Translating Alzheimer's disease–associated polymorphisms into functional candidates: a survey of IGAP genes and SNPs

Check for updates

Yuriko Katsumata [a,b,*], Peter T. Nelson [b,c], Steven Estus [b,d], the Alzheimer's Disease Neuroimaging Initiative (ADNI)[1], David W. Fardo [a,b]

[a] Department of Biostatistics, University of Kentucky, Lexington, KY, USA
[b] Sanders-Brown Center on Aging, University of Kentucky, Lexington, KY, USA
[c] Department of Pathology, University of Kentucky, Lexington, KY, USA
[d] Department of Physiology, University of Kentucky, KY, USA

ABSTRACT

The International Genomics of Alzheimer's Project (IGAP) is a consortium for characterizing the genetic landscape of Alzheimer's disease (AD). The identified and/or confirmed 19 single-nucleotide polymorphisms (SNPs) associated with AD are located on non-coding DNA regions, and their functional impacts on AD are as yet poorly understood. We evaluated the roles of the IGAP SNPs by integrating data from many resources, based on whether the IGAP SNP was (1) a proxy for a coding SNP or (2) associated with altered mRNA transcript levels. For (1), we confirmed that 12 AD-associated coding common SNPs and five nonsynonymous rare variants are in linkage disequilibrium with the IGAP SNPs. For (2), the IGAP SNPs in *CELF1* and *MS4A6A* were associated with expression of their neighboring genes, *MYBPC3* and *MS4A6A*, respectively, in blood. The IGAP SNP in *DSG2* was an expression quantitative trait loci (eQTL) for *DLGAP1* and *NETO1* in the human frontal cortex. The IGAP SNPs in *ABCA7*, *CD2AP*, and *CD33* each acted as eQTL for AD-associated genes in brain. Our approach for identifying proxies and examining eQTL highlighted potentially impactful, novel gene regulatory phenomena pertinent to the AD phenotype.

© 2018 Elsevier Inc. All rights reserved.

1. Introduction

Dementia is a clinical state characterized by a loss of function in memory and behavior, and the clinical state is associated with underlying degeneration of central nervous system synapses and cells. Alzheimer's disease (AD) is the most common form of dementia, accounting for over 50% of dementia cases (Querfurth and LaFerla, 2010). Although it has been more than 100 years since Alois Alzheimer published "About a Peculiar Disease of the Cerebral Cortex" (Alzheimer, 1907), the exact cause of AD has not yet been defined. Amyloid beta (Aβ) protein and hyperphosphorylated tau aggregates in brain are considered the key pathological hallmarks (Reitz et al., 2011; Selkoe et al., 2004). A widely held mechanistic

hypothesis for AD pathogenesis is the "amyloid cascade hypothesis" wherein a key early pathogenetic role is played by parenchymal Aβ peptide accumulation, which causes or exacerbates downstream neuronal injury, enhanced neuroinflammation, tau hyperphosphorylation, and eventually the clinical symptoms of AD (Hardy and Selkoe, 2002).

Familial AD, which often occurs early in life, is linked to mutations in three genes: the amyloid precursor protein (*APP*) gene and the presenilin protein (*PSEN1* and *PSEN2*) genes (van Es and van den Berg, 2009). These genes are associated with altered processing of the APP protein, including a shift in Aβ peptide production from $A\beta_{40}$ to more neurotoxic $A\beta_{42}$ (e.g., Volga German mutation in *PSEN2* and Iberian mutation in *APP*) (Jayadev et al., 2010; Levy-Lahad et al., 1995; Lichtenthaler et al., 1999; Walker et al., 2005), increased total Aβ levels (Swedish mutation in *APP*) (Mullan et al., 1992), and increased Aβ protofibril formation (Arctic mutation in *APP*) (Nilsberth et al., 2001). In contrast, late-onset AD, which accounts for >95% of all AD cases (Mancuso et al., 2008), has a more complex genetic architecture. The ε4 allele of apolipoprotein E (*APOE*) gene is the most well-established susceptibility risk factor for late-onset AD.

A series of genome-wide association studies (GWAS) have identified AD-associated single-nucleotide polymorphisms (SNPs) in addition to the *APOE* alleles (Harold et al., 2009; Hollingworth

et al., 2011; Lambert et al., 2009, 2013; Naj et al., 2011; Seshadri et al., 2010). The study with the largest number of AD and non-AD individuals was the International Genomics of Alzheimer's Project (IGAP), which capitalized on a large, multicenter study design to include 74,046 individuals (Lambert et al., 2013). This study extended associations between the AD phenotype and genetics, finding 21 SNPs as significant by meta-analyzing genetic and phenotype data from four component consortia (Lambert et al., 2013). These SNPs are in or close to *CR1, BIN1, INPP5D, MEF2C, CD2AP, NME8, EPHA1, PTK2B, PICALM, SORL1, FERMT2, SLC24A4-RIN3, DSG2, CASS4, HLA-DRB5-DBR1, CLU, MS4A6A, ABCA7, CD33, ZCWPW1,* and *CELF1* (Supplementary Table 1). Although GWAS have succeeded in revealing numerous susceptibility variants for AD, determining the functional impact of those gene variants and understanding how they contribute to AD pathogenesis represents a barrier to progress in the field.

Genetic variants located in coding regions constitute only ∼1% of gene polymorphisms seen in humans (Rabbani et al., 2014). However, there are many ways that genetic variants in noncoding regions can affect protein expression and structure and thereby exert a protective or disease-inducing impact. Functional variants may be located in a coding region, an alternative splicing region, or a regulatory region such as promoter, operator, insulator, enhancer, or silencer. Nonsynonymous variants, by definition, alter the primary amino acid sequence of a protein and may have effects on the protein structure and function. Synonymous mutations occur in the coding region but do not change the amino acid sequence. These variants were referred to as "silent mutations" until recently (Sauna and Kimchi-Sarfaty, 2011). Several synonymous mutations have been reported to affect mRNA splicing and stability, gene expression, and protein folding and function (Sauna and Kimchi-Sarfaty, 2011). Most of the non-*APOE* AD-associated genetic variants described to date are located in intronic or intergenic regions (i.e., noncoding regions), which may contain regulatory elements. Intronic and intergenic SNPs may act by regulating expression of disease-associated genes and/or modulating translation efficiency and stability (Mockenhaupt and Makeyev, 2015).

In the present study, we analyzed data from multiple sources to gain insights into the roles of noncoding SNPs identified in the IGAP study [including *CD33* and *DSG2*, although their associated SNPs, rs3865444 and rs8093731, respectively, did not reach statistical significance in the combined stages of that study (Lambert et al., 2013)], hereafter referred to as "IGAP SNPs." We hypothesized that each IGAP SNP is potentially: (1) a proxy for an exonic (coding) variant (Supplementary Fig. 1A and B) that has not yet been identified or (2) associated with altered transcript/mRNA levels (Supplementary Fig. 1C). One approach to test the first hypothesis is to identify coding variants in strong linkage disequilibrium (LD) with the variant identified by GWAS, which indicates that the two gene loci are commonly coinherited. For the second hypothesis, expression quantitative trait loci (eQTL) analyses can be used to assess the association between the gene variant and mRNA levels of various transcripts. Thus, eQTL are genetic loci that contribute to variation in gene expression. By mapping eQTL, we investigated how the variants regulate gene expression. Using these combined methods, and multiple data sources, we discovered new evidence of complex gene expression regulation mechanisms in association with previously identified IGAP SNPs.

2. Material and methods

2.1. Genetic data sets

Genetic data were obtained from multiple sources. Whole-exome sequence (WES) data came from the Alzheimer's Disease Sequencing Project (ADSP), composed of 18 cohorts from the Alzheimer's Disease Genetic Consortium (ADGC) and six cohorts from the Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium (Beecham et al., 2017). We also used imputed SNP data (Supplementary Method) from the ADGC comprising 23 different cohorts (Supplementary Table 2). From these sources, there were 28,730 unrelated subjects with imputed GWAS SNP data in ADGC and 10,913 unrelated subjects with WES data in ADSP. We estimated identity-by-descent (IBD) to identify any relatedness and duplicate individuals in the two data sets. Individuals were excluded with estimated IBD ≥0.1875 from ADGC data sets, and two independent data sets were created: an imputed ADGC data set that excluded related individuals and those that potentially overlapped with those in ADSP (hereafter referred to as "ADGC") and WES data in ADSP (hereafter referred to as "ADSP") (Fig. 1). We limited the included subjects to those who had AD diagnosis information and who were 65 years or older at the last visit or at death, yielding a total of 15,343 ADGC subjects with imputed SNP data in the discovery analysis and a total of 10,407 ADSP subjects with WES data in the replication analysis.

2.2. Gene expression data sets

Data were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging, positron emission tomography, other biological markers, and clinical and neuropsychological assessments can be combined to measure the progression of mild cognitive impairment and early AD. Data were included from 763 subjects aged 65 years or older who had both gene expression data from blood (Affymetrix Human Genome U219 Array platform) and whole-genome sequencing data available. Clinical status was determined based on the clinical evaluation at the last examination.

