

OPEN ACCESS

Edited by:

Ellen W. Demerath,
University of Minnesota School of
Public Health, USA

Reviewed by:

Hui-Qi Qu,
The University of Texas School of
Public Health, USA
Andrew DeWan,
Yale School of Public Health, USA

***Correspondence:**

Li Shen,
Center for Neuroimaging, Department
of Radiology and Imaging Sciences,
Indiana University School of Medicine,
355 West 16th Street, Suite 4100,
Indianapolis, IN, USA
shenli@iu.edu

† Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (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.

Specialty section:

This article was submitted to Applied Genetic Epidemiology, a section of the journal *Frontiers in Genetics*

Received: 06 December 2014

Accepted: 09 March 2015

Published: 26 March 2015

Citation:

Yan J, Kim S, Nho K, Chen R, Risacher SL, Moore JH, Saykin AJ and Shen L (2015) Hippocampal transcriptome-guided genetic analysis of correlated episodic memory phenotypes in Alzheimer's disease. *Front. Genet.* 6:117. doi: 10.3389/fgene.2015.00117

Hippocampal transcriptome-guided genetic analysis of correlated episodic memory phenotypes in Alzheimer's disease

Jingwen Yan^{1,2}, Sungeun Kim^{2,3}, Kwangsik Nho^{2,3}, Rui Chen^{2,4}, Shannon L. Risacher^{2,3}, Jason H. Moore⁵, Andrew J. Saykin^{2,3,6}, and Li Shen^{1,2,3,7*} for the Alzheimer's Disease Neuroimaging Initiative †

¹ BioHealth, Indiana University School of Informatics and Computing, Indianapolis, IN, USA, ² Center for Neuroimaging, Department of Radiology and Imaging Sciences, Indiana University School of Medicine, Indianapolis, IN, USA, ³ Indiana Alzheimer Disease Center, Indiana University School of Medicine, Indianapolis, IN, USA, ⁴ Computer Science, Dartmouth College, Hanover, NH, USA, ⁵ Genetics, Community and Family Medicine, Geisel School of Medicine at Dartmouth, Lebanon, NH, USA, ⁶ Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA, ⁷ Center for Computational Biology and Bioinformatics, Indiana University School of Medicine, Indianapolis, IN, USA

As the most common type of dementia, Alzheimer's disease (AD) is a neurodegenerative disorder initially manifested by impaired memory performances. While the diagnosis information indicates a dichotomous status of a patient, memory scores have the potential to capture the continuous nature of the disease progression and may provide more insights into the underlying mechanism. In this work, we performed a targeted genetic study of memory scores on an AD cohort to identify the associations between a set of genes highly expressed in the hippocampal region and seven cognitive scores related to episodic memory. Both main effects and interaction effects of the targeted genetic markers on these correlated memory scores were examined. In addition to well-known AD genetic markers *APOE* and *TOMM40*, our analysis identified a new risk gene *NAV2* through the gene-level main effect analysis. *NAV2* was found to be significantly and consistently associated with all seven episodic memory scores. Genetic interaction analysis also yielded a few promising hits warranting further investigation, especially for the RAVLT list B Score.

Keywords: genetic associate study, correlated phenotypes, hippocampus, episodic memory, Alzheimer's disease

Introduction

Alzheimer's disease (AD) is a neurodegenerative disorder that is initially manifested by impaired memory function. Heritability of AD has been estimated to range from 58 to 79% with the evidence collected from familial aggregation, transmission patterns, and large scale twin studies. The lifetime risk of the first-degree relatives of patients can be twice that of the general population (Gatz et al., 2006; Ertekin-Taner, 2010). Given its high heritability, many genetic studies in AD have been performed, including linkage analysis, candidate gene analysis, and genome-wide association study (GWAS). While apolipoprotein E (*APOE*) has been first found (Saunders et al., 1993) and later frequently replicated in many studies (Shen et al., 2010; Hollingworth et al., 2011; Shi et al., 2012), some other markers have also been reported in recent GWAS studies, such as bridging integrator 1

(*BINI*) (Hu et al., 2011; Lee et al., 2011), clusterin (*CLU*) (Harold et al., 2009; Lambert et al., 2009; Lee et al., 2011), ATP-binding cassette, sub-family A (*ABC1*), member 7 (*ABCA7*) (Hollingworth et al., 2011), complement component (3b/4b) receptor 1 (*CRI*) (Lambert et al., 2009), phosphatidylinositol binding clathrin assembly protein (*PICALM*) (Harold et al., 2009; Lee et al., 2011), EPH receptor A1 (*EPHA1*) (Hollingworth et al., 2011), CD33 molecule (*CD33*) (Hollingworth et al., 2011), Membrane-Spanning 4-Domains (*MS4A4A/MS4A6A*), CD2-associated protein (*CD2AP*) (Naj et al., 2011), and 11 new susceptibility loci recently identified in Lambert et al. (2013).

While traditional genetic association studies mostly focus on analyzing the case control status as the phenotype, they are not designed for revealing genetic risk factors associated with relevant quantitative phenotypes where hidden signals may have appeared long before the disease diagnosis is confirmed. Cognitive measures have been used as highly relevant quantitative traits (QTs) for neuropsychiatric conditions (Gottesman and Gould, 2003; Glahn et al., 2007), including AD (Bennett et al., 2009). These intermediate measures quantitatively capture the progressive nature of the disease and are believed to hold great promise in identifying genetic risk factors. As one of the primary preceding signals, memory decline has been observed in many AD patients prior to diagnostically cognitive, behavioral, and social changes (Backman et al., 2001). Thus, relating genetic markers to memory QTs, rather than diagnosis, has great potential to improve the mechanistic understanding of the pathway from genetics to cognition and then to diagnosis.

Accompanied with advances in high-throughput genotyping techniques, substantial efforts have recently been made to facilitate the reliable identification of memory-related genes. Several genes, including previously identified AD risk genes, have been reported to contribute to the episodic memory disturbances, including *APOE*, *CLU*, *BINI*, brain-derived neurotrophic factor (*BDNF*), and WW and C2 domain containing 1 (*WWC1/KIBRA*) (Egan et al., 2003; Burgess et al., 2011). Recently, a new risk gene, FAST kinase domains 2 (*FASTKD2*), has been reported in the GWAS of a large cohort (Ramanan et al., 2014). However, most of these studies focus only on one specific episodic memory scores. In this study, we aim to identify the genetic factors that are jointly associated with multiple correlated episodic memory scores. This may have the advantage of reducing the biases introduced by variability and outliers in the analysis of a single score, and also to some extent consolidating and integrating the findings.

