

Linking Genetics of Brain Changes to Alzheimer's Disease: Sparse Whole Genome Association Scan of Regional MRI Volumes in the ADNI and AddNeuroMed Cohorts

Mizanur Khondoker^{a,b,*}, Stephen Newhouse^{a,b}, Eric Westman^{a,d}, J-Sebastian Muehlboeck^{a,d}, Patrizia Mecocci^e, Bruno Vellas^f, Magda Tsolaki^f, Iwona Kłoszewska^h, Hilkka Soininenⁱ, Simon Lovestone^j, Richard Dobson^{a,c,1}, Andrew Simmons^{a,c,1} and for the AddNeuroMed consortium and for the Alzheimer's Disease Neuroimaging Initiative²

^aKing's College London, Institute of Psychiatry, Psychology & Neuroscience and NIHR Biomedical Research Centre for Mental Health, London, United Kingdom

^bKing's College London, Institute of Psychiatry, Psychology & Neuroscience, Department of Biostatistics, London, United Kingdom

^cNIHR Biomedical Research Unit Dementia, London, United Kingdom

^dDepartments of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden

^eInstitute of gerontology and Geriatrics, University of Perugia, Perugia, Italy

^fINSERM U 558, University of Toulouse, Toulouse, France

^gAristotle University of Thessaloniki, Thessaloniki, Greece

^hMedical University of Lodz, Lodz, Poland

ⁱDepartment of Neurology, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland

^jDepartment of Psychiatry, University of Oxford, United Kingdom.

Handling Associate Editor: Tânia C. T. Ferraz Alves

Accepted 5 January 2015

Abstract.

Background: Alzheimer's disease (AD) is a highly heritable disease, but until recently few replicated genetic markers have been identified. Markers identified so far are likely to account for only a tiny fraction of the heritability of AD and many more genetic risk alleles are thought to be undiscovered.

Objective: Identifying genetic markers for AD using combined analysis of genetics and brain imaging data.

Methods: Imaging quantitative trait loci (iQTLs) has recently emerged as an interesting research area for linking genetics of brain changes to AD. We consider a genome-wide association scan of 109 brain-wide regional imaging phenotypes to identify genetic susceptibility loci for AD from a combined set of 1,045 subjects from the Alzheimer's Disease Neuroimaging Initiative (ADNI) and the AddNeuroMed studies. We use one-SNP-at-a-time as well as multi-SNP Hyperlasso based iQTL methods for the analysis.

¹Joint last authors.

²The ADNI data used in preparation of this article were obtained from the Alzheimer's disease Neuroimaging Initiative (ADNI) database (<http://adni.loni.usc.edu>). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI

investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

*Correspondence to: Mizanur Khondoker, Department of Biostatistics, Institute of Psychiatry, Psychology & Neuroscience, King's College London, De Crespigny Park, London SE5 8AF, United Kingdom. Tel.: +44 207 848 0305; Fax: +44 207 848 0281; E-mail: mizanur.khondoker@kcl.ac.uk.

Results: We identified several novel markers associated with AD, namely HOMER2 (rs1256429; intronic, $p = 8.7 \times 10^{-10}$), EOMES (rs2724509; flanking), JAM2 (rs2829841; intronic), and WEE1 (rs10770042; coding). The SNP rs1256429 (HOMER2) was one of the top hits in Hyperlasso as well as in the single-SNP analysis showing an association with the volume of the right thalamus and AD, a brain region reported to be linked with AD in several studies.

Conclusion: We believe that the markers identified in this study are novel additions to the existing list of genetic variants associated with AD which can be validated in future replicated studies.