Human brain gene expression and genotype data were obtained from the North American Brain Expression Consortium (NABEC) (Hernandez et al., 2012) and United Kingdom Brain Expression Consortium (UKBEC) (Trabzuni et al., 2011). Details were as described in our previous report (Katsumata et al., 2017). Briefly, the NABEC expression data for the frontal cortex (FCTX) were available at Gene Expression Omnibus (GEO: https://www.ncbi.nlm.nih.gov/geo/) public repository, and the genotype data were obtained from the database of Genotypes and Phenotypes (dbGaP: http://www.ncbi.nlm.nih.gov/gap). Standard quality control (QC) procedures were performed on the NABEC genotype data using PLINK v1.90a (Purcell et al., 2007). Markers were excluded based on the following criteria: (1) minor allele frequency < 1%; (2) call rate per variant (SNPs and indels) < 95%; and (3) Hardy-Weinberg equilibrium test in controls <10−5. Samples were excluded based on the following criteria: (1) call rate per individual <95%, (2) a high degree of relatedness per an estimated proportion of IBD >0.1875, and (3) excess of ±3.0 standard deviations of heterozygosity rate. After performing QC, we imputed using the Michigan Imputation Server (https://imputationserver.sph.umich.edu/start.html) (Das et al., 2016; Loh et al., 2016) with the following parameters: 1000 Genome Phase 3 v5 reference panel, Eagle v2.3 phasing (Loh et al., 2016), and EUR population. Of the 455 neurologically normal donors, 85 subjects who died at the age of 65 years or older and passed QC were included in the analysis (all were U.S. Caucasians). The UKBEC gene expression for three brain regions (FCTX; hippocampus [HIPP]; and temporal cortex [TCTX]) and genotype data were obtained from the BRAINEAC website (http://www.braineac.org/). Dosage genotype data were converted into PLINK file format using
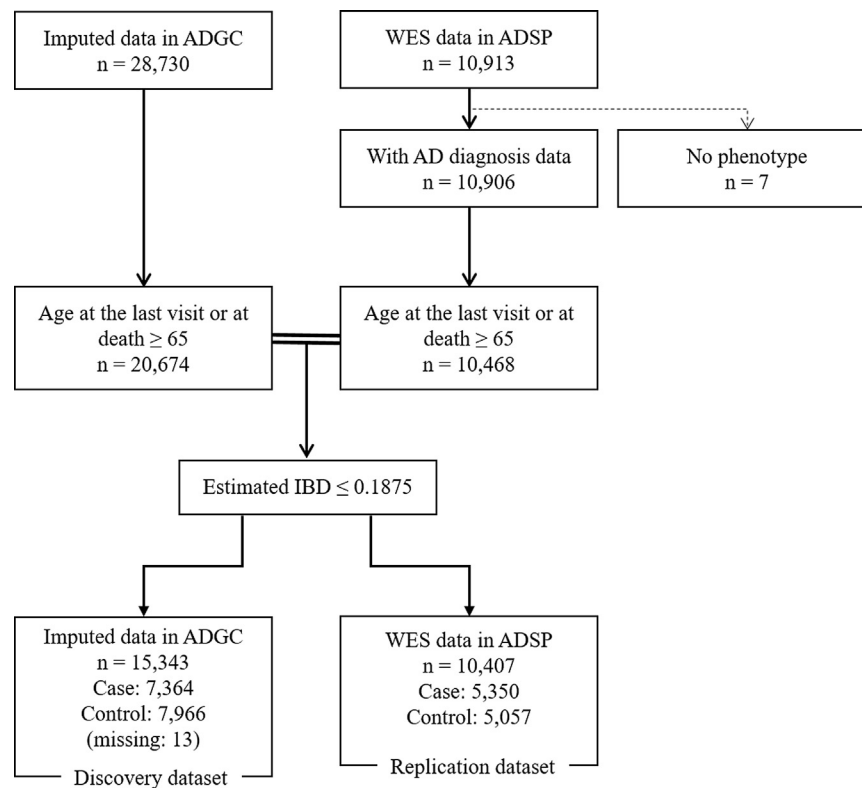
Fig. 1. Flow diagram of the subjects included in the analyses. Abbreviations: ADGC, Alzheimer's Disease Genetics Consortium; ADSP, Alzheimer's Disease Sequencing Project; IBD, identity-by-descent; WES, whole-exome sequence.

Genome-wide Complex Trait Analysis software version 1.24.4 (Yang et al., 2011). Among the 134 neuropathologically normal individuals, 49 subjects who died at the age of 65 years or older were included in the present analyses.

Because the NABEC and UKBEC data sets do not have AD diagnosis information, we retrieved microarray data sets generated from Affymetrix Human Genome U133 Plus 2.0 Array platform (GPL570) regarding AD status from GEO to examine whether the levels of gene expression were different in association with AD status versus controls. We focused on gene expression in four brain regions affected by AD (entorhinal cortex [EC], FCTX, HIPP, and TCTX). We obtained two data sets for EC [GSE48350 (Berchtold et al., 2008) and GSE5281 (Liang et al., 2007)], four data sets for FCTX [GSE48350 (Berchtold et al., 2008), GSE5281 (Liang et al., 2007), GSE66333 (Simpson et al., 2016), and GSE53890 (Lu et al., 2014)], three datasets for HIPP [GSE48350 (Berchtold et al., 2008), GSE5281 (Liang et al., 2007), GSE28146 (Blalock et al., 2011)], and two datasets for TCTX [GSE5281 (Liang et al., 2007) and GSE29652 (Simpson et al., 2011)]. Included were 25 AD cases and 29 controls for EC, 52 cases and 56 controls for FCTX, 50 cases and 44 controls for HIPP, and 34 cases and 11 controls for TCTX, who were 65 years or older at death (Supplementary Table 3). The raw expression data downloaded from GEO (Affymetrix CEL files) were background-corrected and normalized by the RmaBackgroundCorrection and QuantileNormalization functions in "aroma.affymetrix" Bioconductor R package (Bengtsson et al., 2008), and then log2-transformed. The normalized and log2-transformed expression data in each brain region were merged by the Combat function in "sva" Bioconductor R package (Leek et al., 2012). Using principal component analysis, we confirmed that Combat successfully eliminated batch effects in each brain region. We removed one outlier identified in the principal component analysis from GSE5281 in the FCTX (Supplementary Fig. 2).

Probes were excluded that targeted transcripts from different genes (i.e., probes with "_x" suffix) if a more reliable probe was available. We also excluded monoallelically expressed genes including genes on chromosomes X and Y, and HLA genes (i.e., *HLA-A, HLA-B, HLA-C, HLA-DMA, HLA-DMB, HLA-DOA, HLA-DOB, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DQA2, HLA-DQB1, HLA-DQB2, HLA-DRA, HLA-DRB1, HLA-DRB3, HLA-DRB4,* HLA-DRB5, *HLA-E, HLA-F, HLA-G, HLA-J, HLA-P,* and *HLA-T*). The number of probes in each study is shown in Supplementary Table 4.

2.3. Statistical analysis

2.3.1. Hypothesis 1 identified IGAP SNPs are proxies for exonic/coding variants

We applied two separate methods to identify potentially coinherited SNPs with the IGAP SNPs: one was for common SNPs and one for rare variants. We first identified common SNPs in the nearby coding regions showing strong ($r^2 \geq 0.8$) or moderate ($0.4 \leq r^2 < 0.8$) LD with each of the IGAP SNPs by using 1000 Genomes Project Phase 3 in individuals of European ancestry (1000 Genomes EUR) (1000 Genomes Project Consortium, 2010). In the discovery analysis for common SNPs in ADGC, we performed association tests under an additive mode of inheritance using logistic regression adjusted for age at the last visit or death, sex, and the top five principal components computed in PLINK v1.90a (Chang et al., 2015; Purcell et al., 2007). The coding SNPs were evaluated for replication in independent individuals (within the ADSP data set) to limit the possibility of imputation errors. We then identified potentially coinherited rare variants with each of the IGAP SNPs using the Lewontin's D′ estimates (Lewontin, 1964) in 1000 Genomes EUR (1000 Genomes Project Consortium, 2010). Due to the properties of LD metrics, a given common SNP can exhibit disparate patterns of LD (large D′ but low r^2) between it and many rare

variants. We thus focused on nonsynonymous rare variants which are more likely functional and then applied the following criteria: minor allele frequency <0.05, minor allele count ≥5, D' ≥ 0.9, the same direction of effect on AD, and within 1 Mb from the IGAP SNP. Fisher's exact test was used to examine the association between rare variants and AD in ADSP.

Variant Effect Predictor (McLaren et al., 2010) was used to annotate functional consequences of the common coding SNPs and rare variants identified in the association tests. The pathogenetic nature of nonsynonymous common SNPs/rare variants associated with AD was predicted by SIFT (http://sift.jcvi.org/) (Ng and Henikoff, 2003), PolyPhen-2 with HumDiv classifier (http://genetics.bwh.harvard.edu/pph2/) (Adzhubei et al., 2010), and PROVEAN (http://provean.jcvi.org/index.php) (Choi et al., 2012) to evaluate the effect of amino acid substitution on a protein function. We also used Genomic Evolutionary Rate Profiling (GERP)++ (http://mendel.stanford.edu/SidowLab/downloads/gerp/) (Davydov et al., 2010) to examine evolutionary conservation for each of the associated nonsynonymous SNPs/rare variants. Higher score of rejected substitutions score indicates that a site is inferred to have a greater level of evolutionary constraint. We implemented these *in silico* algorithm tools except for PROVEAN for canonical transcripts that are defined as either the longest coding sequence or the longest cDNA in the UCSC Genome Browser (https://genome.ucsc.edu/) (Kent et al., 2002).

2.3.2. Hypothesis 2 identified IGAP SNPs are eQTLs

The goal of these analyses was to evaluate whether the IGAP SNPs were eQTL. We first tested the association between the IGAP SNPs and gene expression on the same chromosome as each of the SNPs, assuming an additive mode of inheritance as implemented in PLINK v1.90a (Chang et al., 2015; Purcell et al., 2007). We then examined whether the levels of gene expression associated with the IGAP SNP status were different from the association with the AD phenotype. An analysis of covariance with age at the death and sex as covariates was applied to test for statistical significance.

For all analyses, we defined associations with false discovery rate (FDR)-adjusted *p*-value <0.05 as statistically significant using the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995).

3. Results

For the present study, individuals with either prevalent or incident AD were considered as AD cases in ADSP. Descriptive characteristics of individuals in the two genetic data sets are shown in Supplementary Table 5. In ADGC and ADSP, 7364 (48.0%) and 5374 (51.4%) were AD cases, respectively.

3.1. Hypothesis 1 identified IGAP SNPs are proxies of coding variants

In the common SNP analyses, 10 exonic SNPs were in strong LD ($r^2 \geq 0.8$) and 16 exonic SNPs in moderate LD ($0.4 \leq r^2 < 0.8$) with IGAP SNPs based on 1000 Genomes EUR (Supplementary Table 6). We first analyzed the imputed genotype data from ADGC to replicate IGAP SNP association with AD and extended the analyses to exonic common SNPs. We confirmed that several proxy SNPs located in coding regions demonstrated statistically significant associations with the AD phenotype. Of these coding SNPs, we replicated 12 loci (rs2296160 in *CR1*, rs1049086 in *HLA-DQB1*, rs2722372 and rs2598044 in *NME8*, rs2405442 and rs1859788 in *PILRA*, rs7982 in *CLU*, rs12453 and rs7232 in *MS4A6A*, and rs3752246, rs4147930, and rs4147934 in *ABCA7*) that surpassed the statistical significance level with FDR adjustment in the separate ADSP data set (Table 1 and Supplementary Table 7). Of these 12

coding SNPs, six SNPs (rs2296160 in *CR1*, rs2722372 in *NME8*, rs1859788 in *PILRA*, rs7232 in *MS4A6A*, and rs3752246 and rs4147934 in *ABCA7*) are missense mutations on at least one of their transcripts (Supplementary Table 8).

In rare variant analyses using ADSP, five rare variants were identified that (1) had D' = 1 and the same direction of effect as an IGAP SNP (rs11575848 in *LY6G6C*, rs2070600 in *AGER*, rs62483572 in *EPO*, rs74547795 in *SYTL2*, and rs111986709 in *DSG3*) (Supplementary Table 9) and (2) were missense mutations on at least one of their transcripts (Supplementary Table 10).