Existing genetic findings mostly consist of individual single nucleotide polymorphism (SNP) markers or genes; and they typically explain a part of heritability but not all. To search for additional heritability, we extended our analysis to incorporate both single marker tests and pair-wise SNP interaction tests. Given the major combinatorial explosion challenge for SNP-SNP interaction analysis, we performed our analysis in a targeted fashion to make computation feasible. In this work, we only examined the genes that are highly preferred to express in the hippocampal region, which is a critical structure related to learning and memory. We performed an association study to examine the relationships between the SNPs in these genes and seven cognitive scores representing episodic memory, including two scores from

Logical Memory Test, and five scores from the Rey Auditory Verbal Learning Test (RAVLT). We performed multiple comparison correction, by considering both the correlation structure within the genotyping data and that among memory scores.

Materials and Methods

Alzheimer's Disease Imaging Initiative (ADNI)

Genotype and QT data used in this study were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations, as a \$60 million, 5-year public-private partnership. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early AD. Determination of sensitive and specific markers of very early AD progression is critical to aid researchers and clinicians to develop new treatments and monitor their effectiveness, and to reduce the time and cost of clinical trials.

The Principal Investigator of this initiative is Michael W. Weiner, MD, VA Medical Center and University of California, San Francisco. ADNI is the result of efforts of many co-investigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the U.S. and Canada. The initial goal of ADNI was to recruit 800 subjects but ADNI has been followed by ADNI-GO and ADNI-2. To date these three protocols have recruited over 1500 adults, ages 55–90, to participate in the research, consisting of cognitively normal older individuals, people with early or late MCI (EMCI or LMCI), and people with early AD. The follow up duration of each group is specified in the protocols for ADNI-1, ADNI-2, and ADNI-GO. Subjects originally recruited for ADNI-1 and ADNI-GO had the option to be followed in ADNI-2. Thousands of longitudinal imaging scans (Jack et al., 2008; Jagust et al., 2010), performance on neuropsychological and clinical assessments (Petersen et al., 2010) and biological samples (Shaw et al., 2009) were collected at baseline and at follow-up visits for all or a subset of participants. Genome-wide genotyping data (Saykin et al., 2010) are available on the full ADNI sample. For up-to-date information, see www.adni-info.org.

Subjects

To eliminate the possible bias introduced by population stratification, this study was restricted to non-Hispanic Caucasian participants from both the ADNI-1 and ADNI-2/GO cohorts. Subjects in other racial/ethnic groups were excluded in the analysis due to the relative small number of those samples (less than 10%). We employed the population stratification approach used in Kim et al. (2013). Briefly, 988 founders with known ancestry information from HapMap phase 3 (HapMap3) release 2 were used as reference data in the population stratification step and merged with the ADNI samples. Multidimensional

scaling (MDS) was performed using PLINK with identity-by-state (IBS) pairwise distance matrix of the merged data. We compared all ADNI participants with self-reported race/ethnicity as “non-Hispanic/white” with HapMap3 samples in the MDS space, and excluded those participants grouped with HapMap3 samples whose ancestries were neither CEU nor TSI. As a result, 1149 non-Hispanic Caucasian participants were included in this study, and their GWAS data passed the above population stratification and all the other quality control (QC) procedures described in Kim et al. (2013). Shown in **Table 1** is the demographic information for these subjects.

Candidate Gene Selection

Candidate genes were extracted based on the expression profile obtained from the Allen Human Brain Atlas (AHBA) (www.brain-map.org). AHBA provides a comprehensive expression mapping of ~60,000 probes (~30,000 genes and transcripts) across the human brain, with ~1000 brain samples collected for two full brains and ~500 brain samples for the other six half brains. In a recent study, extremely high similarity (95%) among the expression profiles of these brain samples was reported (Zeng et al., 2012), and therefore this study employed only one full brain to identify the candidate genes. Expression data of all probes, targeting 29,196 genes and transcripts, across the whole brain were downloaded. Since the expression data was measured based on the probes and each gene has several probes, we excluded those genes whose multiple probe expression profiles do not correlate very well (i.e., correlation coefficient <0.6). The final expression level of each gene was measured based on the average of highly correlated probes, while other probes were considered as outliers. Each brain sample location in AHBA was mapped back to the MarsBaR AAL atlas with 106 brain ROIs (Tzourio-Mazoyer et al., 2002). We calculated the ROI level expression of each gene by averaging the expression of all brain samples within each ROI. Totally 66 out of 948 brain samples were located within the hippocampal region defined by the MarsBaR AAL atlas. Ultimately 1957 genes with average expression level >8 and the hippocampus as one of its top 1% expressed regions were extracted for further analysis. The overall workflow is shown in **Figure 1**.

Genotype and Memory Data

Genotype data of all non-Hispanic Caucasian participants from both ADNI-1 and ADNI-GO/2 were downloaded, quality controlled, imputed to the Illumina 610 Quad platform and combined. As mentioned earlier, population stratification was performed using the approach described in Kim et al. (2013) to make sure that all the participants were non-Hispanic Caucasian. Among all 1957 selected candidate genes, 25,134 SNPs from 1583 genes (boundary: ± 10 kb) were found based on ANNOVAR (<http://www.openbioinformatics.org/annovar/>). Fifty six SNPs from four genes with missing genotype data were excluded. Totally our analysis included 1579 genes and 25,078 SNPs. Seven episodic memory related cognitive scores (**Table 2**) were downloaded as memory QTs, including (1) immediate recall and delayed recall scores from Logical Memory Test, and (2) trials I–V, List B, immediate recall, 30-min delayed recall, and recognition from Rey Auditory Verbal Learning Test (RAVLT). Final analysis was performed using INTERSNP (Herold et al., 2009) to examine both main and interaction effects. Pairwise interactions were evaluated among all 25,078 SNPs.

Statistical Analysis

This work focuses on the analysis of the genes that are highly preferred to express in the hippocampal region. One of the goals is to investigate both main and epistasis effects of these genes/SNPs on the memory performance measured by seven correlated cognitive scores. Traditional linear regression was first performed to test the main effect of each SNP. A full genetic interaction model was then applied on each pair of SNPs. Potential confounding factors, including baseline age, gender, education, and handedness, were incorporated as covariates in both analyses to exclude their effects.

The whole procedure was performed using the INTERSNP software (Herold et al., 2009). It was designed specifically for genome wide interaction analysis of relating SNPs to the case control conditions or various QTs. In this study, with memory scores as QTs, two linear regression models were used to capture the influence of each interaction pair. The first model took into account the main effects of both SNPs plus all the covariates

TABLE 1 | Demographics of participants.

