

**Novel late-onset Alzheimer's disease loci variants associate with brain gene expression**

Mariet Allen, PhD<sup>1\*</sup>, Fanggeng Zou, PhD<sup>1\*</sup>, High Seng Chai, PhD<sup>2\*</sup>, Curtis S. Younkin<sup>1\*</sup>, Julia Crook, PhD<sup>3</sup>, V. Shane Pankratz, PhD<sup>2</sup>, Minerva M. Carrasquillo, PhD<sup>1</sup>, Christopher N. Rowley<sup>1</sup>, Asha A. Nair<sup>2</sup>, Sumit Middha<sup>2</sup>, Sooraj Maharjan<sup>2</sup>, Thuy Nguyen<sup>1</sup>, Li Ma<sup>1</sup>, Kimberly G. Malphrus<sup>1</sup>, Ryan Palusak<sup>1</sup>, Sarah Lincoln<sup>1</sup>, Gina Bisceglia<sup>1</sup>, Constantin Georgescu, PhD<sup>1</sup>, Debra Schultz<sup>4</sup>, Fariborz Rakhshan<sup>4</sup>, Christopher P. Kolbert<sup>4</sup>, Jin Jen, PhD<sup>4</sup>, Jonathan L. Haines, PhD<sup>5</sup>, Richard Mayeux, MD<sup>6</sup>, Margaret A. Pericak-Vance, PhD<sup>7</sup>, Lindsay A. Farrer, PhD<sup>8</sup>, Gerard D. Schellenberg, PhD<sup>9</sup>, Alzheimer's Disease Genetics Consortium<sup>12</sup>, Ronald C. Petersen, MD, PhD<sup>10</sup>, Neill R. Graff-Radford, MD<sup>11</sup>, Dennis W. Dickson, MD<sup>1</sup>, Steven G. Younkin, MD, PhD<sup>1</sup>, Nilüfer Ertekin-Taner, MD, PhD<sup>1, 11#</sup>

- 1) Mayo Clinic Florida, Department of Neuroscience, Jacksonville, FL.
- 2) Mayo Clinic Minnesota, Department of Biostatistics, Rochester, MN.
- 3) Mayo Clinic Florida, Biostatistics Unit, Jacksonville, FL.
- 4) Mayo Clinic Minnesota, Gene Expression Core, Advanced Genome Technology Center, Rochester, MN.
- 5) Vanderbilt University, Department of Molecular Physiology and Biophysics, and Vanderbilt Center for Human Genetics Research, Nashville, TN.
- 6) Columbia University, Gertrude H. Sergievsky Center, Department of Neurology, and Taub Institute on Alzheimer's Disease and the Aging Brain, New York, NY.
- 7) University of Miami, The John P. Hussman Institute for Human Genomics, and Dr. John T. Macdonald Foundation Department of Human Genetics, Miami, FL.
- 8) Boston University, Departments of Biostatistics, Medicine (Genetics Program), Ophthalmology, Neurology, and Epidemiology, Boston, MA.
- 9) University of Pennsylvania Perelman School of Medicine, Department of Pathology and Laboratory Medicine, Philadelphia, PA.

- 10) Mayo Clinic Minnesota, Department of Neurology, Rochester, MN.
- 11) Mayo Clinic Florida, Department of Neurology, Jacksonville, FL.
- 12) Full list of authors from the Alzheimer's Disease Genetics Consortium (ADGC) and their affiliations are presented below.

**Running head:** Novel LOAD loci and brain gene expression

**Title characters:** 87 (includes spaces)

**Running head characters:** 35 (includes spaces)

\*: Equal contribution.

#: Corresponding author

**Contact information:** Mayo Clinic Florida, Departments of Neurology and Neuroscience, 4500 San Pablo Road, Birdsall 3, Jacksonville, FL 32224. [taner.nilufer@mayo.edu](mailto:taner.nilufer@mayo.edu).

**Number of words in abstract:** 256

**Number of words in body of manuscript (excludes abstract, references, tables):**  
2,896

**Number of tables:** 3, **figures:** 1

**Number of supplementary tables:** 5, **supplementary figures:** 3

**Funding:**

Support for this research was provided by the National Institutes of Health grants: National Institute on Aging (R01 032990 to NET and R01 AG018023 to NRG-R and SGY); Mayo Alzheimer's Disease Research Center: (P50 AG016574 to RCP, DWD, NRG-R, SGY, and NET); Mayo Alzheimer's Disease Patient Registry: (U01 AG006576 to RCP); National Institute on Aging (AG025711, AG017216, AG003949 to DWD). This project was also generously supported by the Robert and Clarice Smith and Abigail Van Buren Alzheimer's Disease Research Program (to RCP, DWD, NRG-R, and SGY), and by the Palumbo Professorship in Alzheimer's Disease Research (to SGY). N.E.T. is the recipient of National Institutes of Health (KL2 RR024151), and Siragusa Foundation grants.

**ADGC Author Funding and Acknowledgements:**

The National Institutes of Health, National Institute on Aging (NIH-NIA) supported this work through the following grants: ADGC, U01 AG032984, RC2 AG036528; NACC, U01 AG016976; NCRAD, U24 AG021886; NIA LOAD, U24 AG026395, U24 AG026390; Boston University, P30 AG013846, R01 HG02213, K24 AG027841, U01 AG10483, R01 CA129769, R01 MH080295, R01 AG009029, R01 AG017173, R01 AG025259; Columbia University, P50 AG008702, R37 AG015473; Duke University, P30 AG028377; Emory University, AG025688; Group Health U01 AG06781; Indiana University, P30 AG10133; Johns Hopkins University, P50 AG005146, R01 AG020688; Massachusetts General Hospital, P50 AG005134; Mayo Clinic, P50 AG016574; Mount Sinai School of Medicine, P50 AG005138, P01 AG002219; New York University, P30 AG08051, MO1RR00096, and UL1 RR029893; Northwestern University, P30 AG013854; Oregon Health & Science University, P30 AG008017, R01 AG026916; Rush University, P30 AG010161, R01 AG019085, R01 AG15819, R01 AG17917, R01 AG30146; University of Alabama at Birmingham, P50 AG016582, UL1RR02777; University of Arizona/TGEN, P30 AG019610, R01 AG031581, R01 NS059873; University of California, Davis, P30 AG010129; University of California, Irvine, P50 AG016573, P50, P50 AG016575, P50 AG016576, P50 AG016577; University of California, Los Angeles, P50 AG016570; University of California, San Diego, P50 AG005131; University of California, San Francisco, P50 AG023501, P01 AG019724; University of Kentucky, P30 AG028383; University of Michigan, P50 AG008671; University of Pennsylvania, P30 AG010124; University of Pittsburgh, P50 AG005133, AG030653; University of Southern California, P50 AG005142; University of Texas Southwestern, P30 AG012300; University of Miami, R01 AG027944, AG010491, AG027944, AG021547, AG019757; University of Washington, P50 AG005136, U01 HG004610; Vanderbilt University, R01 AG019085; and Washington University, P50 AG005681, P01 AG03991. ADNI Funding for ADNI is through the Northern California Institute for Research and Education by grants from Abbott, AstraZeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson and Johnson, Eli Lilly and Co., Medpace, Inc., Merck and Co., Inc., Novartis AG, Pfizer Inc, F. Hoffman-La Roche, Schering-Plough, Synarc, Inc., Alzheimer's Association, Alzheimer's Drug Discovery Foundation, the Dana Foundation, and by the National Institute of Biomedical Imaging and Bioengineering and NIA grants U01 AG024904, RC2 AG036535, K01 AG030514. We thank Creighton Phelps, Marcelle Morrison-Bogorad, and Marilyn Miller from NIA who are *ex-officio* ADGC members. Support was also from the Alzheimer's Association (LAF, IIRG-08-89720; MP-V, IIRG-05-14147) and the Veterans Affairs Administration. P.S.G.-H. is supported by Wellcome Trust, Howard Hughes Medical Institute, and the Canadian Institute of Health Research.

**Disclosures:**

N. Graff-Radford, M.D. has served as a consultant to Codman and received grant support from Elan Pharmaceutical Research, Pfizer Pharmaceuticals, Medivation, and Forrest.