The nonsynonymous SNPs and rare variants in the canonical transcript were analyzed *in silico* with SIFT, PolyPhen-2, PROVEAN, and GERP++. Supplementary Table 11 shows the pathogenetic nature prediction only for the canonical transcripts. None of the common nonsynonymous SNPs except for rs7232 in *MS4A6A* were predicted to have a functional impact; the minor allele of rs7232 was predicted to be possibly damaging to the MS4A6A protein according to PolyPhen-2. In contrast, all of the rare variants were predicted to have deleterious effects on protein function.

3.2. Hypothesis 2 identified IGAP SNPs are eQTLs

Table 2 shows transcript levels (gene expression) in blood that were significantly associated with the IGAP SNPs, each reaching FDR-adjusted significance level. The risk allele of rs10838725 in *CELF1* was associated with increased *MYBPC3* expression, and the protective allele of rs983392 in *MS4A6A* was associated with decreased expression of *MS4A6A* itself. *MYBPC3* expression (probe ID: 11725151_at) and *MS4A6A* expression (probe ID: 11716846_a_at) were also significantly associated with AD status.

The significant associations between the IGAP SNPs and brain gene expression in NABEC and UKBEC are shown in Table 3. In FCTX data of NABEC, rs8093731 in *DSG2* acted as an eQTL for two genes, *DLGAP1* and *NETO1*, which were highly correlated ($r^2 = 0.69$). In UKBEC, the risk allele of rs4147929 in *ABCA7*, the risk allele of rs10948363 in *CD2AP*, and the protective allele of rs3865444 in *CD33* were associated with increased *EID2B* expression in FCTX, increased *AK9* expression in FCTX, and decreased *IER2* expression in TCTX, respectively.

In comparison with non-AD samples using the merged data sets, AD cases had significantly lower expressions of *DLGAP1* and *NETO1* (potentially regulated by the *DSG2* SNP) in the four brain regions and of *EID2B* (potentially regulated by the *ABCA7* SNP) in the HIPP. Significantly higher expression of *AK9* (potentially regulated by the *CD2AP* SNP) and *IER2* (potentially regulated by the *CD33* SNP) in AD cases was seen in the EC/HIPP and EC/TCTX, respectively (Table 4).

4. Discussion

Although large GWAS have identified novel loci that are associated with altered AD risk, we have a relatively poor understanding of the functional impact of these loci. In this study, we examined possible functional effects of the IGAP SNPs on AD under two hypotheses: "the IGAP SNP is a proxy of a coding variant" and "the IGAP SNP is an eQTL." For the first hypothesis, rs6656401 in *CR1*, rs9271192 in *HLA-DRB5-DRB1*, rs2718058 in *NME8*, rs1476679 in *ZCWPW1*, rs9331896 in *CLU*, rs983392 in *MS4A6A*, and rs4147929 in *ABCA7* are proxies of common coding SNPs. In addition, the IGAP SNPs rs9271192 in *HLA-DRB5-DRB1*, rs1476679 in *ZCWPW1*, rs10792832 in *PICALM*, and rs8093731 in *DSG2* may reflect the net effect of nonsynonymous rare variants. For the second hypothesis, rs8093731 in *DSG2*, rs4147929 in *ABCA7*, rs10948363 in *CD2AP*, and rs3865444 in *CD33* are associated with gene expression, although whether these SNPs are proxies for the functional regulatory SNP or functional themselves requires further studies.

Table 1
Association of IGAP SNPs and the coding SNPs strongly correlated with the IGAP SNPs with Alzheimer's disease in two data sets: ADGC[a] and ADSP[b]

IGAP SNP				Exonic SNP				
SNP	Closest gene	ADGC[a]		SNP ID	ADGC[a]		ADSP[b]	
		OR	p-value		OR	p-value	OR	p-value
Strong LD ($r^2 \geq 0.8$)								
rs6656401	CR1	1.17	8.49×10^{-7}	rs4844600	1.17	$\mathbf{4.22 \times 10^{-7}}$	-	-
				rs2296160	1.18	$\mathbf{4.24 \times 10^{-8}}$	1.11	$\mathbf{8.56 \times 10^{-3}}$
rs9271192	HLA-DRB5	1.11	2.66×10^{-4}	rs9270303	-	-	1.15	$\mathbf{2.50 \times 10^{-4}}$
rs1476679	ZCWPW1	0.92	1.41×10^{-3}	rs2405442	0.92	$\mathbf{2.12 \times 10^{-3}}$	0.89	$\mathbf{1.09 \times 10^{-3}}$
				rs1859788	0.93	$\mathbf{5.62 \times 10^{-3}}$	0.89	$\mathbf{1.12 \times 10^{-3}}$
rs9331896	CLU	0.92	1.02×10^{-3}	rs7982	0.91	$\mathbf{3.64 \times 10^{-4}}$	0.90	$\mathbf{1.62 \times 10^{-3}}$
rs10838725	CELF1	1.05	0.054	rs2293576	1.03	0.24	1.07	0.040
rs983392	MS4A6A	0.86	7.22×10^{-9}	rs12453	0.86	$\mathbf{1.34 \times 10^{-9}}$	0.89	$\mathbf{3.60 \times 10^{-4}}$
rs4147929	ABCA7	1.12	1.39×10^{-3}	rs3752246	1.11	$\mathbf{2.70 \times 10^{-3}}$	1.18	$\mathbf{9.89 \times 10^{-5}}$
rs3865444	CD33	0.91	4.49×10^{-4}	rs12459419	0.91	$\mathbf{4.02 \times 10^{-4}}$	0.94	0.090

Bold p-value represents the statistical significance after false discovery rate adjustment.
Key: IGAP, International Genomics of Alzheimer's Project; SNP, single-nucleotide polymorphism; ADSP, Alzheimer's Disease Sequencing Project; ADGC, Alzheimer's Disease Genetic Consortium; OR, odds ratio; LD, linkage disequilibrium.
[a] Imputed genotype data from ADGC.
[b] Whole-exome sequencing data from ADSP.

4.1. CR1 SNPs

There were two common coding SNPs, rs4844600 (synonymous) and rs2296160 (nonsynonymous), in strong LD with the IGAP SNP, although the association between rs4844600 and AD could not be assessed for replication because of the lack of WES data in ADSP. CR1, located on chromosome 1q32.2 within a cluster of complement-related genes, encodes complement receptor 1, which typically acts to bind complement-labeled proteins or complexes for their clearance by the immune system (Khera and Das, 2009). Regarding AD, the CR1 protein acts a receptor for complement fragments bound to Aβ, and thus, the change in CR1 protein structure and expression levels may be related to Aβ clearance (Rogers et al., 2006). The synonymous SNP rs4844600 (E60E) is located on exon 2, the IGAP SNP rs6656401 is located between exons 4 and 5, and the nonsynonymous SNP rs2296160, which

causes an alanine-to-threonine amino acid substitution at codon position 2419 (A2419T), is located on exon 44. The SNP rs1408077 in CR1, which is in strong LD with the IGAP SNP rs6656401, was reported to be associated with loss of EC thickness (Biffi et al., 2010), and carriers of the IGAP SNP rs6656401_A had smaller local gray matter volume in the EC in young health adults, which may lead to or reflect increased risk of late-onset AD (Bralten et al., 2011). These results may indicate a causal relationship between CR1 SNPs and AD development.

4.2. HLA-DRB5-DRB1 SNPs

The IGAP SNP rs9271192 is located in an intergenic region (chromosome 6p21.32), near HLA class II genes (HLA-DR, -DQ, and -DP). There were four common coding SNPs that were in strong or moderate LD with the IGAP SNP; one nonsynonymous SNP

Table 2
Significant association of the IGAP SNPs with gene expression in blood in ADNI

IGAP SNP	Closest gene	Probe set ID[a]	Gene	eQTL association		AD association
				β	p-value[b]	p-value[c]
rs6733839	BIN1	11719631_s_at	BIN1	0.071	1.51×10^{-7}	0.28
		11746895_a_at	BIN1	0.084	2.05×10^{-6}	0.56
rs111418223	HLA-DRB5	11730933_a_at	AGPAT1	0.120	1.83×10^{-11}	0.91
		11750187_a_at	AGPAT1	0.114	6.96×10^{-9}	0.68
		11751668_a_at	AGPAT1	0.119	1.71×10^{-8}	0.94
rs1476679	ZCWPW1	11722909_a_at	GATS	0.177	1.53×10^{-17}	0.53
		11736388_a_at	TRIM4	−0.129	3.53×10^{-9}	0.95
		11743311_a_at	PILRB	−0.115	4.15×10^{-9}	0.85
		11730023_s_at	PILRB	−0.107	1.58×10^{-8}	0.78
		11730022_a_at	PILRB	−0.128	3.77×10^{-8}	0.89
		11760665_at	ZKSCAN1	0.147	2.30×10^{-7}	0.15
		11730247_a_at	PVRIG	0.086	6.07×10^{-5}	0.82
rs11771145	EPHA1	11755327_s_at	LOC154761	0.109	3.68×10^{-6}	0.40
rs28834970	PTK2B	11720981_a_at	PTK2B	0.114	6.86×10^{-18}	0.40
		11720982_s_at	PTK2B	0.086	1.18×10^{-17}	0.91
		11720980_a_at	PTK2B	0.094	7.33×10^{-12}	0.47
		11723344_at	TRIM35	−0.070	2.31×10^{-6}	0.45
rs10838725	CELF1	11725151_at	MYBPC3	0.140	1.07×10^{-7}	8.13×10^{-5}
rs983392	MS4A6A	11716846_a_at	MS4A6A	−0.082	2.34×10^{-12}	4.97×10^{-3}
		11751570_a_at	MS4A4A	−0.150	1.15×10^{-6}	0.85
		11732865_a_at	MS4A4A	−0.179	1.59×10^{-6}	0.68

Key: AD, Alzheimer's disease; IGAP, International Genomics of Alzheimer's Project; SNP, single-nucleotide polymorphism; ADNI, Alzheimer's Disease Neuroimaging Initiative; eQTL, expression quantitative trait locus.
[a] Probe set IDs on Affymetrix Human Genome U219 Array.
[b] p-values less than significance level after false discovery rate adjustment were displayed.
[c] p-values calculated by analysis of covariance with the outcome of gene expression and the predictor of Alzheimer's disease status (normal/mild cognitive impairment/AD).