Keywords: Alzheimer's disease, genome wide association study, imaging quantitative trait loci, magnetic resonance imaging, mild cognitive impairment

INTRODUCTION

Genome wide association studies (GWAS) have been successful for understanding the genetic basis of many complex diseases. Alzheimer's disease (AD) is a major cause of dementia around the globe in the elderly. It is a progressive and irreversible brain disorder characterized by gradual degradation of a person's memory, communicability, decision making, and the ability to perform daily activities. Despite being complex and highly heritable (up to 79% [1]), relatively little is known about the genetic basis of AD. The APOE (apolipoprotein E) gene, residing in chromosome 19, was the only important genetic factor known to be linked with late onset AD in the pre-GWAS era. Increasing availability of larger samples and advances in high throughput genomic and brain imaging technologies have recently led to discovery of more genetic susceptibility loci for AD [2–5]. However, the genetic markers identified so far are likely to account for only a tiny fraction of the heritability of AD and many more genetic risk alleles are thought to be undiscovered. Two types of studies have been used to investigate the genetic susceptibility loci for AD: case-control type GWAS and imaging quantitative trait loci (imaging-QTL) analysis. Case-control studies and meta-analysis of case-control studies directly test the association between case-control status and genetic variants [2, 3]. They are intuitively appealing, but require larger samples to achieve power comparable to imaging-QTLs [4, 6, 7]. Imaging-QTLs, on the other hand, test whether the effects of genetic variants, e.g., single nucleotide polymorphism (SNPs), on the quantitative phenotype such as regional brain volumes or cortical thickness measures differ between disease groups. Such analyses are usually done by testing SNP \times Disease interaction in regression type models. As pointed out by Potkin and colleagues [4, 6], imaging-QTLs have several advantages over case-control type studies. Quantitative phenotype measures are more informative than the dichotomous case-control status, and

require smaller sample size to achieve statistical power comparable to case-control association analyses. Gain in statistical power can be up to eight fold in an equivalent size imaging-QTL study [4]. This would imply that an effect requiring several thousand subjects to be detected in a case-control GWAS should be detectable from several hundred individuals in a QTL analysis.

In 2010, we reported two SNP markers (rs1925690, in an intron of ZNF292, p -value = 2.6×10^{-8} ; and rs11129640, flanking ARPP-21, p -value = 2.6×10^{-8}) in an imaging-QTL study showing associations with AD through their effects on entorhinal cortical volume and thickness measures [7]. In this study, we extend our previous work in two main ways. Firstly, our previous study focused on brain regions most commonly associated with AD such as the hippocampus and entorhinal cortex. In the current study we consider a genome-wide scan of 109 brain-wide regional imaging phenotype measures to identify genetic susceptibility loci for AD by using a combined set of 1,045 subjects from the US based Alzheimer's Disease Neuroimaging Initiative (ADNI) and the European AddNeuroMed studies. The current study therefore considers more phenotypes and is based on a larger sample.

Secondly, we undertake a simultaneous analysis of all SNPs and their interactions with the disease status using a penalized regression method (Hyperlasso [8]) in addition to the standard single-SNP imaging QTL analysis. Our previous study [7] and other similar imaging QTL analyses [4] only considered single-SNP analysis. Such analyses test each SNP separately without adjusting for the effects of other SNPs on the disease phenotype. In reality we expect SNPs not to act in isolation and polygenic approaches may hold the key to some of the missing heritability. Simultaneous analyses are intuitively more appealing and appropriate than single-SNP approach for identifying multiple associated variants as they can make a weak association more visible by adjusting for the effects of other associated SNPs, and also can weaken a false signal

by including the stronger effects of the true associated SNPs in the model.

METHODS

Data

The analyses in this study are based on a combined dataset from two studies, the US based ADNI database (<http://adni.loni.usc.edu>) [9] and the European AddNeuroMed study [10]. 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 intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as lessen 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 1,500 adults, ages 55 to 90, to participate in the research, consisting of cognitively normal older individuals, people with early or late MCI, 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. For up-to-date information, see <http://www.adni-info.org/>.