R.C. Petersen, M.D., Ph.D. has been a consultant to GE Healthcare and Elan Pharmaceuticals, has served on a data safety monitoring board in a clinical trial sponsored by Elan Pharmaceuticals, and a safety monitoring board for Wyeth Pharmaceuticals. Please also see below for ADGC Author Disclosures.

### **ADGC Disclosure Statement**

T.D.B. received licensing fees from and is on the speaker's bureau of Athena Diagnostics, Inc. M.R.F. receives research funding from BristolMyersSquibb Company, Danone Research, Elan Pharmaceuticals, Inc., Eli Lilly and Company, Novartis Pharmaceuticals Corporation, OctaPharma AG, Pfizer Inc., and Sonexa Therapeutics, Inc; Receives honoraria as scientific consultant from Accera, Inc., Astellas Pharma US Inc., Baxter, Bayer Pharmaceuticals Corporation, BristolMyersSquibb, Eisai Medical Research, Inc., GE Healthcare, Medavante, Medivation, Inc., Merck & Co., Inc., Novartis Pharmaceuticals Corp., Pfizer, Inc., Prana Biotechnology Ltd., QR Pharma., Inc., The sanofi-aventis Group, and Toyama Chemical Co., Ltd.; and is speaker for Eisai Medical Research, Inc., Forest Laboratories, Pfizer Inc. and Novartis Pharmaceuticals Corporation. A.M.G. has research funding from AstraZeneca, Pfizer and Genentech, and has received remuneration for giving talks at Pfizer and Genentech. R.C.P. is on the Safety Monitoring Committee of Pfizer, Inc. (Wyeth) and a consultant to the Safety Monitoring Committee at Janssen Alzheimer's Immunotherapy Program (Elan), to Elan Pharmaceuticals, and to GE Healthcare. R.E.T. is a consultant to Eisai, Japan in the area of Alzheimer's genetics and a shareholder in, and consultant to Pathway Genomics, Inc, San Diego, CA.

**Author Contributions:**

MA: Drafting/revising the manuscript for content; Study concept; Analysis or interpretation of data; FZ, HSC, CSY, JC, VSP, MMC: Drafting/revising the manuscript for content; Analysis or interpretation of data; CNR, AAN, S. Middha, S. Maharjan, TN, LM, KGM, RP, SL, GB, CG, DS, FR: Analysis or interpretation of data; CPK, JJ, JLH, RM, MAPV, LAF, GDS: Drafting/revising the manuscript for content; Analysis or interpretation of data; ADGC, RCP, NRG-R, DWD: Drafting/revising the manuscript for content; SGY, NET: Drafting/revising the manuscript for content; Study concept; Analysis or interpretation of data

### Abstract

**Objective:** Recent genome-wide association studies (GWAS) of late-onset Alzheimer's disease (LOAD) identified nine novel risk loci. Discovery of functional variants within genes at these loci is required to confirm their role in AD. Single nucleotide polymorphisms that influence gene expression (eSNPs) constitute an important class of functional variants. We therefore investigated the influence of the novel LOAD risk loci on human brain gene expression.

**Methods:** We measured gene expression levels in the cerebellum and temporal cortex of autopsied AD subjects and those with other brain pathologies (~400 total subjects). To determine whether any of the novel LOAD risk variants are eSNPs, we tested their *cis*-association with expression of six nearby LOAD candidate genes detectable in human brain (*ABCA7*, *BIN1*, *CLU*, *MS4A4A*, *MS4A6A*, *PICALM*) and an additional 13 genes  $\pm 100$ kb of these SNPs. To identify additional eSNPs that influence brain gene expression levels of the novel candidate LOAD genes, we identified SNPs  $\pm 100$ kb of their location and tested for *cis*-associations.

**Results:** *CLU* rs11136000 ( $p=7.81 \times 10^{-4}$ ) and *MS4A4A* rs2304933/rs2304935 ( $p=1.48 \times 10^{-4}$ - $1.86 \times 10^{-4}$ ) significantly influence temporal cortex expression levels of these genes. The LOAD-protective *CLU* and risky *MS4A4A* locus alleles associate with higher brain levels of these genes. There are other *cis*-variants that significantly influence brain expression of *CLU* and *ABCA7* ( $p=4.01 \times 10^{-5}$ - $9.09 \times 10^{-9}$ ), some of which also associate with AD risk ( $p=2.64 \times 10^{-2}$ - $6.25 \times 10^{-5}$ ).

**Conclusions:** *CLU* and *MS4A4A* eSNPs may at least partly explain the LOAD risk association at these loci. *CLU* and *ABCA7* may harbor additional strong eSNPs. These

results have implications in the search for functional variants at the novel LOAD risk loci.

## Introduction

Recent genome-wide association studies (GWAS) of late-onset Alzheimer's disease (LOAD) identified nine novel loci, in addition to Apolipoprotein E (APOE)<sup>3, 5-9</sup>. Despite the success of disease GWAS, there remains a knowledge gap with this approach. First, loci identified by disease GWAS may harbor more than one candidate gene<sup>11</sup>. Second, the mechanism of action of the risk variants at the disease loci is not immediately obvious<sup>3, 11</sup>. Third, variants identified in complex disease GWAS do not explain all their underlying genetic component<sup>12, 13</sup>.

An alternative approach is using biologically-relevant, genetically-driven, quantitative phenotypes, i.e. endophenotypes. Gene expression levels are a special group of endophenotypes with a substantial genetic component<sup>11, 14, 15, 16-21</sup>. Identification of variants that influence both gene expression and human disease can discover the actual risk genes at the GWAS loci and their potential mechanism of action<sup>11, 14, 15, 17, 18, 20</sup>.

We postulate that LOAD has multiple risk variants some of which influence gene expression (eSNPs). In this study, we assessed the novel LOAD GWAS loci, first to determine if any of the "top LOAD risk SNPs" influence brain expression of nearby genes *in-cis*. We arbitrarily defined *cis*-associations as those between a transcript and a SNP, which is within the gene encoding that transcript or within 100kb of its 5' start or 3' end sites. Second, we aimed to determine if the brain expression levels of any of the "top LOAD candidate genes" at these loci are influenced by any other eSNPs *in-cis* (*cis*-eSNPs). In this study, we utilize the term eSNPs to define variants associating with expression levels and not to indicate direct evidence of functionality, as the eSNPs are likely markers of other functional, regulatory variants. To determine whether any of the



significant eSNPs that are identified in this second aim also influence AD risk, we utilized meta-analyses results from the LOAD risk GWAS of the Alzheimer's Disease Genetics Consortium (ADGC)<sup>9</sup>. Our findings have implications about the potential mechanism of action for some of the "top LOAD risk SNPs" and for functional regulation of some of the "top LOAD candidate genes".

## Subjects and Methods

### Subjects:

We measured gene expression levels from temporal cortex (TCX) of 399 subjects and from cerebellar tissue (CER) of 374 subjects (**Table 1**). 340 subjects had both TCX and CER measurements. All subjects were participants in the published Mayo LOAD GWAS<sup>22</sup> as part of the autopsy-based series (AUT\_60-80). All subjects had neuropathologic evaluation by DWD. All AD (n=202 for TCX and 197 for CER) subjects had definite diagnosis of AD according to the NINCDS-ADRDA criteria<sup>23</sup> and had Braak scores of  $\geq 4.0$ . All control (non-AD, n=197 for TCX and 177 for CER) subjects had Braak scores of  $\leq 2.5$ , but many had brain pathology unrelated to AD (**Supplementary Table 1**).

### RNA extraction and gene expression measurements:

Total RNA was extracted from frozen TCX or CER samples using the AB (Applied Biosystems) RNA Chemistry for Tissue Samples and an AB PRISM™ 6100 Nucleic Acid PrepStation according to the manufacturer's instructions. The quantity and

quality of the RNA samples were determined by the Agilent 2100 Bioanalyzer using the Agilent RNA 6000 Nano Chip.