	HC	eMCI	LMCI	AD
Number	325	191	427	206
Gender(M/F)	173/152	106/85	274/153	115/91
Handedness(R/L)	302/23	170/21	385/42	189/17
Age(mean \pm std)	75.36 \pm 5.32	70.80 \pm 7.37	74.49 \pm 7.47	75.50 \pm 7.98
Education(mean \pm std)	16.27 \pm 2.67	15.91 \pm 2.66	15.86 \pm 2.94	14.95 \pm 3.07
Log_Delay (mean \pm std)	13.35 \pm 3.34	8.88 \pm 1.71	3.87 \pm 2.69	1.34 \pm 1.95
Log_IMM (mean \pm std)	14.14 \pm 3.28	10.94 \pm 2.83	7.14 \pm 3.14	4.10 \pm 2.86
RAV_Recog (mean \pm std)	12.89 \pm 2.40	12.01 \pm 2.70	9.63 \pm 3.65	7.19 \pm 3.94
RAV_T6 (mean \pm std)	8.22 \pm 3.53	7.07 \pm 3.76	3.87 \pm 3.22	1.70 \pm 1.84
RAV_T30 (mean \pm std)	7.29 \pm 3.78	5.84 \pm 4.09	2.83 \pm 3.31	0.73 \pm 1.56
RAV_TB (mean \pm std)	5.01 \pm 1.87	4.51 \pm 1.96	3.65 \pm 1.54	2.89 \pm 1.31
RAV_TOTAL (mean \pm std)	43.69 \pm s9.94	39.41 \pm 10.70	31.04 \pm 9.32	22.63 \pm 8.16

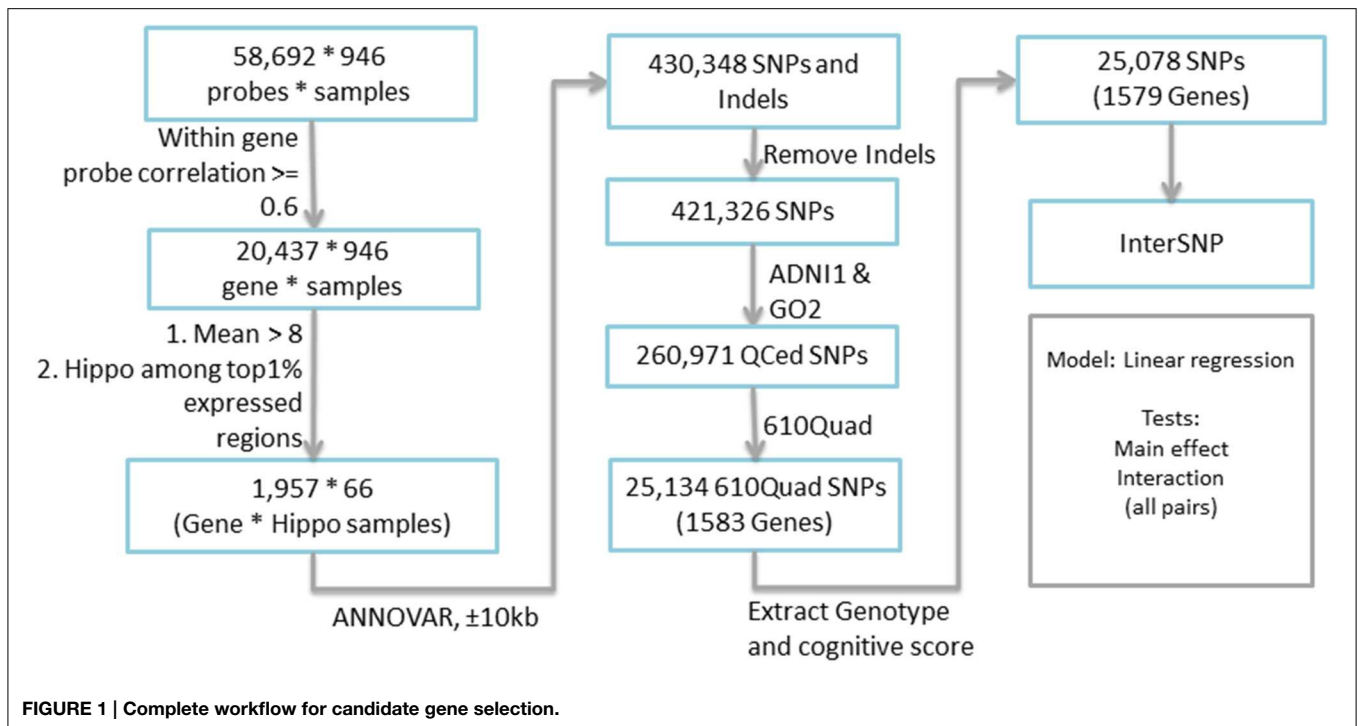


FIGURE 1 | Complete workflow for candidate gene selection.

TABLE 2 | Description of seven episodic memory scores.

Cognitive Score	Description
Log_IMM	Screening logical memory immediate recall
Log_Delay	Screening logical memory delayed recall
RAV_TOT	Baseline RAVLT total score
RAV_TOT6	Baseline RAVLT—trial 6 total number of words recalled
RAV_TOTB	Baseline RAVLT—List B total number of words recalled
RAV_T30	Baseline RAVLT—30 min delay total
RAV_Recog	Baseline RAVLT—30 min delay recognition score

mentioned above, while the second model had an extra interaction term. Comparison between these two models yielded the final interaction effect. Detailed information can be found in <http://intersnp.meb.uni-bonn.de/manual.html>.

Correction for Multiple Testing

Due to the existence of the linkage disequilibrium (LD) structures within SNPs and high correlation among episodic memory measures (Table 3), direct application of either Bonferroni or Benjamin Hochberg correction would be overly conservative and may screen out a lot of potential signals with high false negative rates. In this study, we first estimated the LD blocks in PLINK using an independent data set from the 1000 Genomes Project, and employed the estimated LD block number ($N_{LD} = 5004$) as the independent genetic test number. Similarly for seven cognitive test scores, we estimated the number of independent QTs based on eigenvalues ($N_{iQT} = 2$) by applying the matrix spectral decomposition, using the method described in Van Der Sluis

TABLE 3 | Correlation structure among seven memory scores.

Log_Delay	1.00	0.87	0.55	0.65	0.63	0.42	0.64
Log_IMM	0.87	1.00	0.51	0.62	0.57	0.43	0.66
RAV_Recog	0.55	0.51	1.00	0.64	0.65	0.38	0.63
RAV_T6	0.65	0.62	0.64	1.00	0.86	0.45	0.82
RAV_T30	0.63	0.57	0.65	0.86	1.00	0.50	0.81
RAV_TB	0.42	0.43	0.38	0.45	0.50	1.00	0.61
RAV_TOT	0.64	0.66	0.63	0.82	0.81	0.61	1.00

et al. (2013), Pedraza et al. (2014). In the analysis of main effects, using SNP-level p -values as the input, we performed a gene-based integration based on VEGAS (Liu et al., 2010), where gene level significance was obtained through one million permutations and further corrected for the gene number ($N_G = 1579$) and the estimated independent QT number ($N_{iQT} = 2$) using the Bonferroni procedure (i.e., corrected p -value = uncorrected p -value $\times (N_G \times N_{iQT})$). In the analysis of interaction effects, multiple comparison correction was also performed using the Bonferroni procedure based on $N_{LD}(N_{LD} - 1)/2 = 5004 \times 5003/2 = 12,517,506$ unique LD block pairs and the estimated independent QT number $N_{iQT} = 2$ [i.e., corrected p -value = uncorrected p -value $\times (12,517,506 \times 2)$].