The AddNeuroMed study is a European multi-centre private-partnership project - a pilot for the FP6 Innovative Medicine Initiative. The main aim of AddNeuroMed study is to find biomarkers for AD, in particular for early diagnosis of AD. We integrated brain imaging and common SNP data from the ADNI and AddNeuroMed studies.

Subjects and samples

We started with 1,198 subjects, 818 from the ADNI and 380 from the AddNeuroMed studies. Genotype data was available for each subject on a common set of 551,342 SNPs in the merged dataset. ADNI data for 818 individuals are publicly available from the ADNI database (<http://adni.loni.usc.edu>) which includes brain imaging data (1.5T MRI scans), genotype data based on Illumina platform (Illumina, San Diego, CA), and demographic and neuropsychological information. More details of this data are available in the supplementary materials of Furney et al. [7].

AddNeuroMed is a Europe-wide multi-center study with six data collection centers: University of Kuopio (Finland), University of Perugia (Italy), Aristotle University of Thessaloniki (Greece), King's College London (United Kingdom), and University of Toulouse (France). Data collection at all centers was subject to patient consent and protocol approval by the relevant authority at each site. All subjects went through a neuropsychological screening battery at baseline. All MCI and AD subjects were free from any other form of neurological or psychiatric disorder and were not exposed to alcohol or any other substance misuse. The National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) Alzheimer's Criteria were used to diagnose AD. These criteria are one of the most commonly used in the diagnosis of AD.

Genotyping and QC of genetic data

Illumina HumanHap610-Quad BeadChip platform, containing 620,901 markers, was used for genotyping all samples. A description of the genotyping procedure of the ADNI samples can be found in Shen and colleagues [5]. Genotyping of the AddNeuroMed samples was carried out at the CNG (Centre National de Génotypage, France) as described previously by Furney and colleagues (See Supplementary Materials) [7]. ADNI and AddNeuroMed genotype data were merged using PLINK [11]. After repairing strand-flipping and removing any ambiguous SNPs, the merged data consisted of 1,198 subjects and a common set of 551,342 SNPs.

We then applied the following quality control (QC) steps to filter out subjects and SNPs based on: (i) Genotype call rate, (ii) Minor allele frequency (MAF), (iii) Hardy-Weinberg Equilibrium (HWE) test, (iv) Gender mismatches, (v) Genetic relatedness,