Transcript levels were measured using the Whole Genome DASL assay (Illumina, San Diego, CA). The RNA samples were randomized across the chips and plates using a stratified approach to ensure balance with respect to diagnosis, age, gender, RNA Integrity Numbers (RINs) and APOE genotype. Replicate samples were utilized for quality control (QC). Raw probe-level mRNA expression data of the 399 TCX and 374 CER samples were exported from GenomeStudio software (Illumina Inc.) for preprocessing with background correction, variance stabilizing transformation, quantile normalization and probe filtering using the lumi package of BioConductor<sup>24, 25</sup> (**Supplementary Methods**).

#### **Standard Protocol Approvals, Registrations, and Patient Consents:**

This study was approved by the appropriate institutional review board and appropriate informed consent was obtained from all participants.

#### **Genotype data:**

All of the genotypes used in the SNP/transcript level association studies were extracted from the Mayo LOAD GWAS, since all autopsied subjects with TCX and/or CER gene expression measurements were part of this study<sup>22</sup>. The LOAD GWAS genotypes were generated using Illumina's HumanHap300-Duo Genotyping BeadChips analyzed with an Illumina BeadLab Station (Illumina, San Diego, CA) at the Mayo Clinic

Genotyping Shared Resource according to the manufacturer's protocols. The LOAD GWAS QC methods were previously published<sup>14, 22</sup> (**Supplementary Methods**).

### **Statistical Methods:**

#### *SNP/transcript levels association:*

To test the influence of the “top LOAD risk SNPs” on brain gene expression *in-cis*, we extracted from the Mayo LOAD GWAS the genotypes for these SNPs or their proxies. The Mayo LOAD GWAS had genotypes for the following five “top LOAD risk SNPs”: *BINI*\_rs744373, *CLU*\_rs1136000, *EPHA1*\_rs11767557, *MS4A6A*\_rs610932 and *PICALM*\_rs3851179; and the proxies for *ABCA7*\_rs3764650, *MS4A4A*\_rs4938933, *MS4A6A*\_rs610932, *MS4A4E*\_rs670139 and *PICALM*\_rs561655 (**Supplementary Table 2**). The top “LOAD risk SNPs” or their proxies were tested for *cis*-association with the transcript levels of those genes that resided within  $\pm 100$  kb of their genomic location according to NCBI Build 36. Of the “top LOAD candidate genes” nearby these SNPs, the following had detectable brain expression levels in both TCX and CER: *ABCA7*, *BINI*, *CLU*, *MS4A4A*, *MS4A6A* and *PICALM*. 13 additional nearby genes were also tested for SNP/transcript levels associations. 59 SNP/transcript associations were tested

(**Supplementary Tables 2, 3**).

To test whether the brain expression levels of any of the “top LOAD candidate genes” are influenced by other eSNPs *in-cis* (*cis*-eSNPs), we identified all SNPs in the Mayo LOAD GWAS within  $\pm 100$  kb flanking region of the gene. *ABCA7*, *BINI*, *CLU*, *MS4A4A*, *MS4A6A*, *PICALM* and *CD2AP* were tested for additional eSNPs. A total of 369 SNP/transcript associations were tested (**Supplementary Tables 4, 5**). Study-wide

Bonferroni corrections were done for the total number of SNP/transcript associations tested.

Linear regression analysis to test for SNP/transcript associations were done in PLINK<sup>24</sup>. Preprocessed probe transcript levels were utilized as endophenotypes. Each probe was assessed separately, even though one gene may have multiple probes. An additive model was assumed, with the minor allele dosage (0, 1, 2) as the independent variable, and APOE  $\epsilon$ 4 dosage (0, 1, 2), age at death, gender, PCR plate, RIN, adjusted RIN<sup>2</sup> (defined as  $(RIN - RIN_{mean})^2$ ) as covariates. The CER and TCX results were analyzed separately. The AD and non-AD subjects were analyzed both separately and jointly. The joint analyses included diagnosis as an additional covariate (AD=1, non-AD=0) were considered to be the main analyses. The separate AD and non-AD only analyses are also reported (**Supplementary Tables**).

#### **Alzheimer's Disease Genetics Consortium (ADGC) Meta-Analyses:**

The “top LOAD risk SNPs” that were assessed for their influence on brain gene expression have AD risk association results published in the literature. To determine the AD risk association of the remaining SNPs assessed in this study (i.e. proxies for the top LOAD risk SNPs and the other *cis*-eSNPs), we obtained meta-analyses results from the ADGC<sup>9</sup>. The cohorts that are assessed by ADGC, as well as the methodological details of the meta-analyses are described in detail in a recent publication<sup>9</sup>. Briefly, the meta-analyses of the ADGC dataset results reported here (**Tables 2, 3, Supplementary Tables 2-5**) are generated from the combined analyses of stage 1 and stage 2 cohorts (**Supplementary Methods**), with detailed descriptions provided elsewhere<sup>9, 25</sup>. Stage 1

cohorts are comprised of 8,309 LOAD cases and 7,366 cognitively normal elder controls. Stage 2 has 3,531 LOAD vs. 3,565 control subjects.

Each cohort was tested for AD risk association using a logistic regression approach, assuming an additive model and adjusting for age, sex, APOE  $\epsilon$ 4 dosage and principal components from EIGENSTRAT<sup>26</sup>. The meta-analyses results were generated using the inverse variance method implemented in the software package METAL<sup>27</sup>.

## Results

### ***Cis*-association of the novel “top LOAD risk SNPs” with brain expression levels of genes:**

We evaluated 18 “top LOAD risk SNPs” or their proxies from 6 novel LOAD loci for associations with TCX and CER expression levels of 6 “top LOAD candidate genes” and 13 others within  $\pm 100$ kb of these SNPs. All 19 genes had detectable expression both in TCX and CER of autopsied subjects. Given the assessment of 59 SNP/transcript associations (**Supplementary Table 2**),  $p < 8.47 \times 10^{-4}$  is required to achieve significance after Bonferroni correction. *CLU*\_rs11136000 and two SNPs at the *MS4A* locus (rs2304933 and rs2304935) have significant association with TCX expression levels of *CLU* ( $p = 7.81 \times 10^{-4}$ ) and *MS4A4A* ( $p = 1.48 \times 10^{-4}$  and  $1.86 \times 10^{-4}$ ), respectively (**Table 2**). The *MS4A4A* eSNPs also have nominally significant associations in CER ( $p = 3.65 \times 10^{-2}$ ), but *CLU*\_rs11136000/transcript association is not significant in CER. The significant *MS4A4A* SNP/transcript associations appear to be driven by both AD and non-AD

subjects, however *CLU*\_rs11136000/transcript association achieves significance only in the non-ADs (**Supplementary Tables 2, 3**).

*CLU*\_rs11136000 minor “A” allele associates with both higher TCX expression levels of this gene and lower AD risk (**Table 2, Supplementary Tables 2, 3, Figure 1a**). *CLU*\_rs11136000 reached genome-wide significance in published LOAD GWAS<sup>5, 6</sup> and also has significant association with lower AD risk in the meta-analyses of the ADGC dataset (OR=0.89,  $p=5.23 \times 10^{-7}$ ). The minor alleles of *MS4A4A* SNPs rs2304933 and rs2304935 associate with higher brain expression levels of this and increased AD risk that is nominally significant in the meta-analysis of the ADGC dataset (OR=1.06,  $p=1.11 \times 10^{-2}$ - $1.13 \times 10^{-2}$ ) (**Table 2, Supplementary Table 3, Figures 1a-b**). The *MS4A4A* transcript-associating SNPs rs2304933 and rs2304935 are proxies for rs670139, which is intergenic between *MS4A4E* and *MS4A6A* and which showed genome-wide significant association with increased AD risk<sup>8</sup>.