Results

Main Effects

Linear regression was first used to regress 25,078 SNPs on seven cognitive scores, respectively. SNP level p -values obtained from INTERSNP (see top SNPs in Supplemental Table S1) was further

analyzed through VEGAS to obtain gene level p -values. Reference genome in VEGAS was changed to hg19 to be consistent with our data set. Shown in **Figure 2** are the Q-Q plots for the genetic main effect analyses of seven cognitive scores. In total, eight genes from six different chromosomes have been identified with corrected p -value ≤ 0.05 [i.e., uncorrected gene-level p -values $\leq 0.05/(N_G \times N_{iQT}) = 1.6E-05$]. Shown in **Table 4** are the gene findings along with p -values for seven cognitive scores. As observed, the well-known AD risk gene *APOE* is associated with six cognitive scores. Two genes are consistently associated with all seven memory scores: neuron navigator 2 (*NAV2*) and Translocase Of Outer Mitochondrial Membrane 40 Homolog (*TOMM40*), where *TOMM40* is proximal to *APOE* and has been reported to be associated with the memory impairment (Berbee et al., 2011). Four other genes, Protein Kinase, AMP-Activated, Gamma 2 Non-Catalytic Subunit (*PRKAG2*), protein tyrosine phosphatase, receptor type, D (*PTPRD*), CUGBP, elav-like family member 2 (*CELFB2*), and PDS5, regulator of cohesion maintenance, homolog B (*PDS5B*) are each significantly associated with two cognitive scores. Protocadherin 9 (*PCDH9*) is significantly associated with RAVLT list B scores.

Interaction Effects

Similarly, interaction analysis was also performed based on the linear regression model in INTERSNP, where the association between each pair of 25,078 candidate SNPs and each of seven cognitive scores was examined. After Bonferroni correction using the estimated independent test number, no significant interaction

was observed for Log_Delay, RAV_TOT6, and RAV_T30. Two interactions between gene *PTPRD* (rs598356 and rs610789, in the same LD block) and gene *KHSRP* (rs2075755) were found to pass the significance threshold of corrected $p = 0.05$ (or uncorrected $p = 2.0 \times 10^{-9}$) for Log_IMM. Three interactions, between *FLJ39653* and *SOX5*, *FBXO45* and *SOX5*, *FHIT*, and *PRB1*, were found for RAV_Recog. Only one interaction, between *RCC2* and *ZDHHC21*, passed the significance threshold in RAV_TOT test. A large number of interactions were observed to affect RAV_TOTB, which was the list B recall score in RAVLT test. In total, 71 interactions were identified among 46 genes with corrected $p = 0.05$. Shown in **Table 5** is the list of interactions with uncorrected $p \leq 1 \times 10^{-8}$, where bolded records indicate significant interactions with corrected $p \leq 0.05$. The list of interaction findings for RAV_TOTB is shown in Supplemental Table S2, where 298 interactions have uncorrected $p = 1 \times 10^{-8}$, and 71 have corrected $p = 0.05$.

Discussion

Among all eight genes identified in our main effect analysis, *APOE* is significantly associated with six cognitive scores, while two other genes, *NAV2*, and *TOMM40* are observed to be significantly associated with all seven cognitive scores. Two of them, *APOE* and *TOMM40* have been widely studied and known as AD risk genes. For the new candidate gene *NAV2*, despite no direct association has been previously reported between *NAV2* and AD or episodic memory, its special role in neurite growth and cell migration (Muley et al., 2008; McNeill et al., 2010; Shioya

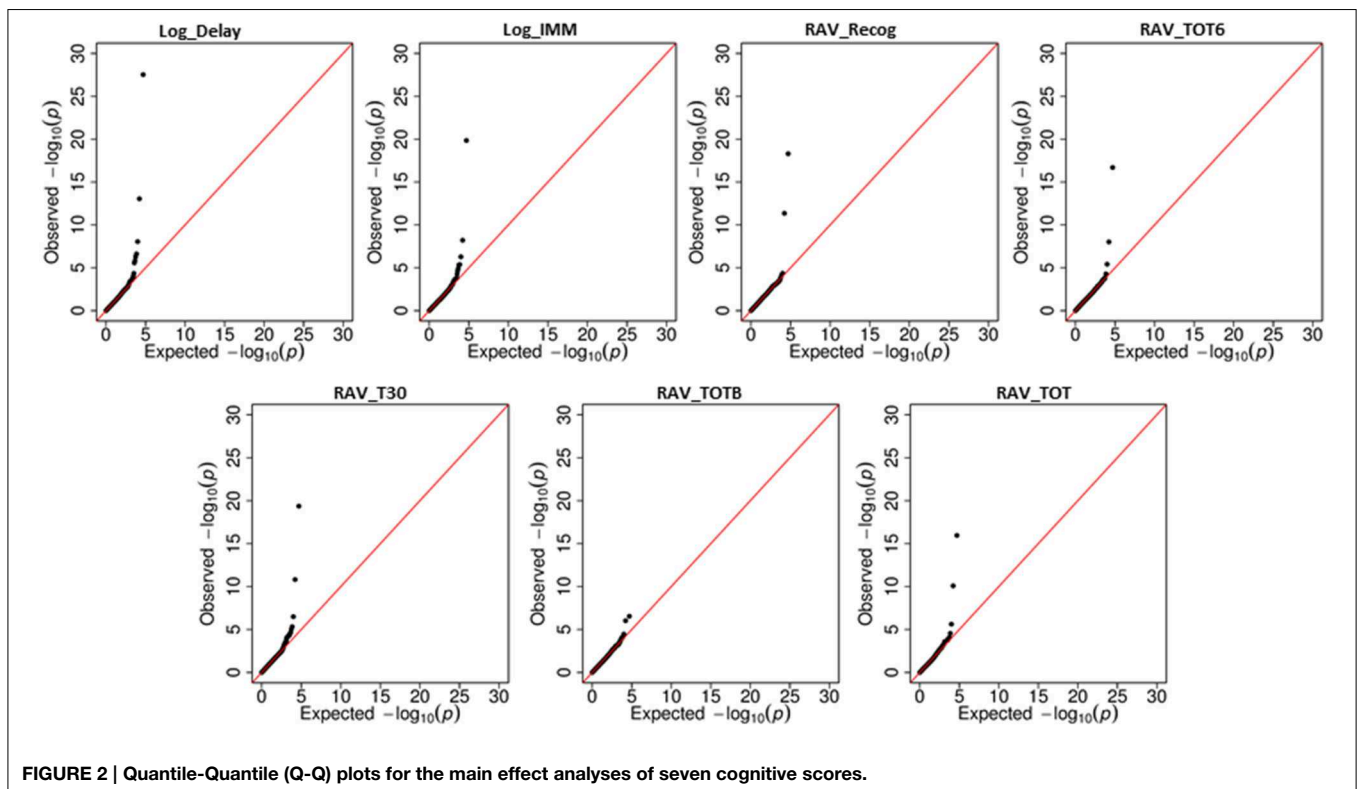


TABLE 4 | Gene-level main effects identified in seven memory scores.