Table 3
Significant association of the IGAP SNPs with brain gene expression in NABEC and UKBEC

IGAP SNP	Closest gene	Probe set ID[a]	Gene expression[b]	$\hat{\beta}$	p-value[c]
NABEC					
rs8093731	DSG2	ILMN_2380779	DLGAP1	0.770	1.36×10^{-8}
		ILMN_1783168	NETO1	0.706	1.11×10^{-5}
UKBEC					
rs4147929	ABCA7	t3862068	EID2B	0.259	3.94×10^{-6}
rs10948363	CD2AP	t2969159	AK9	0.343	1.64×10^{-5}
rs3865444	CD33	t3822216	IER2	−0.255	1.96×10^{-5}

Key: IGAP, International Genomics of Alzheimer's Project; SNP, single nucleotide polymorphism; NABEC, North American Brain Expression Consortium; UKBEC, United Kingdom Brain Expression Consortium; FCTX, frontal cortex; TCTX, temporal cortex.

[a] Probe set IDs on HumanHT-12_v3 Expression BeadChips in NABEC (platform = GPL6947) and on Affymetrix Exon 1.0 ST Arrays in UKBEC (platform = GPL5175).

[b] DLGAP1, NETO1, EID2B, and AK9 expressions in FCTX, and IER2 expression in TCTX.

[c] p-values less than significance level after false discovery rate adjustment are displayed.

rs9270303 in *HLA-DRB1*, three synonymous SNPs rs2308759 in *HLA-DRB1*, rs1049092, and rs1049086 in *HLA-DQB1*. Because of missing data in ADGC, the association tests of rs9270303, rs2308759, and rs1049092 with AD could not be performed in the discovery analysis (Table 1 and Supplementary Table 7). We also identified two nonsynonymous rare variants, rs11575848 and rs2070600, that are potentially coinherited with the IGAP SNP (D' = 1) and are located in the major histocompatibility complex class III region. *LY6G6C* (rs11575848) encodes a leukocyte antigen-6 superfamily member, and *AGER* (rs2070600) encodes a receptor for advanced glycosylation end product. The receptor for advanced glycosylation end product protein is a member of the immunoglobulin superfamily and may regulate Aβ transport across the blood-brain barrier (Tarasoff-Conway et al., 2015).

At least three classes of genes expressed from a single allele (monoallelic) are recognized to exist (Chess, 2012; Gimelbrant et al., 2007). One class is the autosomal imprinted genes regulated in a parent-of-origin–specific manner. The second class is X-inactivated. The third class of these genes is located randomly in autosomes and includes several immune system genes (Chess, 2012;

Gimelbrant et al., 2007). Because we excluded from the analysis *HLA* genes that are randomly monoallelically expressed, we did not examine the associations between the IGAP SNP rs9271192 and expression of *HLA* genes including *HLA-DRB5* and *HLA-DRB1*. Given epigenetic association between DNA methylation in *HLA-DRB5* and AD pathology (Yu et al., 2015), allele-specific expression at these loci may have a strong impact on immunobiological function related to AD.

4.3. CD2AP SNPs

Increased expression of *AK9* in the FCTX was associated with the risk allele of the *CD2AP* IGAP SNP, and *AK9* was significantly overexpressed in the EC and HIPP brain regions of AD cases. *AK9*, located on chromosome 6q21 and more than 60 Mb away from *CD2AP* locus, encodes a member of adenylate kinase family of enzymes. Adenylate kinase reversibly catalyzes the interconversion of adenine nucleotides (ATP + AMP ↔ 2 ADP) (Amiri et al., 2013).

4.4. NME8 SNPs

Two common coding SNPs were in moderate LD with the IGAP *NME8* SNP, one nonsynonymous SNP rs2722372 and one synonymous SNP rs2598044. The nonsynonymous SNP rs2722372 causes arginine-to-lysine amino acid substitution at the codon position 43 (R43K). These coding SNPs were significantly associated with AD and the associations were replicated. *NME8*, located on chromosome 7p14.1, encodes a protein with an N-terminal thioredoxin domain and C-terminal nucleoside diphosphate kinase domains. The NME8 protein is a member of NME/NM23 family. Although the function of this gene is poorly characterized, Liu et al. showed that the IGAP SNP rs2718058 had a neuroprotective effect against cognitive decline, elevated tau levels in cerebrospinal fluid, and hippocampal atrophy (Liu et al., 2014).

4.5. ZCWPW1 SNPs

There were three common coding SNPs that were in strong or moderate LD with the IGAP *ZCWPW1* SNP, one nonsynonymous SNP rs1859788 (in *PILRA*) and two synonymous SNP rs2405442 (in *PILRA*) and rs909152 (in *LRCH4*). The nonsynonymous SNP

Table 4
Associations between gene expressions identified in NABEC and UKBEC and Alzheimer's disease status in the merged data set

Probe set ID[a]	EC		FCTX		HIPP		TCTX	
	$\hat{\beta}$	p-value	$\hat{\beta}$	p-value	$\hat{\beta}$	p-value	$\hat{\beta}$	p-value
Identified in NABEC								
DLGAP1								
206489_s_at	−0.495	**5.68×10^{-3}**	−0.294	**0.013**	−0.635	**2.77×10^{-6}**	−0.318	**1.58×10^{-4}**
206490_at	−0.225	0.071	−0.202	0.039	−0.280	**3.18×10^{-4}**	−0.212	0.052
210750_s_at	−0.343	**0.013**	−0.135	0.13	−0.084	0.27	0.214	0.071
NETO1								
1552736_a_at	−0.433	**5.36×10^{-3}**	−0.210	0.14	−0.597	**1.53×10^{-5}**	−0.088	0.58
1552904_at	−0.411	**2.70×10^{-4}**	−0.115	0.085	−0.359	**5.23×10^{-6}**	−0.119	0.085
1562713_a_at	−0.548	**2.95×10^{-3}**	−0.255	**8.99×10^{-3}**	−0.615	**1.97×10^{-5}**	−0.276	**4.73×10^{-4}**
Identified in UKBEC								
EID2B								
242470_at	−0.383	0.051	−0.319	0.015	−0.424	**6.70×10^{-4}**	−0.166	0.33
AK9								
1552299_at	0.039	0.54	0.014	0.82	0.001	0.99	−0.179	0.048
1564002_a_at	0.197	**5.23×10^{-3}**	0.092	0.084	0.287	**6.38×10^{-3}**	0.083	0.41
IER2								
202081_at	0.292	**7.11×10^{-4}**	0.141	0.060	0.061	0.52	0.949	**4.48×10^{-5}**

Bold p-value represents the statistical significance after FDR adjustment.
Key: NABEC, North American Brain Expression Consortium; UKBEC, United Kingdom Brain Expression Consortium; EC, entorhinal cortex; HIPP, hippocampus; FCTX, frontal cortex; TCTX, temporal cortex.

[a] Probe set IDs on Affymetrix U133 Plus 2.0 array (platform = GPL570).

rs1859788 causes glycine-to-arginine amino acid substitution at codon position 78 (G78R). The coding SNPs which are in strong LD with the IGAP SNP were significantly associated with AD and we confirmed associations in the replication analysis. There was also one nonsynonymous rare variant rs62483572 (in *EPO*) with D' = 1 that causes an amino acid substitution of aspartic acid to asparagine at codon position 70 (D70N). This variant had a more protective effect than the nonsynonymous common SNP rs18597955 (odds ratio = 0.53 for rs62483572 vs. odds ratio = 0.89 for rs1859788 in ADSP). Erythropoietin exhibits a neuroprotective effect under various conditions of neuronal damage such as hypoxia-ischemia, and thus, the nonsynonymous rare variants may be involved in promoting maintenance of homeostasis (Siren et al., 2001).

The IGAP SNP was associated with the expression of several genes including *GATS*, *TRIM4*, *PILRB*, *ZKSCAN1*, and *PVRIG* in blood. However, we did not find significant associations between expression of these genes and AD. *ZCWPW1*, located on chromosome 7q22.1, encodes zinc finger CW (zf-CW)–type and PWWP domain containing 1. Although the function(s) of this gene are unknown, zf-CW may be involved in epigenetics as it is regarded as a member of histone modification reader modules (He et al., 2010).

The IGAP SNP was associated with the expression of several genes including *GATS*, *TRIM4*, *PILRB*, *ZKSCAN1*, and *PVRIG* in blood. However, we did not find significant associations between expression of these genes and AD. *ZCWPW1*, located on chromosome 7q22.1, encodes zf-CW–type and PWWP domain containing 1. Although the function(s) of this gene are unknown, zf-CW may be involved in epigenetics as it is regarded as a member of histone modification reader modules (He et al., 2010).

4.6. CLU SNPs

We confirmed that the synonymous *CLU* SNP rs7982 is in strong LD with the IGAP SNP rs9331896 and was protectively associated with AD. However, we found no evidence of gene expression regulation that was associated with either the *CLU* IGAP SNP or the proxy, synonymous SNP. *CLU* is located on chromosome 8p21.1, and encodes clusterin, also known as apolipoprotein J. Clusterin directly influences Aβ, regulating the conversion of Aβ into insoluble forms (Desikan et al., 2014; Yu and Tan, 2012). *CLU* has two main isoforms, nuclear *CLU* (nCLU, isoform 1) and secretory *CLU* (sCLU, isoform 2) with different functions. The sCLU form is prosurvival, whereas nCLU is proapoptotic (Shannan et al., 2006). Because the coding SNP rs7982 is synonymous, it may affect alternative splicing as Ling et al. showed that the protective SNP rs11136000 (which is in almost perfect LD with rs7982 in 1000 genomes EUR) was associated with an increased nCLU expression level (Ling et al., 2012).

4.7. CELF1 SNPs

We did not find evidence that the *CELF1* SNP rs10838725 and the proxy coding SNPs in LD with the IGAP SNP were associated with AD. However, rs10838725 acted as eQTL for *MYBPC3* expression in blood, which was associated with AD. *MYBPC3* is located on chromosome 11p11.2 and encodes cardiac myosin binding protein C expressed in heart muscle (Gautel et al., 1995). Huang et al. reported that the protective allele of rs1057233 in *CELF1* (r² = 0.17 and D' = 0.97 with the IGAP SNP rs10838725 as shown in Supplementary Table 1) was associated with decreased *MYBPC3* and *SPI1* expressions and with the higher cerebrospinal fluid Aβ$_{42}$ levels (Huang et al., 2017). Our finding that the risk allele of IGAP SNP rs10838725 in *CELF1* was associated with the increased *MYBPC3* expression and also correlated with AD status is concordant with the previous study.

4.8. MS4A6A SNPs

The IGAP SNP rs983392 located downstream of *MS4A6A* was associated with several striking gene regulatory features. We found two coding SNPs, rs12453 and rs7232, in LD with the protective IGAP SNP rs983392. The coding SNP rs7232 is nonsynonymous, causing threonine-to-serine amino acid substitution at codon position 213 (T213S), while the SNP rs12453 is synonymous (L137L). The IGAP SNP was associated with expression in blood of *MS4A6A* and other members of the *MS4A* gene family. Moreover, decreased *MS4A6A* expression was associated with AD risk (Fig. 2). These results help illustrate that the two basic hypotheses we were testing (SNP is a proxy for a coding variant; and, SNP is an eQTL) are not mutually exclusive.