### ***Cis*-association of SNPs with brain expression levels of novel “top LOAD candidate genes”:**

We evaluated expression levels of 7 novel “top LOAD candidate genes”, at six of the recently identified novel LOAD loci, detectable in TCX and CER of autopsied subjects, for their association with *cis*-eSNPs residing within  $\pm 100$  kb of their flanking region. Given the 369 SNP/transcript associations tested,  $p < 1.36 \times 10^{-4}$  is needed to achieve Bonferroni-corrected study-wide significance. There were 9 significant SNP/transcript associations; six for *CLU* expression in TCX ( $p=4.01 \times 10^{-5}$ - $9.00 \times 10^{-9}$ ), and three for *ABCA7* in both TCX and CER ( $p=3.53 \times 10^{-5}$ - $1.03 \times 10^{-7}$ ) (**Table 3**,

**Supplementary Tables 4, 5**). *ABCA7*\_rs7247087/transcript associations were significant in both brain regions, and the other two *ABCA7* SNPs (rs757232 and rs2072102) that were significant in CER had nominal significance in TCX. The direction and magnitude of the effect on *ABCA7* levels were similar in the two brain regions for all three SNPs. None of the five *CLU* SNPs with significant TCX transcript associations showed nominal significance in CER. All significant *CLU* and *ABCA7* SNP/transcript associations showed effects of similar direction and magnitude in both AD and non-AD brains.

*CLU*\_rs569214 which is associated with higher levels of this gene in TCX has nominally significant association with lower AD risk in the meta-analysis of the ADGC dataset (**Table 3, Supplementary Table 5, Figure 1a**). *ABCA7*\_rs757232 and rs2072102 show significant association with higher *ABCA7* levels in two brain regions as well as significantly higher risk of AD (OR=1.11,  $p=6.25 \times 10^{-5}$ - $6.74 \times 10^{-5}$ ) in the meta-analysis of the ADGC dataset (**Table 3, Supplementary Table 5, Figures 1a, 1b**). None of the other significant eSNPs showed any association with AD risk in this study.

## Discussion

The novel disease loci identified in the recent large disease GWAS provide the opportunity to uncover the pathophysiology of complex diseases, such as LOAD, which may be instrumental in the discovery of novel drug targets and development of early diagnostic tools. Nonetheless, significant progress awaits the identification of the true disease genes at these loci, their functional risk alleles and mechanisms of action. In this study, we used the brain gene expression endophenotype and disease GWAS results in a

combined fashion to test whether any of the “top LOAD risk SNPs” (referred henceforth as “LOAD SNPs” for short) are eSNPs and whether the expression of any of the “top LOAD candidate genes” is influenced by any other *cis*-eSNPs that are within  $\pm 100$ kb of their genomic location.

Our approach revealed several important findings: First, the “LOAD SNPs” at the *CLU* and *MS4A* loci also influence brain gene expression of *CLU* and *MS4A4A* genes, respectively. Second, there are additional variants within the arbitrary 100kb *cis*-region that also influence brain expression of *CLU* and *ABCA7*. Third, these additional, strong *CLU* and *ABCA7* *cis*-eSNPs do not appear to be in strong LD with the “LOAD SNPs” at their respective loci.

The strongest “LOAD SNP” at the *ABCA7* locus, rs3764650 was assessed by its proxy rs375229 in our study and not found to have a significant association with brain *ABCA7* levels. On the other hand, two of the strong *ABCA7* *cis*-eSNPs (rs757232, rs2072102) also have significant AD risk association ( $p=6.25 \times 10^{-5}$ - $6.74 \times 10^{-5}$ ). For the *CLU* locus, the top “LOAD SNP”, rs11136000 also associates with *CLU* levels in the temporal cortex. Of the additional strong *CLU* *cis*-eSNPs identified in this study, rs569214 associates with AD risk with modest nominal significance ( $p=2.64 \times 10^{-2}$ ). That the *ABCA7* and *CLU* loci “LOAD SNPs” and additional *cis*-eSNPs, which also associate with AD risk, are not in strong LD with each other suggests that there may be multiple, independent variants at these loci that influence AD risk. Furthermore, at least some of this AD risk is conferred by regulatory variants which affect brain gene expression levels.

The direction of the *CLU* SNP/transcript and SNP/AD risk associations are biologically congruent. The “top LOAD risk SNP” rs11136000 and the other *CLU*



transcript-associating SNPs have alleles which associate with increased brain levels of this gene and also reduce AD risk. *CLU* encodes for clusterin, which is thought to promote neuroprotection in AD via multi-faceted functions including A $\beta$  clearance, prevention of excessive inflammation, inhibition of apoptosis and clearance of neuronal debris<sup>28</sup>. Thus, genetic variants that influence higher brain *CLU* levels could conceivably lower AD risk.

*MS4A* locus SNPs associate with higher levels of *MS4A4A* and increased AD risk. The functions of the *MS4A* locus genes, including *MS4A4A*, are yet to be established, though they are thought to be members of a family of transmembrane proteins which may be parts of oligomeric cell surface complexes involved in signal transduction<sup>29</sup>. Given that the *MS4A* locus harbors 6 genes in an LD block (*MS4A3*, *MS4A2*, *MS4A6A*, *MS4A4E*, *MS4A4A*, *MS4A6E*), it is not possible to distinguish the actual LOAD risk gene(s) in this region based solely on the disease GWAS findings. Our findings suggest that the *MS4A4A* gene might be the strongest AD candidate at this locus, though they need to be interpreted with caution, since brain expression levels of only *MS4A4A* and *MS4A6A* could be tested. Thus, there may be other *MS4A* genes that are influenced by SNPs at this locus. Indeed, we found nominally significant SNP/transcript associations with *MS4A6A* and the downstream *MS4A7*. Others have previously shown nominally significant rs610932/*MS4A6A* transcript associations in brains of 143 neurologically normal, European subjects<sup>8</sup>, though we were not able to confirm this finding in our larger study.

The *ABCA7* transcript-associating SNPs both increase its brain levels and AD risk. *ABCA7* is member of the ATP-binding cassette (ABC) family of proteins

implicated in lipid metabolism<sup>30</sup>, was shown *in-vitro* to regulate cholesterol efflux, inhibit amyloid precursor protein processing<sup>31</sup>; and also to play a role in the phagocytosis of apoptotic cells<sup>32</sup>. ABCA7 has a splice variant, with a distinct expression profile and which lacks the lipid metabolism functionality<sup>33</sup>. Understanding the implications of ABCA7 expressional regulation in AD risk requires further information about the functions of this protein, and clarification about the brain ABCA7 splice variant that is influenced by the ABCA7 SNPs.

Our results also have implications regarding fine mapping studies aimed at functional variant discovery at the novel LOAD risk loci and suggest that screening only for coding polymorphisms may miss important functional variation. The strongest eSNP associations are in the 5' or 3' regions of the transcripts and sometimes in different genes. Bioinformatics focused on transcriptional regulation sites in non-coding regions within and surrounding the *CLU*, *MS4A* and *ABCA7* genes, followed by targeted sequencing and *in-vitro* functional efforts may be fruitful in functional variant discovery at these loci.

It is clear that not all the functional disease variants will be eSNPs. We also note that the strongest LOAD risk SNP was not the strongest eSNP and *vice versa* in this study. This could be due to differences in the populations tested for the LOAD risk and expression associations, sample sizes, LD patterns or a combination of these factors. Nonetheless, it is also possible that the LOAD risk SNPs identified in the disease GWAS may be marking multiple different types of functional variants, i.e. both eSNPs and missense coding variants. Our results suggest that for three of the novel LOAD GWAS loci, eSNPs account for at least part of the disease risk. Not all of the novel LOAD candidate genes were detectable in both CER and TCX of our samples with the DASL

microarrays. It will be important to investigate SNP associations for brain expression levels of the remaining LOAD candidate genes using alternative approaches. It is also necessary to characterize in detail the influence of the eSNPs on brain expression levels of any splice variants.

The strengths of our study include combined use of the gene expression endophenotype and disease GWAS, use of a relatively large sample size of both AD and non-AD brain tissue from two different regions, and investigation of both the “top LOAD risk SNPs” as well as other *cis*-eSNPs for their effects on brain expression of seven “top LOAD candidate genes”. Our study demonstrates the power and utility of the gene expression endophenotypes; identifies brain expression changes in *CLU*, *MS4A4A* and *ABCA7* as one potential mechanism of action at these novel LOAD risk loci and provides direction about functional variant discovery at these loci.