Chr	Gene	7 PRKAG2	9 PTPRD	10 CELF2	11 NAV2	13 PCDH9	13 PDS5B	19 APOE	19 TOMM40
Log_Delay	orig_p	2.5E-03	2.5E-05	4.7E-04	<1.0E-06	5.0E-04	3.0E-06	<1.0E-06	<1.0E-06
	corr_p	NS	NS	NS	<3.2E-03	NS	9.5E-03	<3.2E-03	<3.2E-03
Log_IMM	orig_p	3.2E-04	1.0E-04	2.8E-05	<1.0E-06	4.5E-03	1.3E-05	<1.0E-06	<1.0E-06
	corr_p	NS	NS	NS	<3.2E-03	NS	4.1E-02	<3.2E-03	<3.2E-03
RAV_Recog	orig_p	9.3E-03	2.1E-03	3.2E-05	1.0E-06	1.7E-03	5.8E-02	<1.0E-06	<1.0E-06
	corr_p	NS	NS	NS	3.2E-03	NS	NS	<3.2E-03	<3.2E-03
RAV_T6	orig_p	4.0E-06	8.0E-06	1.1E-04	<1.0E-06	2.0E-02	4.2E-04	<1.0E-06	<1.0E-06
	corr_p	1.3E-02	2.5E-02	NS	<3.2E-03	NS	NS	<3.2E-03	<3.2E-03
RAV_T30	orig_p	4.8E-04	3.0E-06	1.5E-05	<1.0E-06	1.3E-03	2.4E-03	<1.0E-06	<1.0E-06
	corr_p	NS	9.5E-03	4.7E-02	<3.2E-03	NS	NS	<3.2E-03	<3.2E-03
RAV_TB	orig_p	5.4E-02	5.1E-05	1.8E-03	2.0E-06	<1.0E-06	4.3E-04	1.2E-04	<1.0E-06
	corr_p	NS	NS	NS	6.3E-03	<3.2E-03	NS	NS	<3.2E-03
RAV_TOT	orig_p	1.4E-05	2.0E-04	<1.0E-06	<1.0E-06	1.5E-03	3.9E-04	<1.0E-06	<1.0E-06
	corr_p	4.4E-02	NS	<3.2E-03	<3.2E-03	NS	NS	<3.2E-03	<3.2E-03

Both original and corrected p-values are shown. NS indicates “not significant” (i.e., corrected $p > 0.05$). Bolded records indicate significant associations with corrected $p \leq 0.05$.

TABLE 5 | Interaction SNP pairs identified in Log_Delay, Log_IMM, RAV_Recog, RAV_TOT6, and RAV_TOT with uncorrected $p \leq 1 \times 10^{-8}$.

Test	Chr_1	rs_No_1	Gene_1	Dist_1	Chr_2	rs_No_2	Gene_2	Dist_2	p-value	corr_p
Log_Delay	3	rs10866046	<i>FHIT</i>	0	5	rs255211	<i>MCTP1</i>	0	3.58E-09	0.090
	9	rs598356	<i>PTPRD</i>	0	19	rs2075755	<i>KHSRP</i>	0	6.19E-09	0.155
	9	rs610789	<i>PTPRD</i>	0	19	rs2075755	<i>KHSRP</i>	0	7.35E-09	0.184
Log_IMM	9	rs598356	<i>PTPRD</i>	0	19	rs2075755	<i>KHSRP</i>	0	8.74E-10	0.022
	9	rs610789	<i>PTPRD</i>	0	19	rs2075755	<i>KHSRP</i>	0	1.16E-09	0.029
	4	rs1980187	<i>ENOPH1</i>	0	4	rs17579878	<i>ANK2</i>	0	3.24E-09	0.081
	4	rs6827820	<i>ENOPH1</i>	6647	4	rs17579878	<i>ANK2</i>	0	3.54E-09	0.089
	3	rs213376	<i>FHIT</i>	0	6	rs6909677	<i>SASH1</i>	0	4.68E-09	0.117
RAV_Recog	4	rs6822469	<i>FLJ39653</i>	0	12	rs1464502	<i>SOX5</i>	0	4.77E-10	0.012
	3	rs9843585	<i>FBXO45</i>	2925	12	rs16926727	<i>SOX5</i>	0	8.76E-10	0.022
	3	rs17063416	<i>FHIT</i>	0	12	rs2059764	<i>PRB1</i>	-1550	1.44E-09	0.036
	3	rs3915504	<i>FHIT</i>	0	14	rs17510215	<i>NRXN3</i>	0	4.1E-09	0.103
	12	rs1513126	<i>RASSF8</i>	0	17	rs8066154	<i>NUFIP2</i>	0	4.38E-09	0.110
	2	rs4849056	<i>ZC3H8</i>	-6928	9	rs10511506	<i>PTPRD</i>	0	7.99E-09	0.200
	7	rs2190107	<i>GPR37</i>	0	14	rs2268952	<i>AKAP6</i>	0	8.21E-09	0.206
RAV_TOT6	11	rs7118965	<i>OPCML</i>	0	16	rs3866638	<i>WWOX</i>	0	9.71E-09	0.243
RAV_TOT	1	rs2883272	<i>RCC2</i>	8100	9	rs7853156	<i>ZDHHC21</i>	0	1.76E-09	0.044
	6	rs1743448	<i>KHDRBS2</i>	0	18	rs10401068	<i>DOK6</i>	0	3.45E-09	0.086
	1	rs2883272	<i>RCC2</i>	8100	9	rs10961636	<i>ZDHHC21</i>	0	3.74E-09	0.094

No interaction was observed at this threshold for RAV_T30. List of interaction findings for RAV_TB is available in Supplemental Table S2. Bolded records indicate significant interactions with corrected $p \leq 0.05$.

et al., 2010; Marzinke et al., 2013) suggests that it warrants further investigation as a potential target in future analyses. In addition, the expression of NAV3, a paralog of NAV2, was reported to be enhanced in degenerating pyramidal neurons in the cerebral cortex of AD (Shioya et al., 2010).