MS4A6A, located on chromosome 11q12.2, encodes a member of membrane-spanning 4A gene family (membrane-spanning 4A domains, subfamily A, member 6A). *MS4A* genes are highly expressed in hematopoietic cells and involved in the regulation of calcium signaling (Ma et al., 2015). Although functions of the MS4A6A protein are still incompletely understood, it is possible that the *MS4A6A* SNPs are linked to AD via deregulation of calcium signaling implicated in neurodegenerative diseases (LaFerla, 2002; Marambaud et al., 2009).

4.9. PICALM SNPs

There was one nonsynonymous rare variant rs74547795 in *SYTL2* with D' = 1 for the IGAP SNP that causes amino acid substitution of alanine to aspartic acid at codon position 825 (A825D). *PICALM* is located on chromosome 11q14.2 and encodes a phosphatidylinositol cinding clathrin assembly protein that may be involved in Aβ clearance (Zhao et al., 2015) and synaptic neurotransmission release (Sleegers et al., 2010). On the other hand, there is no evidence that *SYTL2* is associated with AD.
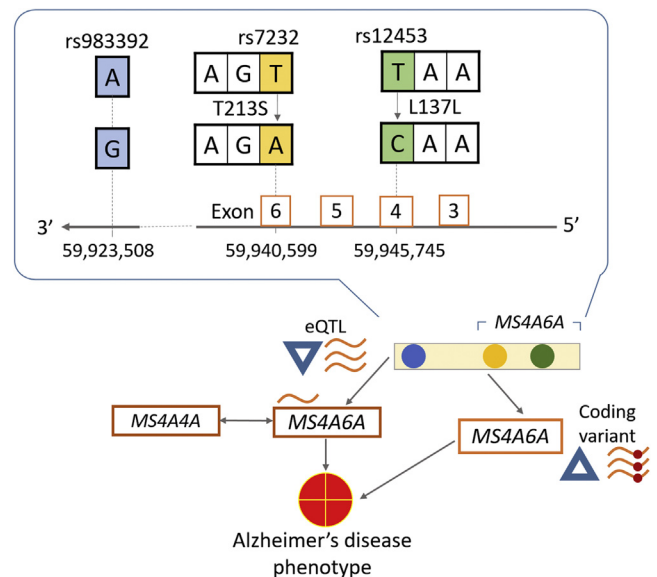
Fig. 2. Potential pathway for the relationship between *MS4A6A* SNPs, blood gene expression, and Alzheimer's disease phenotype. The relationship between *MS4A6A* IGAP SNP rs983392, blood gene expression, and Alzheimer's disease phenotype is complex and possibly multifactorial. The rs983392 SNP constitutes an eQTL for *MS4A6A* and also a proxy for a nonsynonymous exonic *MS4A6A* SNP rs7232. It is possible that the nonsynonymous SNP rs7232 is not, or is only partially responsible, for the eQTL phenomena, which may indicate parallel gene regulatory mechanism(s). Abbreviations: SNP, single-nucleotide polymorphism; IGAP, International Genomics of Alzheimer's Project; eQTL, expression quantitative trait locus.

4.10. DSG2 SNPs

There was one nonsynonymous rare variant rs111986709 located in *DSG3* with D' = 1 for the IGAP SNP. The variant causes serine-to-phenylalanine amino acid substitution at codon position 771 (S771F), and the mutation was predicted to have an impact on the DSG3 protein. *DSG2* and *DSG3* encode members of the desmoglein family. The role of these genes in AD is unknown.

The expression of both *DLGAP1* and *NETO1* were strongly associated with the *DSG2* IGAP SNP and were highly correlated with each other in the FCTX. Interestingly, these genes were significantly underexpressed in four brain regions of AD cases. Located on chromosome 18p11.31 more than 25 Mb away from *DSG2*, *DLGAP1* encodes disks large-associated protein 1 (also known as guanylate kinase−associated protein [GKAP]). *NETO1* is located on chromosome 18q22.3, more than 40 Mb away from *DSG2,* and encodes neuropilin and tolloid-like 1. Both *DLGAP1* and *NETO1* are mainly expressed in neurons of human brains (http://web.stanford.edu/group/barres_lab/brainseqMariko/brainseq2.html) (Bennett et al., 2016) and may be involved in N-methyl-D-aspartate receptor−dependent synaptic plasticity (Ng et al., 2009; Shin et al., 2012).

4.11. ABCA7 SNPs

Of six coding SNPs in strong or moderate LD with the IGAP SNP, we confirmed that two nonsynonymous SNPs [rs3752246 causing alanine-to-glycine amino acid substitution at codon position 1527 (A1527G) and rs4147934 causing serine-to-alanine amino acid substitution at codon position 2045 (S2045A)] and one synonymous SNP [rs4147930 (L1995L)] were associated with AD. The IGAP SNP acted as an eQTL for *EID2B* expression (the risk allele was associated with increased *EID2B* expression in the FCTX). However, decreased expression of *EID2B* in HIPP was associated with AD risk (i.e., the association directions are in conflict). *ABCA7*, located on chromosome 19p13.3, encodes a member of the superfamily of ATP-binding cassette transporters. *ABCA7* is expressed in hippocampal CA1 neurons and in microglia (Kim et al., 2006). *ABCA7* is involved in lipid efflux from cells to lipoproteins and has been associated with Aβ accumulation (Kim et al., 2013).

4.12. CD33 SNPs

There were two nonsynonymous SNPs in strong or moderate LD with the IGAP *CD33* SNP, rs12459419 causing alanine-to-valine amino acid substitution at codon position 14 (A14V) and rs35112940 causing glycine-to-arginine amino acid substitution at codon position 304 (G304R). Although we did not find sufficient evidence that the proxy coding SNPs in LD with the IGAP SNP rs3865444 were associated with AD, the protective allele of rs3865444 was associated with decreased *IER2* expression in the TCTX and, further, decreased *IER2* expression in the EC and TCTX had a protective effect on AD. *IER2* is located on chromosome 19p13.2 more than 35 Mb away from *CD33* and encodes immediate early response 2. *IER2* may function as a transcription factor (Takaya et al., 2009).

Malik et al. reported that the IGAP SNP rs3865444 modulated exon 2 splicing by showing that the proportion of *CD33* expressed as a *CD33* isoform−lacking exon 2 was increased in the protective allele of rs3865444; the proxy nonsynonymous SNP rs12459419 was shown to modulate exon 2 splicing efficiency (Malik et al., 2013). Additional studies are warranted to examine the association between *CD33* isoform and *IER2* expression.

There are limitations to this study. We aggregated data from many rich resources that aid in establishing a confluence of related information; however, these data sets are heterogeneous and can exhibit biases from their respective parent study designs, analytic protocols, and participant pools. A major limitation of our study is that we have limited our assessment to subjects with European-type genomic characteristics, which is connected to the fact that many research centers and clinics that contribute to the study share this underlying bias. In addition, according to the commonly used, but inexact convention, we focused on genes closest to the identified IGAP SNP.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neurobiolaging.2018.10.017.

References

Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., Sunyaev, S.R., 2010. A method and server for predicting damaging missense mutations. Nat. Methods 7, 248–249.

Alzheimer, A., 1907. Über eine eigenartige Erkrankung der Hirnrinde. Allgemeine Z. Psychiatrie Psychisch-gerichtliche Med. 64, 146–148.

Amiri, M., Conserva, F., Panayiotou, C., Karlsson, A., Solaroli, N., 2013. The human adenylate kinase 9 is a nucleoside mono- and diphosphate kinase. Int. J. Biochem. Cell Biol. 45, 925–931.

Beecham, G.W., Bis, J.C., Martin, E.R., Choi, S.H., DeStefano, A.L., van Duijn, C.M., Fornage, M., Gabriel, S.B., Koboldt, D.C., Larson, D.E., Naj, A.C., Psaty, B.M., Salerno, W., Bush, W.S., Foroud, T.M., Wijsman, E., Farrer, L.A., Goate, A., Haines, J.L., Pericak-Vance, M.A., Boerwinkle, E., Mayeux, R., Seshadri, S., Schellenberg, G., 2017. The Alzheimer's disease sequencing project: study design and sample selection. Neurol. Genet. 3, e194.

Bengtsson, H., Simpson, K., Bullard, J., Hansen, K., 2008. aroma.affymetrix: A Generic Framework in R for Analyzing Small to Very Large Affymetrix Data Sets in Bounded Memory. Tech Report #745. Department of Statistics, University of California, Berkeley.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. A. Stat. Soc. 57, 289–300.

Bennett, M.L., Bennett, F.C., Liddelow, S.A., Ajami, B., Zamanian, J.L., Fernhoff, N.B., Mulinyawe, S.B., Bohlen, C.J., Adil, A., Tucker, A., Weissman, I.L., Chang, E.F., Li, G., Grant, G.A., Hayden Gephart, M.G., Barres, B.A., 2016. New tools for studying microglia in the mouse and human CNS. Proc. Natl. Acad. Sci. U. S. A. 113, E1738–E1746.

Berchtold, N.C., Cribbs, D.H., Coleman, P.D., Rogers, J., Head, E., Kim, R., Beach, T., Miller, C., Troncoso, J., Trojanowski, J.Q., Zielke, H.R., Cotman, C.W., 2008. Gene expression changes in the course of normal brain aging are sexually dimorphic. Proc. Natl. Acad. Sci. U. S. A. 105, 15605–15610.

Biffi, A., Anderson, C.D., Desikan, R.S., Sabuncu, M., Cortellini, L., Schmansky, N., Salat, D., Rosand, J., Alzheimer's Disease Neuroimaging, I., 2010. Genetic variation and neuroimaging measures in Alzheimer disease. Arch. Neurol. 67, 677–685.

Blalock, E.M., Buechel, H.M., Popovic, J., Geddes, J.W., Landfield, P.W., 2011. Microarray analyses of laser-captured hippocampus reveal distinct gray and white matter signatures associated with incipient Alzheimer's disease. J. Chem. Neuroanat. 42, 118–126.

Bralten, J., Franke, B., Arias-Vasquez, A., Heister, A., Brunner, H.G., Fernandez, G., Rijpkema, M., 2011. CR1 genotype is associated with entorhinal cortex volume in young healthy adults. Neurobiol. Aging 32, 2106.e7–2106.e11.

Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., Lee, J.J., 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 4, 7.

Chess, A., 2012. Mechanisms and consequences of widespread random monoallelic expression. Nat. Rev. Genet. 13, 421–428.