	Temporal Cortex		Cerebellum	
	AD	NON-AD	AD	NON-AD
<b>N</b>	202	197	197	177
<b>Females (%)</b>	108 (53%)	78 (40%)	101 (51%)	63 (36%)
<b>Mean Age <math>\pm</math> SD</b>	73.6 $\pm$ 5.5	71.6 $\pm$ 5.6	73.6 $\pm$ 5.6	71.7 $\pm$ 5.5
<b>ApoE4 dose 0 (%)</b>	79 (39%)	146 (74%)	71 (36%)	130 (73%)
<b>ApoE4 dose 1 (%)</b>	96 (48%)	46 (23%)	98 (50%)	43 (24%)
<b>ApoE4 dose 2 (%)</b>	27 (13%)	3 (2%)	28 (14%)	2 (1%)
<b>ApoE4 dose unknown (%)</b>	0 (0%)	2 (1%)	0 (0%)	2 (1%)
<b>Mean RIN <math>\pm</math> SD</b>	6.3 $\pm$ 0.9	6.9 $\pm$ 1.0	7.2 $\pm$ 1	7.2 $\pm$ 1

**Table1. Subjects with temporal cortex and cerebellar expression measurements.** SD=Standard deviation; RIN=RNA Integrity Number

CHR	eSNP	Tested Allele	PROBE	SYMBOL	TX ALL		CER ALL		AD-Locus Association					LD (eSNP & AD SNP)		AD-eSNP Association	
					P	Beta	P	Beta	AD LOCUS	AD SNP	OR	P	Reference	R <sup>2</sup>	D'	OR (95% CI)	P
11	rs2304933	A	ILMN_2370336	MS4A4A	1.48E-04	0.22	3.65E-02	0.09	MS4A4E	rs670139	1.09	1.40E-09	Hollingworth et al.	0.621	0.907	1.06 (1.01 - 1.11)	1.13E-02
11	rs2304935	G	ILMN_2370336	MS4A4A	1.86E-04	0.22	3.65E-02	0.09	MS4A4E	rs670139	1.09	1.40E-09	Hollingworth et al.	0.621	0.907	1.06 (1.01 - 1.11)	1.11E-02
8	rs11136000	A	ILMN_1667058	CLU	7.81E-04	0.17	ns	0.03	CLU	rs11136000	0.84	1.40E-09	Harold/Lambert et al.	na	na	0.89 (0.85 - 0.93)	5.226E-07

**Table 2: Significant expression level associations for the top AD locus SNPs.** The results that have study-wide significance for TCX, CER or both after Bonferroni correction for 59 SNP/transcript association tests are depicted. eSNP=SNP tested for expression level associations. PROBE=DASL expression probe, SYMBOL= gene tested for its brain expression level association with the eSNP. TX ALL = eSNP/expression associations in the TCX of all subjects. CER ALL = eSNP/expression associations in the CER of all subjects. AD - Locus Association: Published AD risk-Locus association results. AD SNP=Published AD locus SNP with strongest AD risk association. This SNP may be different than the eSNP. LD=linkage disequilibrium. If eSNP and AD SNP are different, then their LD is depicted with R<sup>2</sup> and D'. AD-eSNP association=Association of the eSNP with AD risk in the meta-analysis of the ADGC dataset. Beta=effect size, OR=odds ratio, P=p value, CI=confidence interval, ns=not significant, na=not applicable.

CHR	eSNP	Tested Allele	PROBE	SYMBOL	TX ALL		CER ALL		AD-eSNP Association	
					P	BETA	P	BETA	OR (95%CI)	P
8	rs894019	A	ILMN_1667058	CLU	<b>9.00E-09</b>	<b>0.28</b>	ns	0.01	1.00 (0.95 - 1.05)	ns
8	rs569214	A	ILMN_1667058	CLU	<b>1.98E-06</b>	<b>0.24</b>	ns	-0.02	0.95 (0.91 - 0.99)	2.64E-02
8	rs542876	A	ILMN_1667058	CLU	<b>2.13E-06</b>	<b>0.21</b>	ns	-0.01	na	na
8	rs473024	G	ILMN_1667058	CLU	<b>1.04E-05</b>	<b>0.24</b>	ns	0.06	1.04 (0.99 - 1.10)	ns
8	rs2582369	A	ILMN_1667058	CLU	<b>2.02E-05</b>	<b>0.23</b>	ns	0.07	1.04 (0.98 - 1.09)	ns
19	rs7247087	A	ILMN_1743205	ABCA7	<b>3.53E-05</b>	<b>0.14</b>	<b>1.03E-07</b>	<b>0.18</b>	0.97 (0.91 - 1.03)	ns
8	rs570197	G	ILMN_1667058	CLU	<b>4.01E-05</b>	<b>0.22</b>	ns	0.06	1.03 (0.98 - 1.09)	ns
19	rs757232	A	ILMN_1743205	ABCA7	1.61E-03	0.10	<b>2.62E-06</b>	<b>0.15</b>	1.11 (1.05 - 1.17)	6.25E-05
19	rs2072102	A	ILMN_1743205	ABCA7	1.12E-03	0.11	<b>3.57E-06</b>	<b>0.15</b>	1.11 (1.05 - 1.17)	6.74E-05

**Table 3: Significant associations for the *cis*-eSNPs and brain expression levels of the top AD loci genes.** The results that have study-wide significance for TCX, CER or both after Bonferroni correction for 369 SNP/transcript association tests are depicted. eSNP=SNP tested for expression level associations. PROBE=DASL expression probe, SYMBOL=gene tested for its brain expression level association with the eSNP. TX ALL = eSNP/expression associations in the TCX of all subjects. CER ALL = eSNP/expression associations in the CER of all subjects. AD-eSNP association=Association of the eSNP with AD risk in the meta-analysis of the ADGC dataset. Beta=effect size, OR=odds ratio, P=p value, CI=confidence interval, ns=not significant, na=not available.

### Figure Legends

**Figure 1: Box plots of the residuals for the preprocessed expression values for *ABCA7*, *CLU* and *MS4A4A* genes in the (a) Temporal Cortex and (b) Cerebellum.** Residuals were obtained following adjustment for age at death, gender, *APOE*  $\epsilon 4$  dose, PCR plate, RIN, adjusted RIN<sup>2</sup> and diagnosis in ALL subjects included in the analysis, for each genotype (0,1 or 2 alleles) of the targeted SNP (rs#) in a linear regression model. Median values are represented by a thick, black, horizontal line within the box, whilst the box represents the upper and lower quartiles. The whiskers represent the maximum and minimum values (excluding outliers) defined as 1.5 times the inter-quartile range. Outliers are represented as circles. The subject counts for each genotype are indicated (N=) above each box.

**Supplementary Figures 1-3: Data plots of SNPs tested for association with expression levels of *CLU* (Supplementary Figure 1), *MS4A4A* (Supplementary Figure 2) and *ABCA7* (Supplementary Figure 3) in the Temporal Cortex and Cerebellum.** P-values were transformed using  $-\log_{10}$  and are plotted against the position of each SNP along the chromosome. Genes found within the locus boundaries from the UCSC genome browser (<http://genome.ucsc.edu/>). The LD across the locus is represented by a plot generated with Haploview<sup>34</sup>, using data from the Mayo GWAS<sup>22</sup>. The top reported LOAD SNP at the locus is highlighted on the P-value plot with red circles and a red box around the SNP in the list of rs numbers. The top eSNP in this study is highlighted on the p-value plot by blue squares and a blue box around the SNP in the list of rs numbers. Thirty-five SNPs were tested for *CLU*, 27 for *MS4A4A*, and 23 for *ABCA7*.

**Supplementary Figure 2: Data plot of 27 SNPs tested for association with expression levels of *MS4A4A* in the Temporal Cortex and Cerebellum.** \* Indicates that the SNP highlighted here is actually a proxy for rs4938933, the reported top LOAD SNP at the *MS4A4A* locus. # indicates that the top eSNP highlighted here is also a proxy for the reported top LOAD SNP rs670139 at the *MS4A4E* locus.

**Supplementary Figure 3: Data plot of 23 SNPs tested for association with expression levels of *ABCA7* in the Temporal Cortex and Cerebellum.** \* Indicates that the SNP highlighted here is actually a proxy for rs3764650, one of the reported top LOAD SNPs at the *ABCA7* locus.



**Acknowledgements:**

The authors thank all the study participants and their families, without whom, this work would not be possible. The authors acknowledge members of the Mayo Clinic Jacksonville Memory Disorders Clinic and Mayo Clinic Alzheimer's Disease Research Center for their help in the collection of the samples.