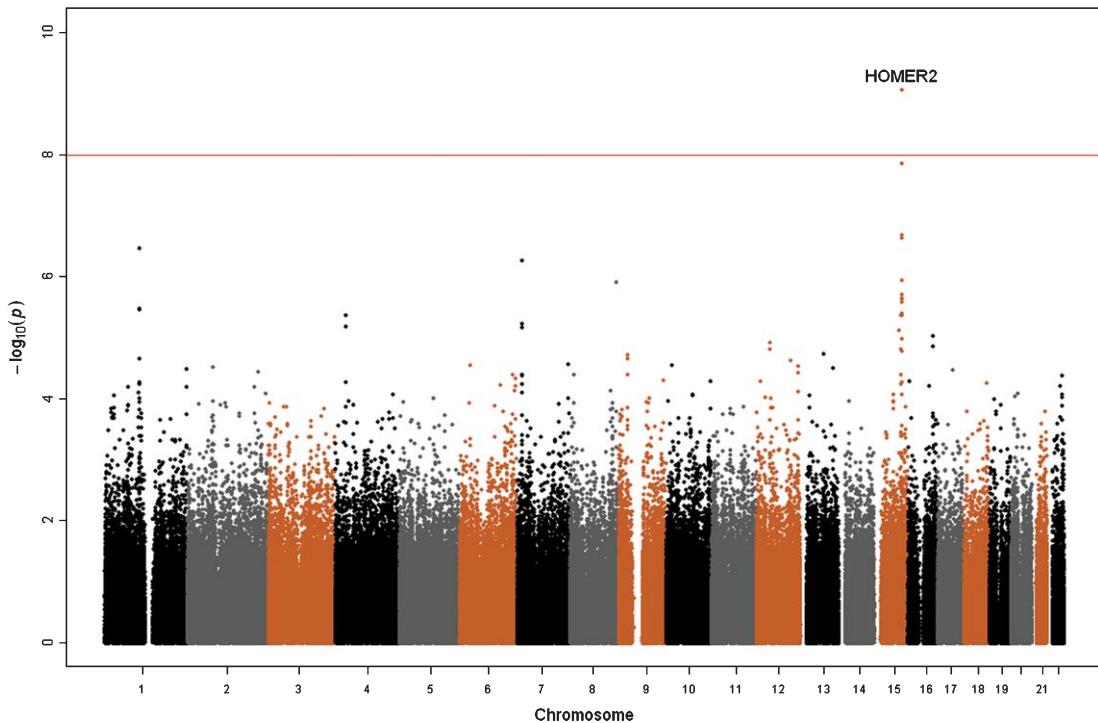


Fig. 1. Manhattan plot for the phenotype right thalamus.

(vi) Autosomal heterozygosity, and (vii) Divergent ancestry (population structure). We excluded SNPs and subjects with call rate <98%, MAF <1%, and failing Hardy-Weinberg Equilibrium (HWE) test at $p < 0.001$. We also excluded individuals whose clinical gender does not match with the genetically inferred gender, those who are genetically related, and individuals with unusually high or low autosomal heterozygosity score. For genetic relatedness, we calculated the Identical-by-descent (IBD) score and dropped one individual (lowest SNP call rate) of each pair with IBD >0.1875, which is the half-way point between 2nd and 3rd degree relatives. For filtering based on autosomal heterozygosity, we excluded subjects with unusually high or low (outside 4SD about mean) heterozygosity score (F) which can be obtained from PLINK [11] software using the $-het$ flag. For filtering divergent ancestry outliers we combined our study samples with the HapMap3 data, and excluded outliers based on the first two principal components (PCs), calculated using the EIGENSOFT package (http://genetics.med.harvard.edu/reich/Reich_Lab/Software.html). A scatterplot of the first two PCs based on the combined study samples and HapMap3 data is shown in Supplementary Fig. 1. We calculated the medians of the corresponding PCs based on the

European (CEU, ADNI, and AddNeuroMed) samples and excluded all samples in ADNI and AddNeuroMed data outside $8 \times IQR$ (inter-quartile range) about the medians as outliers.

In summary, call rate and MAF thresholds based filtering removed 22,186 SNPs. Another 2,282 SNPs failed the HWE test, leaving around 504,874 SNPs in the analysis. 17 subjects were removed for gender mismatches and for being genetically related, 41 subjects for having outlying autosomal heterozygosity score and genotype call rate < 98%, 43 for having ancestry differences (population structure), and finally 52 subjects because no imaging data were available for those subjects. This left 1,045 subjects for the analysis. A summary of baseline demographics (age, gender), Mini-Mental State Examination (MMSE) test scores, and the number of APOE ϵ 4 alleles by disease status of the available subjects is shown in Table 1.

Acquisition and QC of imaging data

Imaging data acquisition of the AddNeuroMed samples was compatible with that of the ADNI study and detailed protocols were described previously in Liu and colleagues [12]. Briefly, the imaging protocol included a high resolution sagittal 3-dimensional

Table 1

Summary of baseline demographics, Mini-Mental State Examination (MMSE) test scores, and the number of APOE ϵ 4 alleles by disease status of the selected sample