Choi, Y., Sims, G.E., Murphy, S., Miller, J.R., Chan, A.P., 2012. Predicting the functional effect of amino acid substitutions and indels. PLoS One 7, e46688.

Das, S., Forer, L., Schonherr, S., Sidore, C., Locke, A.E., Kwong, A., Vrieze, S.I., Chew, E.Y., Levy, S., McGue, M., Schlessinger, D., Stambolian, D., Loh, P.R., Iacono, W.G., Swaroop, A., Scott, L.J., Cucca, F., Kronenberg, F., Boehnke, M., Abecasis, G.R., Fuchsberger, C., 2016. Next-generation genotype imputation service and methods. Nat. Genet. 48, 1284–1287.

Davydov, E.V., Goode, D.L., Sirota, M., Cooper, G.M., Sidow, A., Batzoglou, S., 2010. Identifying a high fraction of the human genome to be under selective constraint using GERP++. PLoS Comput. Biol. 6, e1001025.

Desikan, R.S., Thompson, W.K., Holland, D., Hess, C.P., Brewer, J.B., Zetterberg, H., Blennow, K., Andreassen, O.A., McEvoy, L.K., Hyman, B.T., Dale, A.M., Alzheimer's

Disease Neuroimaging Initiative, G, 2014. The role of clusterin in amyloid-beta-associated neurodegeneration. JAMA Neurol. 71, 180—187.

Gautel, M., Zuffardi, O., Freiburg, A., Labeit, S., 1995. Phosphorylation switches specific for the cardiac isoform of myosin binding protein-C: a modulator of cardiac contraction? EMBO J. 14, 1952—1960.

1000 Genomes Project Consortium, 2010. A map of human genome variation from population-scale sequencing. Nature 467, 1061—1073.

Gimelbrant, A., Hutchinson, J.N., Thompson, B.R., Chess, A., 2007. Widespread monoallelic expression on human autosomes. Science 318, 1136—1140.

Hardy, J., Selkoe, D.J., 2002. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. Science 297, 353—356.

Harold, D., Abraham, R., Hollingworth, P., Sims, R., Gerrish, A., Hamshere, M.L., Pahwa, J.S., Moskvina, V., Dowzell, K., Williams, A., Jones, N., Thomas, C., Stretton, A., Morgan, A.R., Lovestone, S., Powell, J., Proitsi, P., Lupton, M.K., Brayne, C., Rubinsztein, D.C., Gill, M., Lawlor, B., Lynch, A., Morgan, K., Brown, K.S., Passmore, P.A., Craig, D., McGuinness, B., Todd, S., Holmes, C., Mann, D., Smith, A.D., Love, S., Kehoe, P.G., Hardy, J., Mead, S., Fox, N., Rossor, M., Collinge, J., Maier, W., Jessen, F., Schurmann, B., Heun, R., van den Bussche, H., Heuser, I., Kornhuber, J., Wiltfang, J., Dichgans, M., Frolich, L., Hampel, H., Hull, M., Rujescu, D., Goate, A.M., Kauwe, J.S., Cruchaga, C., Nowotny, P., Morris, J.C., Mayo, K., Sleegers, K., Bettens, K., Engelborghs, S., De Deyn, P.P., Van Broeckhoven, C., Livingston, G., Bass, N.J., Gurling, H., McQuillin, A., Gwilliam, R., Deloukas, P., Al-Chalabi, A., Shaw, C.E., Tsolaki, M., Singleton, A.B., Guerreiro, R., Muhleisen, T.W., Nothen, M.M., Moebus, S., Jockel, K.H., Klopp, N., Wichmann, H.E., Carrasquillo, M.M., Pankratz, V.S., Younkin, S.G., Holmans, P.A., O'Donovan, M., Owen, M.J., Williams, J., 2009. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nat. Genet. 41, 1088—1093.

He, F., Umehara, T., Saito, K., Harada, T., Watanabe, S., Yabuki, T., Kigawa, T., Takahashi, M., Kuwasako, K., Tsuda, K., Matsuda, T., Aoki, M., Seki, E., Kobayashi, N., Guntert, P., Yokoyama, S., Muto, Y., 2010. Structural insight into the zinc finger CW domain as a histone modification reader. Structure 18, 1127—1139.

Hernandez, D.G., Nalls, M.A., Moore, M., Chong, S., Dillman, A., Trabzuni, D., Gibbs, J.R., Ryten, M., Arepalli, S., Weale, M.E., Zonderman, A.B., Troncoso, J., O'Brien, R., Walker, R., Smith, C., Bandinelli, S., Traynor, B.J., Hardy, J., Singleton, A.B., Cookson, M.R., 2012. Integration of GWAS SNPs and tissue specific expression profiling reveal discrete eQTLs for human traits in blood and brain. Neurobiol. Dis. 47, 20—28.

Hollingworth, P., Harold, D., Sims, R., Gerrish, A., Lambert, J.C., Carrasquillo, M.M., Abraham, R., Hamshere, M.L., Pahwa, J.S., Moskvina, V., Dowzell, K., Jones, N., Stretton, A., Thomas, C., Richards, A., Ivanov, D., Widdowson, C., Chapman, J., Lovestone, S., Powell, J., Proitsi, P., Lupton, M.K., Brayne, C., Rubinsztein, D.C., Gill, M., Lawlor, B., Lynch, A., Brown, K.S., Passmore, P.A., Craig, D., McGuinness, B., Todd, S., Holmes, C., Mann, D., Smith, A.D., Beaumont, H., Warden, D., Wilcock, G., Love, S., Kehoe, P.G., Hooper, N.M., Vardy, E.R., Hardy, J., Mead, S., Fox, N.C., Rossor, M., Collinge, J., Maier, W., Jessen, F., Ruther, E., Schurmann, B., Heun, R., Kolsch, H., van den Bussche, H., Heuser, I., Kornhuber, J., Wiltfang, J., Dichgans, M., Frolich, L., Hampel, H., Gallacher, J., Hull, M., Rujescu, D., Giegling, I., Goate, A.M., Kauwe, J.S., Cruchaga, C., Nowotny, P., Morris, J.C., Mayo, K., Sleegers, K., Bettens, K., Engelborghs, S., De Deyn, P.P., Van Broeckhoven, C., Livingston, G., Bass, N.J., Gurling, H., McQuillin, A., Gwilliam, R., Deloukas, P., Al-Chalabi, A., Shaw, C.E., Tsolaki, M., Singleton, A.B., Guerreiro, R., Muhleisen, T.W., Nothen, M.M., Moebus, S., Jockel, K.H., Klopp, N., Wichmann, H.E., Pankratz, V.S., Sando, S.B., Aasly, J.O., Barcikowska, M., Wszolek, Z.K., Dickson, D.W., Graff-Radford, N.R., Petersen, R.C., , Alzheimer's Disease Neuroimaging, I, van Duijn, C.M., Breteler, M.M., Ikram, M.A., DeStefano, A.L., Fitzpatrick, A.L., Lopez, O., Launer, L.J., Seshadri, S., , Consortium, C, Berr, C., Campion, D., Epelbaum, J., Dartigues, J.F., Tzourio, C., Alperovitch, A., Lathrop, M., consortium, E., Feulner, T.M., Friedrich, P., Riehle, C., Krawczak, M., Schreiber, S., Mayhaus, M., Nicolhaus, S., Wagenpfeil, S., Steinberg, S., Stefansson, H., Stefansson, K., Snaedal, J., Bjornsson, S., Jonsson, P.V., Chouraki, V., Genier-Boley, B., Hiltunen, M., Soininen, H., Combarros, O., Zelenika, D., Delepine, M., Bullido, M.J., Pasquier, F., Mateo, I., Frank-Garcia, A., Porcellini, E., Hanon, O., Coto, E., Alvarez, V., Bosco, P., Siciliano, G., Mancuso, M., Panza, F., Solfrizzi, V., Nacmias, B., Sorbi, S., Bossu, P., Piccardi, P., Arosio, B., Annoni, G., Seripa, D., Pilotto, A., Scarpini, E., Galimberti, D., Brice, A., Hannequin, D., Licastro, F., Jones, L., Holmans, P.A., Jonsson, T., Riemenschneider, M., Morgan, K., Younkin, S.G., Owen, M.J., O'Donovan, M., Amouyel, P., Williams, J., 2011. Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. Nat. Genet. 43, 429—435.

Huang, K.L., Marcora, E., Pimenova, A.A., Di Narzo, A.F., Kapoor, M., Jin, S.C., Harari, O., Bertelsen, S., Fairfax, B.P., Czajkowski, J., Chouraki, V., Grenier-Boley, B., Bellenguez, C., Deming, Y., McKenzie, A., Raj, T., Renton, A.E., Budde, J., Smith, A., Fitzpatrick, A., Bis, J.C., DeStefano, A., Adams, H.H.H., Ikram, M.A., van der Lee, S., Del-Aguila, J.L., Fernandez, M.V., , International Genomics of Alzheimer's, P., Alzheimer's Disease Neuroimaging, I, Ibanez, L., Sims, R., Escott-Price, V., Mayeux, R., Haines, J.L., Farrer, L.A., Pericak-Vance, M.A., Lambert, J.C., van Duijn, C., Launer, L., Seshadri, S., Williams, J., Amouyel, P., Schellenberg, G.D., Zhang, B., Borecki, I., Kauwe, J.S.K., Cruchaga, C., Hao, K., Goate, A.M., 2017. A common haplotype lowers PU.1 expression in myeloid cells and delays onset of Alzheimer's disease. Nat. Neurosci. 20, 1052—1061.

Jayadev, S., Leverenz, J.B., Steinbart, E., Stahl, J., Klunk, W., Yu, C.E., Bird, T.D., 2010. Alzheimer's disease phenotypes and genotypes associated with mutations in presenilin 2. Brain 133 (Pt 4), 1143—1154.

Katsumata, Y., Nelson, P.T., Ellingson, S.R., Fardo, D.W., 2017. Gene-based association study of genes linked to hippocampal sclerosis of aging neuropathology: GRN, TMEM106B, ABCC9, and KCNMB2. Neurobiol. Aging 53, 193.e17—193.e25.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., Haussler, D., 2002. The human genome browser at UCSC. Genome Res. 12, 996—1006. Article published online before print in May 2002.

Khera, R., Das, N., 2009. Complement Receptor 1: disease associations and therapeutic implications. Mol. Immunol. 46, 761—772.

Kim, W.S., Guillemin, G.J., Glaros, E.N., Lim, C.K., Garner, B., 2006. Quantitation of ATP-binding cassette subfamily-A transporter gene expression in primary human brain cells. Neuroreport 17, 891—896.

Kim, W.S., Li, H., Ruberu, K., Chan, S., Elliott, D.A., Low, J.K., Cheng, D., Karl, T., Garner, B., 2013. Deletion of Abca7 increases cerebral amyloid-beta accumulation in the J20 mouse model of Alzheimer's disease. J. Neurosci. 33, 4387—4394.