## ADGC Authors

Liana G. Apostolova, MD,<sup>1</sup> Steven E. Arnold, MD,<sup>2</sup> Clinton T. Baldwin, PhD,<sup>3</sup> Robert Barber, PhD,<sup>4</sup> Michael M. Barmada, PhD,<sup>5</sup> Thomas Beach, MD, PhD,<sup>6</sup> Gary W. Beecham, PhD,<sup>7,8</sup> Duane Beekly, BS,<sup>9</sup> David A. Bennett, MD,<sup>10,11</sup> Eileen H. Bigio, MD,<sup>12</sup> Thomas D. Bird, MD,<sup>13</sup> Deborah Blacker, MD,<sup>14,15</sup> Bradley F. Boeve, MD,<sup>16</sup> James D. Bowen, MD,<sup>17</sup> Adam Boxer, MD, PhD,<sup>18</sup> James R. Burke, MD, PhD,<sup>19</sup> Jacqueline Buros, BS,<sup>3</sup> Joseph D. Buxbaum, PhD,<sup>20,21,22</sup> Nigel J. Cairns, PhD, FRCPath,<sup>23</sup> Laura B. Cantwell, MPH,<sup>24</sup> Chuanhai Cao, PhD,<sup>25</sup> Chris S. Carlson, PhD,<sup>26</sup> Regina M. Carney, MD,<sup>27</sup> Steven L. Carroll, MD, PhD,<sup>28</sup> Helena C. Chui, MD,<sup>29</sup> David G. Clark, MD,<sup>30</sup> Jason Corneveaux, BS,<sup>31</sup> Carl W. Cotman, PhD,<sup>32</sup> Paul K. Crane, MD, MPH,<sup>33</sup> Carlos Cruchaga, PhD,<sup>34</sup> Jeffrey L. Cummings, MD,<sup>1</sup> Philip L. De Jager, MD, PhD,<sup>35,36</sup> Charles DeCarli, MD,<sup>37</sup> Steven T. DeKosky, MD,<sup>38</sup> F. Yesim Demirci, MD,<sup>5</sup> Ramon Diaz-Arrastia, MD, PhD,<sup>39</sup> Malcolm Dick, PhD,<sup>32</sup> Beth A. Dombroski, PhD,<sup>24</sup> Ranjan Duara, MD,<sup>40</sup> William G. Ellis, MD,<sup>41</sup> Denis Evans, MD,<sup>42</sup> Kelley M. Faber, MS,<sup>33</sup> Kenneth B. Fallon, MD,<sup>28</sup> Martin R. Farlow, MD,<sup>43</sup> Steven Ferris, PhD,<sup>44</sup> Tatiana M. Foroud, PhD,<sup>33</sup> Matthew P. Frosch, MD, PhD,<sup>45</sup> Douglas R. Galasko, MD,<sup>46</sup> Paul J. Gallins, MS,<sup>7</sup> Mary Ganguli, MD,<sup>47</sup> Marla Gearing, PhD,<sup>48,49</sup> Daniel H. Geschwind, MD, PhD,<sup>50</sup> Bernardino Ghetti, MD,<sup>51</sup> John R. Gilbert, PhD,<sup>7,8</sup> Sid Gilman, MD, FRCP,<sup>52</sup> Bruno Giordani, PhD,<sup>53</sup> Jonathan D. Glass, MD,<sup>54</sup> Alison M. Goate, D.Phil,<sup>34</sup> Robert C. Green, MD,<sup>3,55,56</sup> John H. Growdon, MD,<sup>57</sup> Hakon Hakonarson, MD, PhD,<sup>58</sup> Ronald L. Hamilton, MD,<sup>59</sup> John Hardy, PhD,<sup>60</sup> Lindy E. Harrell, MD, PhD,<sup>30</sup> Elizabeth Head, PhD,<sup>61</sup> Lawrence S. Honig, MD, PhD,<sup>62</sup> Matthew J. Huentelman, PhD,<sup>31</sup> Christine M. Hulette, MD,<sup>63</sup> Bradley T. Hyman, MD, PhD,<sup>57</sup> Gail P. Jarvik, MD, PhD,<sup>64,65</sup> Gregory A. Jicha, MD, PhD,<sup>66</sup> Lee-Way Jin, MD, PhD,<sup>41</sup> Nancy Johnson, PhD,<sup>67</sup> Gyungah Jun, PhD,<sup>68,3,69</sup> M. Ilyas Kamboh, PhD,<sup>5,70</sup> Jason Karlawish, MD,<sup>71</sup> Anna Karydas, BA,<sup>18</sup> John S.K. Kauwe, PhD,<sup>72</sup> Jeffrey A. Kaye, MD,<sup>73,74</sup> Ronald Kim, MD,<sup>75</sup> Edward H. Koo, MD,<sup>46</sup> Neil W. Kowall, MD,<sup>55,76</sup> Patricia Kramer, PhD,<sup>77,73</sup> Walter A. Kukull, PhD,<sup>78</sup> James J. Lah, MD, PhD,<sup>54</sup> Eric B. Larson, MD, MPH,<sup>79</sup> Allan I. Levey, MD, PhD,<sup>54</sup> Andrew P. Lieberman, MD, PhD,<sup>80</sup> Oscar L. Lopez, MD,<sup>70</sup> Kathryn L. Lunetta, PhD,<sup>68</sup> Wendy J. Mack, PhD,<sup>81</sup> Daniel C. Marson, JD, PhD,<sup>30</sup> Eden R. Martin, PhD,<sup>7,8</sup> Frank Martiniuk, PhD,<sup>82</sup> Deborah C. Mash, PhD,<sup>83</sup> Eliezer Masliah, MD,<sup>46,84</sup> Wayne C. McCormick, MD, MPH,<sup>33</sup> Susan M. McCurry, PhD,<sup>85</sup> Andrew N. McDavid, BA,<sup>26</sup> Ann C. McKee, MD,<sup>55,76</sup> Marsel Mesulam, MD,<sup>86,87</sup> Bruce L. Miller, MD,<sup>18</sup> Carol A. Miller, MD,<sup>88</sup> Joshua W. Miller, PhD,<sup>41</sup> Thomas J. Montine, MD, PhD,<sup>89</sup> John C. Morris, MD,<sup>23,90</sup> Amanda J. Myers, PhD,<sup>91</sup> Adam C. Naj, PhD,<sup>7</sup> Petra Nowotny, PhD,<sup>34</sup> Joseph E. Parisi, MD,<sup>92,93</sup> Daniel P. Perl, MD,<sup>94</sup> Elaine Peskind, MD,<sup>95</sup> Wayne W. Poon, PhD,<sup>32</sup> Huntington Potter, PhD,<sup>25</sup> Joseph F. Quinn, MD,<sup>73</sup> Ashok Raj, MD,<sup>25</sup> Ruchita A. Rajbhandary, MPH,<sup>7</sup> Murray Raskind, MD,<sup>95</sup> Eric M. Reiman, MD,<sup>31,96,97,98</sup> Barry Reisberg, MD,<sup>44,99</sup> Christiane Reitz, MD, PhD,<sup>100,101,102</sup> John M. Ringman, MD,<sup>1</sup> Erik D. Roberson, MD, PhD,<sup>30</sup> Ekaterina Rogaeva, PhD,<sup>103</sup> Roger N. Rosenberg, MD,<sup>39</sup> Mary Sano, PhD,<sup>21</sup> Andrew J. Saykin, PsyD,<sup>33,104</sup> Julie A. Schneider, MD,<sup>105,10</sup> Lon S. Schneider, MD,<sup>29,106</sup> William Seeley, MD,<sup>18</sup> Michael L. Shelanski, MD, PhD,<sup>107</sup> Michael A. Slifer, MD, PhD,<sup>7,8</sup> Charles D. Smith, MD,<sup>66</sup> Joshua A. Sonnen, MD,<sup>89</sup> Salvatore Spina, MD,<sup>51</sup> Peter St George-Hyslop, MD, FRCP,<sup>103,108</sup> Robert A. Stern, PhD,<sup>55</sup> Rudolph E. Tanzi, PhD,<sup>57</sup> John Q. Trojanowski, MD, PhD,<sup>24</sup> Juan C. Troncoso, MD,<sup>109</sup> Debby W. Tsuang, MD,<sup>95</sup> Viviana M. Van Deerlin,