	Male			Female		
	Control	MCI	AD	Control	MCI	AD
n (% within Gender)	155(27)	282(50)	132(23)	145(31)	183(38)	148(31)
Mean (SD) – Age	75.2 (5.7)	75.3 (6.8)	75.9 (7.0)	74.9 (5.9)	73.6 (6.8)	74.6 (7.1)
MMSE	28.9 (1.5)	27.2 (1.7)	23.01 (3.3)	29.11 (1.4)	26.9 (1.7)	21.7 (3.9)
Count (% within Gender)						
APOE ϵ 4 : 0	111 (37)	139 (47)	48 (16)	102 (42)	86 (35)	55 (23)
APOE ϵ 4 : 1	38 (18)	113 (54)	60 (28)	40 (22)	76 (42)	66 (36)
APOE ϵ 4 : 2	6 (10)	30 (50)	24 (40)	3 (6)	21 (41)	27 (53)

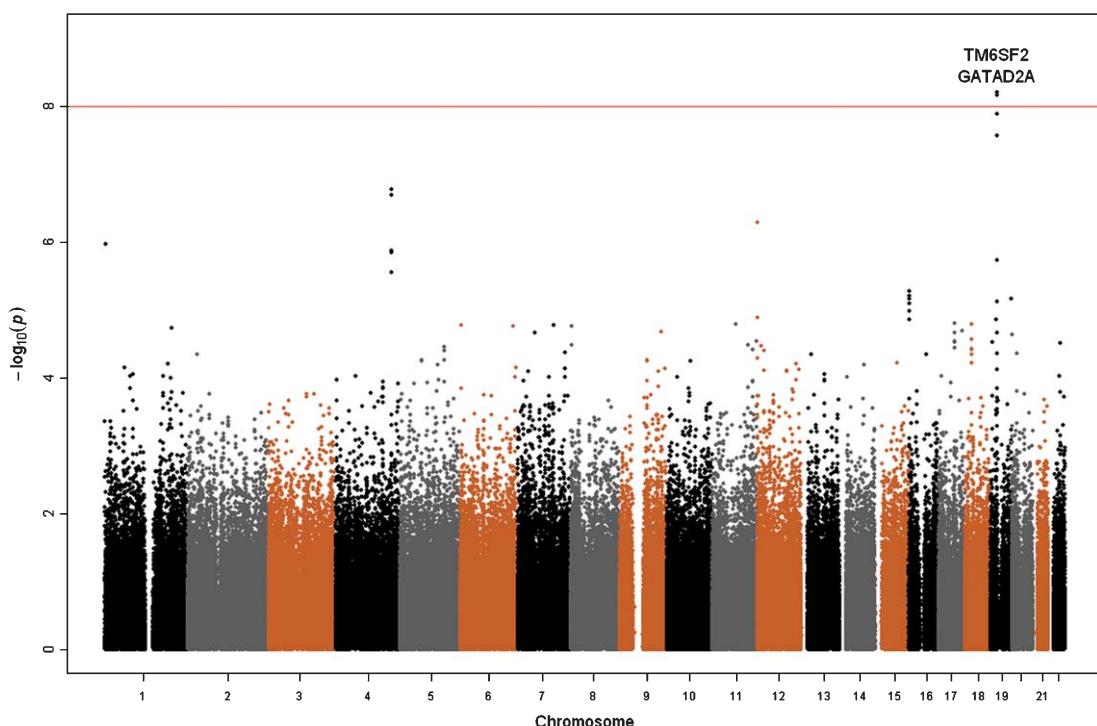


Fig. 2. Manhattan plot for the phenotype lateral orbitofrontal cortex.

T1-weighted magnetization prepared rapid gradient echo (MPRAGE) volume and axial proton density/T2-weighted fast spin echo images using 1.5-T MR systems. Image analysis uses an automated atlas based segmentation scheme developed by Fischl and colleagues [13–15] to define regional volumes and regional surface area patches. The regional volumes were normalized measures by the intracranial volume (e.g., hippocampal volume divided by intracranial volume), while the regional surface measures are characterized by average cortical thickness and were not normalized [16]. Each volume or cortical thickness measure is described by an anatomical region. Exam-

ples of the regional volume measures and regional cortical thickness measures are shown in Supplementary Figure 2.

