SG, GS, KN, KNN, SK, YS, QSL, KE, AS, JFW. All authors read and approved the final manuscript.

### Funding

National Cell Repository for Alzheimer's Disease (NCRAD) is supported by a cooperative agreement grant (U24 AG21886) awarded by the National Institute on Aging (NIA). Funding was also obtained from ADNI (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012).

### Availability of data and materials

The datasets generated and analyzed in the current study are available in the LONI repository (<http://www.loni.usc.edu/>).

### Ethics approval and consent to participate

The institutional review boards of all participating ADNI sites reviewed and approved the data collection protocol provided by ADNI. For up-to-date information, see [www.adni-info.org](http://www.adni-info.org)

### Consent for publication

Not applicable

### Competing interests

AV, JWD, KI, EA, BRG, SG, GS, and JFW are all employees of AbbVie and may hold stock. YS, QSL are employees of Janssen, and KE was an employee of Biogen.

### Author details

<sup>1</sup>Genomics Research Center, AbbVie, North Chicago, IL, USA. <sup>2</sup>Exploratory Statistics, AbbVie, North Chicago, IL, USA. <sup>3</sup>Center for Neuroimaging, Department of Radiology and Imaging Sciences, Indiana University School of Medicine, Indianapolis, IN, USA. <sup>4</sup>National Centralized Repository for Alzheimer's Disease and Related Dementias, Indiana University School of Medicine, Indianapolis, IN, USA. <sup>5</sup>Neuroscience Therapeutic Area, Janssen Research & Development, Pennington, NJ 08534, USA. <sup>6</sup>Research Information Technology, Janssen Research & Development, Pennington, NJ 08534, USA. <sup>7</sup>Biogen, Cambridge, MA 02142, USA. <sup>8</sup>Electrical and Computer Engineering, State University of New York, Oswego, NY 13126, USA. <sup>9</sup>Currently at Biomarin Pharmaceuticals, Novato, CA 94949, USA.

Received: 4 September 2019 Accepted: 14 May 2020

Published online: 15 June 2020

### References

- Prince M, Wimo A, Guerchet M, Ali G-C, Wu Y-T, Prina M, et al. The global impact of dementia an analysis of prevalence, incidence, cost and trends. *World Alzheimer Rep.* 2015;2015.
- Godyn J, Jonczyk J, Panek D, Malawska B. Therapeutic strategies for Alzheimer's disease in clinical trials. *Pharmacol Rep.* 2016;68:127–38.
- Lehmann S, Teunissen CE. Editorial: Biomarkers of Alzheimer's disease: the present and the future. *Front Neurol.* 2016;7:1–3.
- Saykin AJ, Shen L, Yao X, Kim S, Nho K, Risacher SL, et al. Genetic studies of quantitative MCI and AD phenotypes in ADNI: Progress, opportunities, and plans. *Alzheimers Dement.* 2015;11:792–814.
- Weiner MW, Veitch DP, Aisen PS, Beckett LA, Cairns NJ, Green RC, et al. Recent publications from the Alzheimer's Disease Neuroimaging Initiative:

- reviewing progress toward improved AD clinical trials. *Alzheimers Dement*. 2017;13:e1–85 Elsevier.
6. Allis CD, Jenuwein T. The molecular hallmarks of epigenetic control. *Nat Rev Genet*. 2016;17:487–500 Nature Publishing Group.
  7. Razin A, Riggs AD. DNA methylation and gene function. *Science*. 1980;210:604–10.
  8. Klein CJ, Botuyan M, Wu Y, Ward CJ, Nicholson GA, Hammans S, et al. Mutations in DNMT1 cause hereditary sensory neuropathy with dementia and hearing loss. *Chromosome*. 2011;27:819–26.
  9. Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet*. 1999;23:185–8.
  10. Meehan RR, Thomson JP, Lentini A, Nestor CE, Pennings S. DNA methylation as a genomic marker of exposure to chemical and environmental agents. *Curr Opin Chem Biol*. 2018;45:48–56 Elsevier Ltd.
  11. Barouki R, Melén E, Herceg Z, Beckers J, Chen J, Karagas M, et al. Epigenetics as a mechanism linking developmental exposures to long-term toxicity. *Environ Int*. 2018;114:77–86 Elsevier.
  12. Delgado-Morales R, Esteller M. Opening up the DNA methylome of dementia. *Mol Psychiatry*. 2016; In press. Nature Publishing Group.
  13. Roubroeks, Lunnon K. Epigenetics and DNA methylomic profiling in Alzheimer's disease and other neurodegenerative diseases. *J Neurochem*. 2017;38:42–9.
  14. Qazi TJ, Quan Z, Mir A, Qing H. Epigenetics in Alzheimer's disease: perspective of DNA methylation. *Mol Neurobiol*. 2017.
  15. Fransquet PD, Lacaze P, Saffery R, McNeil J, Woods R, Ryan J. Blood DNA methylation as a potential biomarker of dementia: A systematic review. *Alzheimers Dement*. 2017;1–23 Elsevier Inc.
  16. 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;17:1164–70 Nature Publishing Group.
  17. 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;17:1156–63 Nature Publishing Group.
  18. Davies MN, Volta M, Pidsley R, Lunnon K, Dixit A, Lovestone S, et al. Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. *Genome Biol*. 2012;13:R43.
  19. Aryee MJ, Liu W, Engelmann JC, Nuhn P, Gurel M, Haffner MC, et al. DNA methylation alterations exhibit intra-individual stability and inter-individual heterogeneity in prostate cancer metastases. *Sci Transl Med*. 2013;5.
  20. Xu X, Wells AB, O'Brien DR, Nehorai A, Dougherty JD. Cell type-specific expression analysis to identify putative cellular mechanisms for neurogenetic disorders. *J Neurosci*. 2014;34:1420–31.
  21. Dougherty JD, Schmidt EF, Nakajima M, Heintz N. Analytical approaches to RNA profiling data for the identification of genes enriched in specific cells. *Nucleic Acids Res*. 2010;38:4218–30.
  22. Aberg KA, Shabalina AA, Chan RF, Zhao M, Kumar G, van Groenou G, et al. Convergence of evidence from a methylome-wide CpG-SNP association study and GWAS of major depressive disorder. *Transl Psychiatry*. 2018;8 Springer US.
  23. Ciuculete DM, Boström AE, Voisin S, Philipps H, Titova OE, Bandstein M, et al. A methylome-wide mQTL analysis reveals associations of methylation sites with GAD1 and HDAC3 SNPs and a general psychiatric risk score. *Transl Psychiatry*. 2017;7 Nature Publishing Group.
  24. Prosper F, Agirre X, Fernandez AF, Assenov Y, Martin-Subero JJ, Balint B, et al. A DNA methylation fingerprint of 1628 human samples. *Genome Res*. 2012;22:407–19.
  25. 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; 10:1024–32.
  26. Braun PR, Han S, Hing B, Nagahama Y, Gaul LN, Heinzman JT, et al. Genome-wide DNA methylation comparison between live human brain and peripheral tissues within individuals. *Transl Psychiatry*. 2019;9.
  27. Farré P, Jones MJ, Meaney MJ, Emberly E, Turecki G, Kobor MS. Concordant and discordant DNA methylation signatures of aging in human blood and brain. *Epigenetics Chromatin*. 2015;8:1–17.