MD, PhD,<sup>24</sup> Badri Narayan Vardarajan , MS,<sup>3</sup> Harry V. Vinters, MD,<sup>1,110</sup> Jean Paul Vonsattel, MD,<sup>111</sup> Li-San Wang, PhD,<sup>24</sup> Sandra Weintraub, PhD,<sup>86,87</sup> Kathleen A. Welsh-Bohmer, PhD,<sup>19,112</sup> Jennifer Williamson, MS,<sup>62</sup> Randall L. Woltjer, MD, PhD<sup>113</sup>

**ADGC Author Affiliations**

<sup>1</sup>Department of Neurology, University of California Los Angeles, Los Angeles, California; <sup>2</sup>Department of Psychiatry, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania; <sup>3</sup>Department of Medicine (Genetics Program), Boston University, Boston, Massachusetts; <sup>4</sup>Department of Pharmacology and Neuroscience, University of Texas Southwestern, Fort Worth, Texas; <sup>5</sup>Department of Human Genetics, University of Pittsburgh, Pittsburgh, Pennsylvania; <sup>6</sup>Civin Laboratory for Neuropathology, Banner Sun Health Research Institute, Phoenix, Arizona; <sup>7</sup>The John P. Hussman Institute for Human Genomics, University of Miami, Miami, Florida; <sup>8</sup>Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami, Miami, Florida; <sup>9</sup>National Alzheimer's Coordinating Center, University of Washington, Seattle, Washington; <sup>10</sup>Department of Neurological Sciences, Rush University Medical Center, Chicago, Illinois; <sup>11</sup>Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, Illinois; <sup>12</sup>Department of Pathology, Northwestern University, Chicago, Illinois; <sup>13</sup>Department of Neurology, University of Washington, Seattle, Washington; <sup>14</sup>Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts; <sup>15</sup>Department of Psychiatry, Massachusetts General Hospital/Harvard Medical School, Boston, Massachusetts; <sup>16</sup>Department of Neurology, Mayo Clinic, Rochester, Minnesota; <sup>17</sup>Swedish Medical Center, Seattle, Washington; <sup>18</sup>Department of Neurology, University of California San Francisco, San Francisco, California; <sup>19</sup>Department of Medicine, Duke University, Durham, North Carolina; <sup>20</sup>Department of Neuroscience, Mount Sinai School of Medicine, New York, New York; <sup>21</sup>Department of Psychiatry, Mount Sinai School of Medicine, New York, New York; <sup>22</sup>Departments of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, New York; <sup>23</sup>Department of Pathology and Immunology, Washington University, St. Louis, Missouri; <sup>24</sup>Department of Pathology and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania; <sup>25</sup>Byrd Alzheimer Institute, University of Southern Florida Health, Tampa, Florida; <sup>26</sup>Fred Hutchinson Cancer Research Center, Seattle, Washington; <sup>27</sup>Department of Psychiatry, Vanderbilt University, Nashville, Tennessee; <sup>28</sup>Department of Pathology, University of Alabama at Birmingham, Birmingham, Alabama; <sup>29</sup>Department of Neurology, University of Southern California, Los Angeles, California; <sup>30</sup>Department of Neurology, University of Alabama at Birmingham, Birmingham, Alabama; <sup>31</sup>Neurogenomics Division, Translational Genomics Research Institute, Phoenix, Arizona; <sup>32</sup>Institute for Memory Impairments and Neurological Disorders, University of California Irvine, Irvine, California; <sup>33</sup>Department of Medicine, University of Washington, Seattle, Washington; <sup>34</sup>Department of Psychiatry and Hope Center Program on Protein Aggregation and Neurodegeneration, Washington University School of Medicine, St. Louis, Missouri; <sup>35</sup>Program in Translational NeuroPsychiatric Genomics, Department of Neurology, Brigham and Women's Hospital, Boston, Massachusetts; <sup>36</sup>Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts; <sup>37</sup>Department of Neurology, University of California Davis, Sacramento, California; <sup>38</sup>University of Virginia School of Medicine, Charlottesville, Virginia; <sup>39</sup>Department of Neurology, University of Texas Southwestern, Dallas, Texas; <sup>40</sup>Wien

Center for Alzheimer's Disease and Memory Disorders, Mount Sinai Medical Center, Miami Beach, Florida; <sup>41</sup>Department of Pathology and Laboratory Medicine, University of California Davis, Sacramento, California; <sup>42</sup>Rush Institute for Healthy Aging, Department of Internal Medicine, Rush University Medical Center, Chicago, Illinois; <sup>43</sup>Department of Neurology, Indiana University, Indianapolis, Indiana; <sup>44</sup>Department of Psychiatry, New York University, New York, New York; <sup>45</sup>C.S. Kubik Laboratory for Neuropathology, Massachusetts General Hospital, Charlestown, Massachusetts; <sup>46</sup>Department of Neurosciences, University of California San Diego, La Jolla, California; <sup>47</sup>Department of Psychiatry, University of Pittsburgh, Pittsburgh, Pennsylvania; <sup>48</sup>Department of Pathology and Laboratory Medicine, Emory University, Atlanta, Georgia; <sup>49</sup>Emory Alzheimer's Disease Center, Emory University, Atlanta, Georgia; <sup>50</sup>Neurogenetics Program, University of California Los Angeles, Los Angeles, California; <sup>51</sup>Department of Pathology and Laboratory Medicine, Indiana University, Indianapolis, Indiana; <sup>52</sup>Department of Neurology, University of Michigan, Ann Arbor, Michigan; <sup>53</sup>Department of Psychiatry, University of Michigan, Ann Arbor, Michigan; <sup>54</sup>Department of Neurology, Emory University, Atlanta, Georgia; <sup>55</sup>Department of Neurology, Boston University, Boston, Massachusetts; <sup>56</sup>Department of Epidemiology, Boston University, Boston, Massachusetts; <sup>57</sup>Department of Neurology, Massachusetts General Hospital/Harvard Medical School, Boston, Massachusetts; <sup>58</sup>Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania; <sup>59</sup>Department of Pathology (Neuropathology), University of Pittsburgh, Pittsburgh, Pennsylvania; <sup>60</sup>Institute of Neurology, University College London, Queen Square, London; <sup>61</sup>Department of Molecular and Biomedical Pharmacology, University of California Irvine, Irvine, California; <sup>62</sup>Taub Institute on Alzheimer's Disease and the Aging Brain, Department of Neurology, Columbia University, New York, New York; <sup>63</sup>Department of Pathology, Duke University, Durham, North Carolina; <sup>64</sup>Department of Genome Sciences, University of Washington, Seattle, Washington; <sup>65</sup>Department of Medicine (Medical Genetics), University of Washington, Seattle, Washington; <sup>66</sup>Department of Neurology, University of Kentucky, Lexington, Kentucky; <sup>67</sup>Department of Psychiatry and Behavioral Sciences, Northwestern University, Chicago, Illinois; <sup>68</sup>Department of Biostatistics, Boston University, Boston, Massachusetts; <sup>69</sup>Department of Ophthalmology, Boston University, Boston, Massachusetts; <sup>70</sup>University of Pittsburgh Alzheimer's Disease Research Center, Pittsburgh, Pennsylvania; <sup>71</sup>Department of Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania; <sup>72</sup>Department of Biology, Brigham Young University, Provo, Utah; <sup>73</sup>Department of Neurology, Oregon Health & Science University, Portland, Oregon; <sup>74</sup>Department of Biomedical Engineering, Oregon Health & Science University, Portland, Oregon; <sup>75</sup>Department of Pathology and Laboratory Medicine, University of California Irvine, Irvine, California; <sup>76</sup>Department of Pathology, Boston University, Boston, Massachusetts; <sup>77</sup>Department of Molecular & Medical Genetics, Oregon Health & Science University, Portland, Oregon; <sup>78</sup>Department of Epidemiology, University of Washington, Seattle, Washington; <sup>79</sup>Group Health Research Institute, Seattle, Washington; <sup>80</sup>Department of Pathology, University of Michigan, Ann Arbor, Michigan; <sup>81</sup>Department of Preventive Medicine, University of Southern California, Los Angeles, California; <sup>82</sup>Department of Medicine - Pulmonary, New York University, New York, New York; <sup>83</sup>Department of Neurology, University of Miami, Miami, Florida;