LaFerla, F.M., 2002. Calcium dyshomeostasis and intracellular signalling in Alzheimer's disease. Nat. Rev. Neurosci. 3, 862—872.

Lambert, J.C., Heath, S., Even, G., Campion, D., Sleegers, K., Hiltunen, M., Combarros, O., Zelenika, D., Bullido, M.J., Tavernier, B., Letenneur, L., Bettens, K., Berr, C., Pasquier, F., Fievet, N., Barberger-Gateau, P., Engelborghs, S., De Deyn, P., Mateo, I., Franck, A., Helisalmi, S., Porcellini, E., Hanon, O., , European Alzheimer's Disease Initiative, I., de Pancorbo, M.M., Lendon, C., Dufouil, C., Jaillard, C., Leveillard, T., Alvarez, V., Bosco, P., Mancuso, M., Panza, F., Nacmias, B., Bossu, P., Piccardi, P., Annoni, G., Seripa, D., Galimberti, D., Hannequin, D., Licastro, F., Soininen, H., Ritchie, K., Blanche, H., Dartigues, J.F., Tzourio, C., Gut, I., Van Broeckhoven, C., Alperovitch, A., Lathrop, M., Amouyel, P., 2009. Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. Nat. Genet. 41, 1094—1099.

Lambert, J.C., Ibrahim-Verbaas, C.A., Harold, D., Naj, A.C., Sims, R., Bellenguez, C., DeStafano, A.L., Bis, J.C., Beecham, G.W., Grenier-Boley, B., Russo, G., Thornton-Wells, T.A., Jones, N., Smith, A.V., Chouraki, V., Thomas, C., Ikram, M.A., Zelenika, D., Vardarajan, B.N., Kamatani, Y., Lin, C.F., Gerrish, A., Schmidt, H., Kunkle, B., Dunstan, M.L., Ruiz, A., Bihoreau, M.T., Choi, S.H., Reitz, C., Pasquier, F., Cruchaga, C., Craig, D., Amin, N., Berr, C., Lopez, O.L., De Jager, P.L., Deramecourt, V., Johnston, J.A., Evans, D., Lovestone, S., Letenneur, L., Moron, F.J., Rubinsztein, D.C., Eiriksdottir, G., Sleegers, K., Goate, A.M., Fievet, N., Huentelman, M.W., Gill, M., Brown, K., Kamboh, M.I., Keller, L., Barberger-Gateau, P., McGuiness, B., Larson, E.B., Green, R., Myers, A.J., Dufouil, C., Todd, S., Wallon, D., Love, S., Rogaeva, E., Gallacher, J., St George-Hyslop, P., Clarimon, J., Lleo, A., Bayer, A., Tsuang, D.W., Yu, L., Tsolaki, M., Bossu, P., Spalletta, G., Proitsi, P., Collinge, J., Sorbi, S., Sanchez-Garcia, F., Fox, N.C., Hardy, J., Deniz Naranjo, M.C., Bosco, P., Clarke, R., Brayne, C., Galimberti, D., Mancuso, M., Matthews, F., , European Alzheimer's Disease, I., Genetic, Environmental Risk in Alzheimer's, D., Alzheimer's Disease Genetic, C., Cohorts for, H., Aging Research in Genomic, E., Moebus, S., Mecocci, P., Del Zompo, M., Maier, W., Hampel, H., Pilotto, A., Bullido, M., Panza, F., Caffarra, P., Nacmias, B., Gilbert, J.R., Mayhaus, M., Lannefelt, L., Hakonarson, H., Pichler, S., Carrasquillo, M.M., Ingelsson, M., Beekly, D., Alvarez, V., Zou, F., Valladares, O., Younkin, S.G., Coto, E., Hamilton-Nelson, K.L., Gu, W., Razquin, C., Pastor, P., Mateo, I., Owen, M.J., Faber, K.M., Jonsson, P.V., Combarros, O., O'Donovan, M.C., Cantwell, L.B., Soininen, H., Blacker, D., Mead, S., Mosley Jr., T.H., Bennett, D.A., Harris, T.B., Fratiglioni, L., Holmes, C., de Bruijn, R.F., Passmore, P., Montine, T.J., Bettens, K., Rotter, J.I., Brice, A., Morgan, K., Foroud, T.M., Kukull, W.A., Hannequin, D., Powell, J.F., Nalls, M.A., Ritchie, K., Lunetta, K.L., Kauwe, J.S., Boerwinkle, E., Riemenschneider, M., Boada, M., Hiltuenen, M., Martin, E.R., Schmidt, R., Rujescu, D., Wang, L.S., Dartigues, J.F., Mayeux, R., Tzourio, C., Hofman, A., Nothen, M.M., Graff, C., Psaty, B.M., Jones, L., Haines, J.L., Holmans, P.A., Lathrop, M., Pericak-Vance, M.A., Launer, L.J., Farrer, L.A., van Duijn, C.M., Van Broeckhoven, C., Moskvina, V., Seshadri, S., Williams, J., Schellenberg, G.D., Amouyel, P., 2013. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat. Genet. 45, 1452—1458.

Leek, J.T., Johnson, W.E., Parker, H.S., Jaffe, A.E., Storey, J.D., 2012. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics 28, 882—883.

Levy-Lahad, E., Wasco, W., Poorkaj, P., Romano, D.M., Oshima, J., Pettingell, W.H., Yu, C.E., Jondro, P.D., Schmidt, S.D., Wang, K., Crowley, A.C., Fu, Y.H., Guenette, S.Y., Galas, D., Nemens, E., Wijsman, E.M., Bird, T.D., Schellenberg, G.D., Tanzi, R.E., 1995. Candidate gene for the chromosome 1 familial Alzheimer's disease locus. Science 269, 973—977.

Lewontin, R.C., 1964. The interaction of selection and linkage. I. General considerations; heterotic models. Genetics 49, 49—67.

Liang, W.S., Dunckley, T., Beach, T.G., Grover, A., Mastroeni, D., Walker, D.G., Caselli, R.J., Kukull, W.A., McKeel, D., Morris, J.C., Hulette, C., Schmechel, D., Alexander, G.E., Reiman, E.M., Rogers, J., Stephan, D.A., 2007. Gene expression profiles in anatomically and functionally distinct regions of the normal aged human brain. Physiol. Genomics 28, 311—322.

Lichtenthaler, S.F., Wang, R., Grimm, H., Uljon, S.N., Masters, C.L., Beyreuther, K., 1999. Mechanism of the cleavage specificity of Alzheimer's disease gamma-secretase identified by phenylalanine-scanning mutagenesis of the transmembrane domain of the amyloid precursor protein. Proc. Natl. Acad. Sci. U. S. A. 96, 3053—3058.

Ling, I.F., Bhongsatiern, J., Simpson, J.F., Fardo, D.W., Estus, S., 2012. Genetics of clusterin isoform expression and Alzheimer's disease risk. PLoS One 7, e33923.

Liu, Y., Yu, J.T., Wang, H.F., Hao, X.K., Yang, Y.F., Jiang, T., Zhu, X.C., Cao, L., Zhang, D.Q., Tan, L., 2014. Association between NME8 locus polymorphism and cognitive decline, cerebrospinal fluid and neuroimaging biomarkers in Alzheimer's disease. PLoS One 9, e114777.

Loh, P.R., Danecek, P., Palamara, P.F., Fuchsberger, C., A Reshef, Y., K Finucane, H., Schoenherr, S., Forer, L., McCarthy, S., Abecasis, G.R., Durbin, R., L Price, A., 2016. Reference-based phasing using the haplotype reference consortium panel. Nat. Genet. 48, 1443–1448.

Lu, T., Aron, L., Zullo, J., Pan, Y., Kim, H., Chen, Y., Yang, T.H., Kim, H.M., Drake, D., Liu, X.S., Bennett, D.A., Colaiacovo, M.P., Yankner, B.A., 2014. REST and stress resistance in ageing and Alzheimer's disease. Nature 507, 448–454.

Ma, J., Yu, J.T., Tan, L., 2015. MS4A cluster in Alzheimer's disease. Mol. Neurobiol. 51, 1240–1248.

Malik, M., Simpson, J.F., Parikh, I., Wilfred, B.R., Fardo, D.W., Nelson, P.T., Estus, S., 2013. CD33 Alzheimer's risk-altering polymorphism, CD33 expression, and exon 2 splicing. J. Neurosci. 33, 13320–13325.

Mancuso, M., Orsucci, D., Siciliano, G., Murri, L., 2008. Mitochondria, mitochondrial DNA and Alzheimer's disease. What comes first? Curr. Alzheimer Res. 5, 457–468.

Marambaud, P., Dreses-Werringloer, U., Vingtdeux, V., 2009. Calcium signaling in neurodegeneration. Mol. Neurodegener. 4, 20.

McLaren, W., Pritchard, B., Rios, D., Chen, Y., Flicek, P., Cunningham, F., 2010. Deriving the consequences of genomic variants with the ensembl API and SNP effect predictor. Bioinformatics 26, 2069–2070.

Mockenhaupt, S., Makeyev, E.V., 2015. Non-coding functions of alternative pre-mRNA splicing in development. Semin. Cell Dev. Biol. 47-48, 32–39.

Mullan, M., Crawford, F., Axelman, K., Houlden, H., Lilius, L., Winblad, B., Lannfelt, L., 1992. A pathogenic mutation for probable Alzheimer's disease in the APP gene at the N-terminus of beta-amyloid. Nat. Genet. 1, 345–347.