  28. 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:e1187–10 Nature Publishing Group.
  29. Horvath S, Zhang Y, Langfelder P, Kahn RS, Boks MP, van Eijk K, et al. Aging effects on DNA methylation modules in human brain and blood tissue. *Genome Biol*. 2012;13:R97.
  30. Jia T. Epigenome-wide meta-analysis of blood DNA methylation and its association with subcortical volumes: findings from the ENIGMA Epigenetics Working Group; 2019.
  31. Saykin AJ, Shen L, Foroud TM, Potkin SG, Swaminathan S, Kim S, et al. Alzheimer's disease neuroimaging initiative biomarkers as quantitative phenotypes: genetics core aims, progress, and plans. *Alzheimers Dement*. 2010;6:265–73 Elsevier Ltd.
  32. Horgusluoglu-Moloch E, Nho K, Risacher SL, Kim S, Foroud T, Shaw LM, et al. Targeted neurogenesis pathway-based gene analysis identifies ADORA2A associated with hippocampal volume in mild cognitive impairment and Alzheimer's disease. *Neurobiol Aging*. 2017;60:92–103 Elsevier Inc.
  33. Apostolova LG, Risacher SL, Duran T, Stage EC, Goukasanian N, West JD, et al. Associations of the top 20 Alzheimer disease risk variants with brain amyloidosis. *JAMA Neurol*. 2018;75:328–41.
  34. Pan Y, Liu G, Zhou F, Su B, Li Y. DNA methylation profiles in cancer diagnosis and therapeutics. *Clin Exp Med*. 2018;18:1–14 Springer International Publishing.
  35. Somnineni HK, Venkateswaran S, Kilaru V, Marigorta UM, Mo A, Okou DT, et al. Blood-Derived DNA Methylation Signatures of Crohn's Disease and Severity of Intestinal Inflammation. *Gastroenterology*. 2019;156:2254–2265.e3.
  36. Meder B, Haas J, Sedaghat-Hamedani F, Kayvanpour E, Frese K, Lai A, et al. Epigenome-Wide Association Study Identifies Cardiac Gene Patterning and a Novel Class of Biomarkers for Heart Failure. *Circulation*. 2017;136:1528–44.
  37. 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;10:585–95.
  38. Lo RY, Hubbard AE, Shaw LM, Trojanowski JQ, Petersen RC, Aisen PS, et al. Longitudinal change of biomarkers in cognitive decline. *Arch Neurol*. 2011; 68:1257–66.
  39. Dumurgier J, Hanseeuw BJ, Hatling FB, Judge KA, Schultz AP, Chhatwal JP, Blacker D, Sperling RA, Johnson KA, Hyman BT-GT. Alzheimer's Biomarkers and Future Decline in Cognitive Normal Older Adults. *J Alzheimers Dis*. 2018;60:1451–9.
  40. Petersen RC, Aisen PS, Beckett LA. Alzheimer's Disease Neuroimaging Initiative (ADNI) Clinical characterization; 2011. p. 201–9.
  41. Pidsley R, Wong CCY, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450 K methylation array data. *BMC Genomics*. 2013;14:293.
  42. Tian Y, Morris TJ, Webster AP, Yang Z, Beck S, Feber A, et al. ChAMP: Updated methylation analysis pipeline for Illumina BeadChips. *Bioinformatics*. 2017;33:3982–4.
  43. Fortin JP, Triche TJ, Hansen KD. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. *Bioinformatics*. 2017;33:558–60.
  44. Davis S, Du P, Bilke S, Triche T Jr, Bootwalla M. Methylumi: Handle Illumina methylation data. R package version 2.32.0; 2019.
  45. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43:e47.
  46. Phipson B, Lee S, Majewski IJ, Alexander WS, Smyth G. Robust hyperparameter estimation protects. 2016;10:946–63.
  47. Liu R, Holik AZ, Su S, Jansz N, Chen K, Leong HS, et al. Why weight? Modelling sample and observational level variability improves power in RNA-seq analyses. *Nucleic Acids Res*. 2015;43.
  48. Smyth GK, Michaud J, Scott HS. Use of within-array replicate spots for assessing differential expression in microarray experiments. *Bioinformatics*. 2005;21:2067–75.
  49. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci*. 2005;102:15545–50.
  50. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. *Bioinformatics*. 2011;27:1739–40.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.