Each of the other five genes (*PRKAG2*, *PTPRD*, *CELF2*, *PDS5B*, and *PCDH9*) is associated with one or two cognitive

scores in our study. Most of them have already been previously reported as being associated with cognitive impairment. The *PRKAG2* gene provides instructions for making one part (the gamma-2 subunit) of a larger enzyme called AMP-activated protein kinase (*AMPK*), which is a master switch of energy and plays an important role in metabolism functions. Better performance in verbal memories and attention tasks has been found

for specific genotype groups –26 polymorphism *PRKAG2*, indicating an active role of *PRKAG2* in cognitive impairment (Kim et al., 2012). This finding has been confirmed recently in another network based analysis (Caberlotto et al., 2013). Mostly known as tumor suppressor gene (Stallings et al., 2006; Kohno et al., 2010), *PTPRD* was also discovered to have possible functional connections with neurological disorders (Ghani et al., 2012) and may have potential interaction with AD marker tau protein (Shulman et al., 2014). *CELF2* functions to induce the exon 2/3 skipping in *MAPT* gene, which encodes the AD risk protein tau (Ladd, 2013). *PDS5B*, regulator of cohesion maintenance, has been found to be significantly associated with brain atrophy in Furney et al. (2011).

Unlike the main effect tests, no interactions were observed to be associated with all cognitive scores. After Bonferroni correction based on estimated independent genetic and memory QT numbers, no interaction was found for Log_Delay, RAV_TOT6, and RAV_T30. A few interaction signals, ranging from 1 to 3, were found to be associated with Log_IMM, RAV_Recog, or RAV_TOT. Interestingly, a large number of interactions were found in RAV_TOTB. Compared with interactions from the STRING database (Franceschini et al., 2013), no overlap has

been found. We mapped SNPs to genes, and plotted the interaction network using Cytoscape (Shannon et al., 2003); see **Figure 3**. Enrichment analysis was performed using MetaCore from Thomson Reuters and GSEA (Mootha et al., 2003; Subramanian et al., 2005), respectively. In MetaCore, we did not find any enriched pathways with FDR q -value < 0.05 . Gene set enrichment analysis in GSEA yielded 77 enriched sets with FDR q -value < 0.05 , which can generally be categorized into 3 groups: (1) genes with promoter regions containing a specific motif, (2) targets of microRNAs, and (3) genes in cancer module.

In sum, we have performed single SNP/gene analysis and SNP interaction analysis on genes highly expressed in the hippocampal region to identify the genetic factors that are jointly associated with multiple correlated episodic memory scores, in order to reduce the biases introduced by the noise and outliers in the individual analysis of each single score. Three genes were identified to be significantly associated with most of the cognitive scores. *NAV2* is a novel candidate risk gene whereas the other two (*APOE* and *TOMM40*) have been previously reported in AD studies. Although it is not well studied in the AD field, the essential role of *NAV2* in neurite outgrowth and cell migration

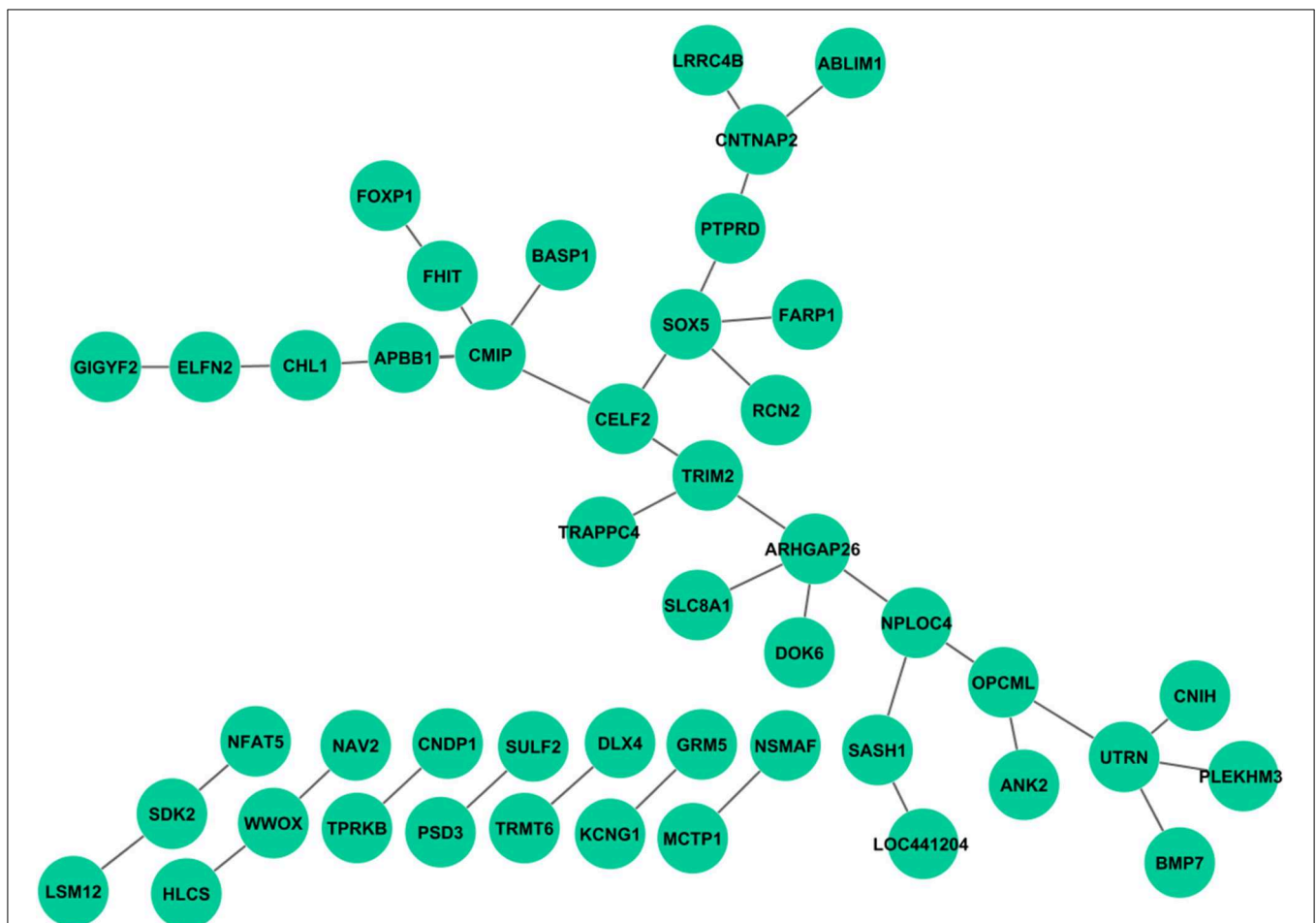


FIGURE 3 | Genetic interactions associated with RAV_TOTB. Each link indicates at least one pair of SNPs between two genes reaches statistical significance with corrected $p < 0.05$.

makes it a potential target warranting further investigation. Unlike the main effect analysis, we did not find any interaction signals that consistently influence all or most of the seven memory scores. Most memory scores have none or few interaction signals observed except RAV_TOTB, which has over 70 interaction pairs passing the significance threshold. Replication and validation of newly identified interactions warrant further investigation.