Naj, A.C., Jun, G., Beecham, G.W., Wang, L.S., Vardarajan, B.N., Buros, J., Gallins, P.J., Buxbaum, J.D., Jarvik, G.P., Crane, P.K., Larson, E.B., Bird, T.D., Boeve, B.F., Graff-Radford, N.R., De Jager, P.L., Evans, D., Schneider, J.A., Carrasquillo, M.M., Ertekin-Taner, N., Younkin, S.G., Cruchaga, C., Kauwe, J.S., Nowotny, P., Kramer, P., Hardy, J., Huentelman, M.J., Myers, A.J., Barmada, M.M., Demirci, F.Y., Baldwin, C.T., Green, R.C., Rogaeva, E., St George-Hyslop, P., Arnold, S.E., Barber, R., Beach, T., Bigio, E.H., Bowen, J.D., Boxer, A., Burke, J.R., Cairns, N.J., Carlson, C.S., Carney, R.M., Carroll, S.L., Chui, H.C., Clark, D.G., Corneveaux, J., Cotman, C.W., Cummings, J.L., DeCarli, C., DeKosky, S.T., Diaz-Arrastia, R., Dick, M., Dickson, D.W., Ellis, W.G., Faber, K.M., Fallon, K.B., Farlow, M.R., Ferris, S., Frosch, M.P., Galasko, D.R., Ganguli, M., Gearing, M., Geschwind, D.H., Ghetti, B., Gilbert, J.R., Gilman, S., Giordani, B., Glass, J.D., Growdon, J.H., Hamilton, R.L., Harrell, L.E., Head, E., Honig, L.S., Hulette, C.M., Hyman, B.T., Jicha, G.A., Jin, L.W., Johnson, N., Karlawish, J., Karydas, A., Kaye, J.A., Kim, R., Koo, E.H., Kowall, N.W., Lah, J.J., Levey, A.I., Lieberman, A.P., Lopez, O.L., Mack, W.J., Marson, D.C., Martiniuk, F., Mash, D.C., Masliah, E., McCormick, W.C., McCurry, S.M., McDavid, A.N., McKee, A.C., Mesulam, M., Miller, B.L., Miller, C.A., Miller, J.W., Parisi, J.E., Perl, D.P., Peskind, E., Petersen, R.C., Poon, W.W., Quinn, J.F., Rajbhandary, R.A., Raskind, M., Reisberg, B., Ringman, J.M., Roberson, E.D., Rosenberg, R.N., Sano, M., Schneider, L.S., Seeley, W., Shelanski, M.L., Slifer, M.A., Smith, C.D., Sonnen, J.A., Spina, S., Stern, R.A., Tanzi, R.E., Trojanowski, J.Q., Troncoso, J.C., Van Deerlin, V.M., Vinters, H.V., Vonsattel, J.P., Weintraub, S., Welsh-Bohmer, K.A., Williamson, J., Woltjer, R.L., Cantwell, L.B., Dombroski, B.A., Beekly, D., Lunetta, K.L., Martin, E.R., Kamboh, M.I., Saykin, A.J., Reiman, E.M., Bennett, D.A., Morris, J.C., Montine, T.J., Goate, A.M., Blacker, D., Tsuang, D.W., Hakonarson, H., Kukull, W.A., Foroud, T.M., Haines, J.L., Mayeux, R., Pericak-Vance, M.A., Farrer, L.A., Schellenberg, G.D., 2011. Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. Nat. Genet. 43, 436–441.

Ng, P.C., Henikoff, S., 2003. SIFT: predicting amino acid changes that affect protein function. Nucleic Acids Res. 31, 3812–3814.

Ng, D., Pitcher, G.M., Szilard, R.K., Sertie, A., Kanisek, M., Clapcote, S.J., Lipina, T., Kalia, L.V., Joo, D., McKerlie, C., Cortez, M., Roder, J.C., Salter, M.W., McInnes, R.R., 2009. Neto1 is a novel CUB-domain NMDA receptor-interacting protein required for synaptic plasticity and learning. Plos Biol. 7, e41.

Nilsberth, C., Westlind-Danielsson, A., Eckman, C.B., Condron, M.M., Axelman, K., Forsell, C., Stenh, C., Luthman, J., Teplow, D.B., Younkin, S.G., Naslund, J., Lannfelt, L., 2001. The 'Arctic' APP mutation (E693G) causes Alzheimer's disease by enhanced Abeta protofibril formation. Nat. Neurosci. 4, 887–893.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., Sham, P.C., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575.

Querfurth, H.W., LaFerla, F.M., 2010. Alzheimer's disease. N. Engl. J. Med. 362, 329–344.

Rabbani, B., Tekin, M., Mahdieh, N., 2014. The promise of whole-exome sequencing in medical genetics. J. Hum. Genet. 59, 5–15.

Reitz, C., Brayne, C., Mayeux, R., 2011. Epidemiology of Alzheimer disease. Nat. Rev. Neurol. 7, 137–152.

Rogers, J., Li, R., Mastroeni, D., Grover, A., Leonard, B., Ahern, G., Cao, P., Kolody, H., Vedders, L., Kolb, W.P., Sabbagh, M., 2006. Peripheral clearance of amyloid beta peptide by complement C3-dependent adherence to erythrocytes. Neurobiol. Aging 27, 1733–1739.

Sauna, Z.E., Kimchi-Sarfaty, C., 2011. Understanding the contribution of synonymous mutations to human disease. Nat. Rev. Genet. 12, 683–691.

Selkoe, D.J., American College of, P., American Physiological, S., 2004. Alzheimer disease: mechanistic understanding predicts novel therapies. Ann. Intern. Med. 140, 627–638.

Seshadri, S., Fitzpatrick, A.L., Ikram, M.A., DeStefano, A.L., Gudnason, V., Boada, M., Bis, J.C., Smith, A.V., Carassquillo, M.M., Lambert, J.C., Harold, D., Schrijvers, E.M., Ramirez-Lorca, R., Debette, S., Longstreth Jr., W.T., Janssens, A.C., Pankratz, V.S., Dartigues, J.F., Hollingworth, P., Aspelund, T., Hernandez, I., Beiser, A., Kuller, L.H., Koudstaal, P.J., Dickson, D.W., Tzourio, C., Abraham, R., Antunez, C., Du, Y., Rotter, J.I., Aulchenko, Y.S., Harris, T.B., Petersen, R.C., Berr, C., Owen, M.J., Lopez-Arrieta, J., Varadarajan, B.N., Becker, J.T., Rivadeneira, F., Nalls, M.A., Graff-Radford, N.R., Campion, D., Auerbach, S., Rice, K., Hofman, A., Jonsson, P.V., Schmidt, H., Lathrop, M., Mosley, T.H., Au, R., Psaty, B.M., Uitterlinden, A.G., Farrer, L.A., Lumley, T., Ruiz, A., Williams, J., Amouyel, P., Younkin, S.G., Wolf, P.A., Launer, L.J., Lopez, O.L., van Duijn, C.M., Breteler, M.M., Consortium, C., Consortium, G., Consortium, E, 2010. Genome-wide analysis of genetic loci associated with Alzheimer disease. JAMA 303, 1832–1840.

Shannan, B., Seifert, M., Boothman, D.A., Tilgen, W., Reichrath, J., 2006. Clusterin and DNA repair: a new function in cancer for a key player in apoptosis and cell cycle control. J. Mol. Histol. 37, 183–188.

Shin, S.M., Zhang, N., Hansen, J., Gerges, N.Z., Pak, D.T., Sheng, M., Lee, S.H., 2012. GKAP orchestrates activity-dependent postsynaptic protein remodeling and homeostatic scaling. Nat. Neurosci. 15, 1655–1666.

Simpson, J.E., Ince, P.G., Shaw, P.J., Heath, P.R., Raman, R., Garwood, C.J., Gelsthorpe, C., Baxter, L., Forster, G., Matthews, F.E., Brayne, C., Wharton, S.B., Function, M.R.C.C., Ageing Neuropathology Study, G, 2011. Microarray analysis of the astrocyte transcriptome in the aging brain: relationship to Alzheimer's pathology and APOE genotype. Neurobiol. Aging 32, 1795–1807.

Simpson, J.E., Ince, P.G., Minett, T., Matthews, F.E., Heath, P.R., Shaw, P.J., Goodall, E., Garwood, C.J., Ratcliffe, L.E., Brayne, C., Rattray, M., Wharton, S.B., Function, M.R.C.C., Ageing Neuropathology Study, G, 2016. Neuronal DNA damage response-associated dysregulation of signalling pathways and cholesterol metabolism at the earliest stages of Alzheimer-type pathology. Neuropathol. Appl. Neurobiol. 42, 167–179.

Siren, A.L., Knerlich, F., Poser, W., Gleiter, C.H., Bruck, W., Ehrenreich, H., 2001. Erythropoietin and erythropoietin receptor in human ischemic/hypoxic brain. Acta Neuropathol. 101, 271–276.

Sleegers, K., Lambert, J.C., Bertram, L., Cruts, M., Amouyel, P., Van Broeckhoven, C., 2010. The pursuit of susceptibility genes for Alzheimer's disease: progress and prospects. Trends Genet. 26, 84–93.

Takaya, T., Kasatani, K., Noguchi, S., Nikawa, J., 2009. Functional analyses of immediate early gene ETR101 expressed in yeast. Biosci. Biotechnol. Biochem. 73, 1653–1660.

Tarasoff-Conway, J.M., Carare, R.O., Osorio, R.S., Glodzik, L., Butler, T., Fieremans, E., Axel, L., Rusinek, H., Nicholson, C., Zlokovic, B.V., Frangione, B., Blennow, K., Menard, J., Zetterberg, H., Wisniewski, T., de Leon, M.J., 2015. Clearance systems in the brain-implications for Alzheimer disease. Nat. Rev. Neurol. 11, 457–470.

Trabzuni, D., Ryten, M., Walker, R., Smith, C., Imran, S., Ramasamy, A., Weale, M.E., Hardy, J., 2011. Quality control parameters on a large dataset of regionally dissected human control brains for whole genome expression studies. J. Neurochem. 119, 275–282.

van Es, M.A., van den Berg, L.H., 2009. Alzheimer's disease beyond APOE. Nat. Genet. 41, 1047–1048.

Walker, E.S., Martinez, M., Brunkan, A.L., Goate, A., 2005. Presenilin 2 familial Alzheimer's disease mutations result in partial loss of function and dramatic changes in Abeta 42/40 ratios. J. Neurochem. 92, 294–301.

Yang, J., Lee, S.H., Goddard, M.E., Visscher, P.M., 2011. GCTA: a tool for genome-wide complex trait analysis. Am. J. Hum. Genet. 88, 76–82.

Yu, J.T., Tan, L., 2012. The role of clusterin in Alzheimer's disease: pathways, pathogenesis, and therapy. Mol. Neurobiol. 45, 314–326.

Yu, L., Chibnik, L.B., Srivastava, G.P., Pochet, N., Yang, J., Xu, J., Kozubek, J., Obholzer, N., Leurgans, S.E., Schneider, J.A., Meissner, A., De Jager, P.L., Bennett, D.A., 2015. Association of Brain DNA methylation in SORL1, ABCA7, HLA-DRB5, SLC24A4, and BIN1 with pathological diagnosis of Alzheimer disease. JAMA Neurol. 72, 15–24.

Zhao, Z., Sagare, A.P., Ma, Q., Halliday, M.R., Kong, P., Kisler, K., Winkler, E.A., Ramanathan, A., Kanekiyo, T., Bu, G., Owens, N.C., Rege, S.V., Si, G., Ahuja, A., Zhu, D., Miller, C.A., Schneider, J.A., Maeda, M., Maeda, T., Sugawara, T., Ichida, J.K., Zlokovic, B.V., 2015. Central role for PICALM in amyloid-beta blood-brain barrier transcytosis and clearance. Nat. Neurosci. 18, 978–987.