<sup>84</sup>Department of Pathology, University of California San Diego, La Jolla, California;  
<sup>85</sup>School of Nursing Northwest Research Group on Aging, University of Washington, Seattle, Washington; <sup>86</sup>Alzheimer's Disease Center, Northwestern University, Chicago, Illinois; <sup>87</sup>Cognitive Neurology, Northwestern University, Chicago, Illinois; <sup>88</sup>Department of Pathology, University of Southern California, Los Angeles, California; <sup>89</sup>Department of Pathology, University of Washington, Seattle, Washington; <sup>90</sup>Department of Neurology, Washington University, St. Louis, Missouri; <sup>91</sup>Department of Psychiatry & Behavioral Sciences, University of Miami, Miami, Florida; <sup>92</sup>Department of Anatomic Pathology, Mayo Clinic, Rochester, Minnesota; <sup>93</sup>Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota; <sup>94</sup>Department of Pathology, Mount Sinai School of Medicine, New York, New York; <sup>95</sup>Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, Washington; <sup>96</sup>Department of Psychiatry, University of Arizona, Phoenix, Arizona; <sup>97</sup>Arizona Alzheimer's Consortium, Phoenix, Arizona; <sup>98</sup>Banner Alzheimer's Institute, Phoenix, Arizona; <sup>99</sup>Alzheimer's Disease Center, New York University, New York, New York; <sup>100</sup>Taub Institute on Alzheimer's Disease and the Aging Brain, Columbia University, New York, New York; <sup>101</sup>Gertrude H. Sergievsky Center, Columbia University, New York, New York; <sup>102</sup>Department of Neurology, Columbia University, New York, New York; <sup>103</sup>Tanz Centre for Research in Neurodegenerative Disease, University of Toronto, Toronto, Ontario; <sup>104</sup>Department of Radiology and Imaging Sciences, Indiana University, Indianapolis, Indiana; <sup>105</sup>Department of Pathology (Neuropathology), Rush University Medical Center, Chicago, Illinois; <sup>106</sup>Department of Psychiatry, University of Southern California, Los Angeles, California; <sup>107</sup>Department of Pathology, Columbia University, New York, New York; <sup>108</sup>Cambridge Institute for Medical Research and Department of Clinical Neurosciences, University of Cambridge, Cambridge, Massachusetts; <sup>109</sup>Department of Pathology, Johns Hopkins University, Baltimore, Maryland; <sup>110</sup>Department of Pathology & Laboratory Medicine, University of California Los Angeles, Los Angeles, California; <sup>111</sup>Taub Institute on Alzheimer's Disease and the Aging Brain, Department of Pathology, Columbia University, New York, New York; <sup>112</sup>Department of Psychiatry & Behavioral Sciences, Duke University, Durham, North Carolina; <sup>113</sup>Department of Pathology, Oregon Health & Science University, Portland, Oregon.

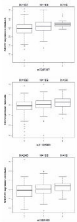
## References

1. Ferri CP, Sousa R, Albanese E, Ribeiro WS, Honyashiki M. World Alzheimer Report 2009 Executive Summary: Alzheimer's Disease International, 2009.
2. Gatz M, Reynolds CA, Fratiglioni L, et al. Role of genes and environments for explaining Alzheimer disease. *Arch Gen Psychiatry* 2006;63:168-174.
3. Ertekin-Taner N. Genetics of Alzheimer disease in the pre- and post-GWAS era. *Alzheimers Res Ther* 2010;2:3.
4. Ertekin-Taner N. Genetics of Alzheimer's disease: a centennial review. *Neurol Clin* 2007;25:611-667.
5. Harold D, Abraham R, Hollingworth P, et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat Genet* 2009.
6. Lambert JC, Heath S, Even G, et al. Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. *Nat Genet* 2009;41:1094-1099.
7. Seshadri S, Fitzpatrick AL, Ikram MA, et al. Genome-wide analysis of genetic loci associated with Alzheimer disease. *JAMA* 2010;303:1832-1840.
8. Hollingworth P, Harold D, Sims R, et al. Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. *Nat Genet* 2011;43:429-435.
9. Naj AC, Jun G, Beecham GW, et al. Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. *Nat Genet* 2011;43:436-441.
10. Harold D, Jehu L, Turic D, et al. Interaction between the ADAM12 and SH3MD1 genes may confer susceptibility to late-onset Alzheimer's disease. *Am J Med Genet B Neuropsychiatr Genet* 2007;144:448-452.
11. Ertekin-Taner N. Gene expression endophenotypes: a novel approach for gene discovery in Alzheimer's disease. *Mol Neurodegener* 2011;6:31.
12. Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. *Science* 2008;322:881-888.
13. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. *Nature* 2009;461:747-753.
14. Zou F, Carrasquillo MM, Pankratz VS, et al. Gene expression levels as endophenotypes in genome-wide association studies of Alzheimer disease. *Neurology* 2010;74:480-486.
15. Carrasquillo MM, Belbin O, Zou F, et al. Concordant association of insulin degrading enzyme gene (IDE) variants with IDE mRNA, Abeta, and Alzheimer's disease. *PLoS One* 2010;5:e8764.
16. Stranger BE, Nica AC, Forrest MS, et al. Population genomics of human gene expression. *Nat Genet* 2007;39:1217-1224.
17. Dixon AL, Liang L, Moffatt MF, et al. A genome-wide association study of global gene expression. *Nat Genet* 2007;39:1202-1207.

18. Goring HH, Curran JE, Johnson MP, et al. Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes. *Nat Genet* 2007;39:1208-1216.
19. Emilsson V, Thorleifsson G, Zhang B, et al. Genetics of gene expression and its effect on disease. *Nature* 2008;452:423-428.
20. Schadt EE, Molony C, Chudin E, et al. Mapping the genetic architecture of gene expression in human liver. *PLoS Biol* 2008;6:e107.
21. Myers AJ, Gibbs JR, Webster JA, et al. A survey of genetic human cortical gene expression. *Nat Genet* 2007;39:1494-1499.
22. Carrasquillo MM, Zou F, Pankratz VS, et al. Genetic variation in PCDH11X is associated with susceptibility to late-onset Alzheimer's disease. *Nat Genet* 2009;41:192-198.
23. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 1984;34:939-944.
24. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81:559-575.
25. Jun G, Naj AC, Beecham GW, et al. Meta-analysis Confirms CR1, CLU, and PICALM as Alzheimer Disease Risk Loci and Reveals Interactions With APOE Genotypes. *Arch Neurol* 2010.
26. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006;38:904-909.
27. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010;26:2190-2191.
28. Nuutinen T, Suuronen T, Kauppinen A, Salminen A. Clusterin: a forgotten player in Alzheimer's disease. *Brain Res Rev* 2009;61:89-104.
29. Liang Y, Buckley TR, Tu L, Langdon SD, Tedder TF. Structural organization of the human MS4A gene cluster on Chromosome 11q12. *Immunogenetics* 2001;53:357-368.
30. Abe-Dohmae S, Ueda K, Yokoyama S. ABCA7, a molecule with unknown function. *FEBS Lett* 2006;580:1178-1182.
31. Chan SL, Kim WS, Kwok JB, et al. ATP-binding cassette transporter A7 regulates processing of amyloid precursor protein in vitro. *J Neurochem* 2008;106:793-804.
32. Jehle AW, Gardai SJ, Li S, et al. ATP-binding cassette transporter A7 enhances phagocytosis of apoptotic cells and associated ERK signaling in macrophages. *J Cell Biol* 2006;174:547-556.
33. Ikeda Y, Abe-Dohmae S, Munehira Y, et al. Posttranscriptional regulation of human ABCA7 and its function for the apoA-I-dependent lipid release. *Biochem Biophys Res Commun* 2003;311:313-318.
34. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21:263-265.



1a



1b