Acknowledgments

This work was supported by NIH R01 LM011360, U01 AG024904, RC2 AG036535, R01 AG19771, P30 AG10133, UL1TR001108, DOD W81XWH-14-2-0151, NCAA 14132004, and NSF IIS-1117335 at IU, and by NIH R01 LM011360, R01 LM009012, and R01 LM010098 at Dartmouth. 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: Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech;

BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; 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; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. 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). The grantee organization is the Northern Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Supplementary Material

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fgene.2015.00117/abstract>

References

- Backman, L., Small, B. J., and Fratiglioni, L. (2001). Stability of the preclinical episodic memory deficit in Alzheimer's disease. *Brain* 124, 96–102. doi: 10.1093/brain/124.1.96
- Bennett, D. A., De Jager, P. L., Leurgans, S. E., and Schneider, J. A. (2009). Neuropathologic intermediate phenotypes enhance association to Alzheimer susceptibility alleles. *Neurology* 72, 1495–1503. doi: 10.1212/WNL.0b013e3181a2e87d
- Berbee, J. F., Vanmierlo, T., Abildayeva, K., Blokland, A., Jansen, P. J., Lutjohann, D., et al. (2011). Apolipoprotein CI knock-out mice display impaired memory functions. *J. Alzheimers. Dis.* 23, 737–747. doi: 10.3233/JAD-2010-100576
- Burgess, J. D., Pedraza, O., Graff-Radford, N. R., Hirpa, M., Zou, F. G., Miles, R., et al. (2011). Association of common KIBRA variants with episodic memory and AD risk. *Neurobiol. Aging* 32, 557.e1–557.e9. doi: 10.1016/j.neurobiolaging.2010.11.004
- Caberlotto, L., Lauria, M., Nguyen, T. P., and Scotti, M. (2013). The central role of AMP-Kinase and energy homeostasis impairment in Alzheimer's disease: a multifactor network analysis. *PLoS ONE* 8:e78919. doi: 10.1371/journal.pone.0078919
- Egan, M. F., Kojima, M., Callicott, J. H., Goldberg, T. E., Kolachana, B. S., Bertolino, A., et al. (2003). The BDNF val66met polymorphism affects activity-dependent secretion of BDNF and human memory and hippocampal function. *Cell* 112, 257–269. doi: 10.1016/S0092-8674(03)00035-7
- Ertekin-Taner, N. (2010). Genetics of Alzheimer disease in the pre- and post-GWAS era. *Alzheimers Res. Ther.* 2, 3. doi: 10.1186/alzrt26
- Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A., et al. (2013). STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* 41, D808–D815. doi: 10.1093/nar/gks1094
- Furney, S. J., Simmons, A., Breen, G., Pedrosa, I., Lunnon, K., Proitsi, P., et al. (2011). Genome-wide association with MRI atrophy measures as a quantitative trait locus for Alzheimer's disease. *Mol. Psychiatry* 16, 1130–1138. doi: 10.1038/mp.2010.123
- Gatz, M., Reynolds, C. A., Fratiglioni, L., Johansson, B., Mortimer, J. A., Berg, S., et al. (2006). Role of genes and environments for explaining Alzheimer disease. *Arch. Gen. Psychiatry* 63, 168–174. doi: 10.1001/archpsyc.63.2.168
- Ghani, M., Pinto, D., Lee, J. H., Grinberg, Y., Sato, C., Moreno, D., et al. (2012). Genome-wide survey of large rare copy number variants in Alzheimer's disease among caribbean hispanics. *G3* 2, 71–78. doi: 10.1534/g3.111.000869
- Glahn, D. C., Almasy, L., Blangero, J., Burk, G. M., Estrada, J., Peralta, J. M., et al. (2007). Adjudicating neurocognitive endophenotypes for schizophrenia. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 144B, 242–249. doi: 10.1002/ajmg.b.30446
- Gottesman, I. I., and Gould, T. D. (2003). The endophenotype concept in psychiatry: etymology and strategic intentions. *Am. J. Psychiatry* 160, 636–645. doi: 10.1176/appi.ajp.160.4.636
- Harold, D., Abraham, R., Hollingworth, P., Sims, R., Gerrish, A., Hamshere, M. L., et al. (2009). Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat. Genet.* 41, 1088–1093. doi: 10.1038/ng.440
- Herold, C., Steffens, M., Brockschmidt, F. F., Baur, M. P., and Becker, T. (2009). INTERSNP: genome-wide interaction analysis guided by a priori information. *Bioinformatics* 25, 3275–3281. doi: 10.1093/bioinformatics/btp596
- Hollingworth, P., Harold, D., Sims, R., Gerrish, A., Lambert, J. C., Carrasquillo, M. M., et al. (2011). Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. *Nat. Genet.* 43, 429–435. doi: 10.1038/ng.803
- Hu, X. L., Pickering, E., Liu, Y. C., Hall, S., Fournier, H., Katz, E., et al. (2011). Meta-analysis for genome-wide association study identifies multiple variants at the BIN1 locus associated with late-onset Alzheimer's disease. *PLoS ONE* 6:e16616. doi: 10.1371/journal.pone.0016616
- Jack, C. R. Jr., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., et al. (2008). The Alzheimer's disease Neuroimaging Initiative (ADNI): MRI methods. *J. Magn. Reson. Imaging* 27, 685–691. doi: 10.1002/jmri.21049
- Jagust, W. J., Bandy, D., Chen, K., Foster, N. L., Landau, S. M., Mathis, C. A., et al. (2010). The Alzheimer's disease Neuroimaging Initiative positron emission tomography core. *Alzheimers Dement.* 6, 221–229. doi: 10.1016/j.jalz.2010.03.003

- Kim, E., Lee, S. H., Lee, K. S., Cheong, H. K., Namkoong, K., Hong, C. H., et al. (2012). AMPK gamma 2 subunit gene PRKAG2 polymorphism associated with cognitive impairment as well as diabetes in old age. *Psychoneuroendocrinology* 37, 358–365. doi: 10.1016/j.psyneuen.2011.07.005
- Kim, S., Swaminathan, S., Inlow, M., Risacher, S. L., Nho, K., Shen, L., et al. (2013). Influence of genetic variation on plasma protein levels in older adults using a multi-analyte panel. *PLoS ONE* 8:e70269. doi: 10.1371/journal.pone.0070269
- Kohno, T., Otsuka, A., Girard, L., Sato, M., Iwakawa, R., Ogiwara, H., et al. (2010). A catalog of genes homozygously deleted in human lung cancer and the candidacy of PTPRD as a tumor suppressor gene. *Genes Chromosomes Cancer* 49, 342–352. doi: 10.1002/gcc.20746
- Ladd, A. N. (2013). CUG-BP, Elav-like family (CELF)-mediated alternative splicing regulation in the brain during health and disease. *Mol. Cell. Neurosci.* 56, 456–464. doi: 10.1016/j.mcn.2012.12.003
- Lambert, J. C., Heath, S., Even, G., Campion, D., Sleegers, K., Hiltunen, M., et al. (2009). Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. *Nat. Genet.* 41, U1094–U1068. doi: 10.1038/ng.439
- Lambert, J. C., Ibrahim-Verbaas, C. A., Harold, D., Naj, A. C., Sims, R., Bellenguez, C., et al. (2013). Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* 45, 1452–1458. doi: 10.1038/ng.2802
- Lee, J. H., Cheng, R., Barral, S., Reitz, C., Medrano, M., Lantigua, R., et al. (2011). Identification of novel loci for Alzheimer disease and replication of CLU, PICALM, and BIN1 in Caribbean Hispanic individuals. *Arch. Neurol.* 68, 320–328. doi: 10.1001/archneurol.2010.292
- Liu, J. Z., Mcrae, A. F., Nyholt, D. R., Medland, S. E., Wray, N. R., Brown, K. M., et al. (2010). A versatile gene-based test for genome-wide association studies. *Am. J. Hum. Genet.* 87, 139–145. doi: 10.1016/j.ajhg.2010.06.009
- Marzinke, M. A., Mavencamp, T., Duratinsky, J., and Clagett-Dame, M. (2013). 14-3-3epsilon and NAV2 interact to regulate neurite outgrowth and axon elongation. *Arch. Biochem. Biophys.* 540, 94–100. doi: 10.1016/j.abb.2013.10.012
- McNeill, E. M., Roos, K. P., Moechars, D., and Clagett-Dame, M. (2010). Nav2 is necessary for cranial nerve development and blood pressure regulation. *Neural Dev.* 5:6. doi: 10.1186/1749-8104-5-6
- Mootha, V. K., Lindgren, C. M., Eriksson, K. F., Subramanian, A., Sihag, S., Lehar, J., et al. (2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* 34, 267–273. doi: 10.1038/ng1180
- Muley, P. D., McNeill, E. M., Marzinke, M. A., Knobel, K. M., Barr, M. M., and Clagett-Dame, M. (2008). The aTRA-responsive gene neuron navigator 2 functions in neurite outgrowth and axonal elongation. *Dev. Neurobiol.* 68, 1441–1453. doi: 10.1002/dneu.20670
- Naj, A. C., Jun, G., Beecham, G. W., Wang, L. S., Vardarajan, B. N., Buross, J., et al. (2011). Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. *Nat. Genet.* 43, 436–441. doi: 10.1038/ng.801
- Pedraza, O., Allen, M., Jennette, K., Carrasquillo, M., Crook, J., Serie, D., et al. (2014). Evaluation of memory endophenotypes for association with CLU, CR1, and PICALM variants in black and white subjects. *Alzheimers Dement.* 10, 205–213. doi: 10.1016/j.jalz.2013.01.016
- Petersen, R. C., Aisen, P. S., Beckett, L. A., Donohue, M. C., Gamst, A. C., Harvey, D. J., et al. (2010). Alzheimer's disease Neuroimaging Initiative (ADNI): clinical characterization. *Neurology* 74, 201–209. doi: 10.1212/WNL.0b013e3181cb3e25
- Ramanan, V. K., Nho, K., Shen, L., Risacher, S. L., Kim, S., McDonald, B. C., et al. (2014). FASTKD2 is associated with memory and hippocampal structure in older adults. *Mol. Psychiatry*. doi: 10.1038/mp.2014.142. [Epub ahead of print].
- Saunders, A. M., Strittmatter, W. J., Schmechel, D., Georgehyslop, P. H. S., Pericakvance, M. A., Joo, S. H., et al. (1993). Association of Apolipoprotein-E Allele Epsilon-4 with Late-Onset Familial and Sporadic Alzheimers-disease. *Neurology* 43, 1467–1472. doi: 10.1212/WNL.43.8.1467
- Saykin, A. J., Shen, L., Foroud, T. M., Potkin, S. G., Swaminathan, S., Kim, S., et al. (2010). Alzheimer's disease Neuroimaging Initiative biomarkers as quantitative phenotypes: genetics core aims, progress, and plans. *Alzheimers Dement.* 6, 265–273. doi: 10.1016/j.jalz.2010.03.013
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303
- Shaw, L. M., Vanderstichele, H., Knapiak-Czajka, M., Clark, C. M., Aisen, P. S., Petersen, R. C., et al. (2009). Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects. *Ann. Neurol.* 65, 403–413. doi: 10.1002/ana.21610
- Shen, L., Kim, S., Risacher, S. L., Nho, K., Swaminathan, S., West, J. D., et al. (2010). Whole genome association study of brain-wide imaging phenotypes for identifying quantitative trait loci in MCI and AD: a study of the ADNI cohort. *Neuroimage* 53, 1051–1063. doi: 10.1016/j.neuroimage.2010.01.042
- Shi, H., Belbin, O., Medway, C., Brown, K., Kalsheker, N., Carrasquillo, M., et al. (2012). Genetic variants influencing human aging from late-onset Alzheimer's disease (LOAD) genome-wide association studies (GWAS). *Neurobiol. Aging* 33, 1849.e5–1849.e18. doi: 10.1016/j.neurobiolaging.2012.02.014
- Shioya, M., Obayashi, S., Tabunoki, H., Arima, K., Saito, Y., Ishida, T., et al. (2010). Aberrant microRNA expression in the brains of neurodegenerative diseases: miR-29a decreased in Alzheimer disease brains targets neurone navigator 3. *Neuropathol. Appl. Neurobiol.* 36, 320–330. doi: 10.1111/j.1365-2990.2010.01076.x
- Shulman, J. M., Imboyya, S., Giagtzoglou, N., Powers, M. P., Hu, Y. H., Devenport, D., et al. (2014). Functional screening in Drosophila identifies Alzheimer's disease susceptibility genes and implicates Tau-mediated mechanisms. *Hum. Mol. Genet.* 23, 870–877. doi: 10.1093/hmg/ddt478
- Stallings, R. L., Nair, P., Maris, J. M., Catchpole, D., Mcdermott, M., O'meara, A., et al. (2006). High-resolution analysis of chromosomal breakpoints and genomic instability identifies PTPRD as a candidate tumor suppressor gene in neuroblastoma. *Cancer Res.* 66, 3673–3680. doi: 10.1158/0008-5472.CAN-05-4154
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* 102, 15545–15550. doi: 10.1073/pnas.0506580102
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289. doi: 10.1006/nimg.2001.0978
- Van Der Sluis, S., Posthuma, D., and Dolan, C. V. (2013). TATES: efficient multivariate genotype-phenotype analysis for genome-wide association studies. *PLoS Genet.* 9:e1003235. doi: 10.1371/journal.pgen.1003235
- Zeng, H., Shen, E. H., Hohmann, J. G., Oh, S. W., Bernard, A., Royall, J. J., et al. (2012). Large-scale cellular-resolution gene profiling in human neocortex reveals species-specific molecular signatures. *Cell* 149, 483–496. doi: 10.1016/j.cell.2012.02.052

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Yan, Kim, Nho, Chen, Risacher, Moore, Saykin and Shen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.