QC of the imaging data was carried out as previously described [17, 18]. We checked the imaging data for violation of any standard assumption of the statistical model (linear regression) that we used for the analysis. Appropriate transformation, e.g., log-transformation was applied where necessary to enhance symmetry. Each of the (transformed) volume and thickness measures was normalized to have mean zero and unit standard deviation. Some example phenotype distributions and their transformed versions are shown in

Supplementary Figure 3. Volume and thickness measures outside ± 3 were treated as outliers and set to missing (9) in the plink phenotype file to be removed from the analysis.

Imaging QTL analysis

Image analysis and subsequent QC procedures resulted in 109 quantitative phenotype measures (34 left cortical thickness, 34 right cortical thickness, and 41 regional volume measures). Analyses were performed under an additive genetic model where genotypes were coded as 0, 1, and 2 representing the number of minor alleles in the genotype. We were interested in testing whether the effects of SNPs on the quantitative phenotype differ between the disease groups. Therefore our main interest is in the SNP \times Disease interaction terms in a regression type model. Such interaction effects are usually tested for each SNP separately in single-SNP regression models [4, 7]. Single-SNP analysis tests each SNP separately without taking account of the effects of other (possibly correlated) SNPs on the disease phenotype. We therefore also considered simultaneous analysis of the entire set of SNPs and the corresponding SNP \times Disease interactions using penalized regression models (Hyperlasso [8]) as a supporting or complementary analysis tool.

Single-SNP imaging QTLs

We initially considered the commonly used single-SNP analyses where we test SNP \times Disease interaction for every SNP separately considering a multiple linear regression model of the form:

$$\begin{aligned} E(y) = & \beta_0 + \beta_1 SNP + \beta_2 Disease + \beta_3 Age \\ & + \beta_4 Sex + \beta_5 APOE\epsilon 4 + \sum_{j=1}^k \beta_{pj} PC_j \\ & + \beta_6 SNP \times Disease \end{aligned} \quad (1)$$

where y is the quantitative phenotype measure, β_1 and β_6 are the SNP effect and SNP \times Disease interaction effect, respectively. PC_j is the j -th principal component to take account of population structure or any other omitted confounders ($j = 1, \dots, k$). The value of k is equal to the number of statistically significant PCs based on the Tracy-Widom distribution [19, 20] implemented in the EIGENSOFT package (http://genetics.med.harvard.edu/reich/Reich_Lab/Software.html). The EIGENSOFT package combines

functionality from population genetics methods [21] and EIGENSTRAT stratification correction method [22]. We also adjusted for Age, Sex, and APOE ϵ 4 genotype. Although APOE ϵ 4 is a well-known genetic risk factor for AD, some studies (e.g., de-Almada [23]; Berge et al.[24]) suggested protective effect of APOE ϵ 2 and APOE ϵ 3 genotype markers. To identify possible (protective) effects of APOE ϵ 2 and APOE ϵ 3 markers, we tested Disease \times APOE ϵ 2 and Disease \times APOE ϵ 3 interaction effects within the single-SNP imaging QTLs models (1) controlling for the effects of Age, Sex, APOE ϵ 4 genotype, and population structure (via the principal components).

In model (1), the disease variable is treated as quantitative (0 = Control, 1 = MCI, 2 = AD). The quantitative coding of the disease variable automatically imposes the effect homogeneity assumption, that is, it assumes that (MCI versus Control) effect is the same as (AD versus Control) effect. This, however, is a very strong assumption and may not hold in general. We therefore performed a complementary analysis by considering the disease variable as categorical. This requires replacing the original disease variable by two dummy variables: MCI (coded as MCI = 1, if MCI; 0, otherwise) and AD (coded as AD = 1, if AD; 0, otherwise) leading to an alternative form of the single-SNP model (1):

$$\begin{aligned} E(y) = & \beta_0 + \beta_1 SNP + \beta_2 MCI + \beta_3 AD + \beta_4 Age \\ & + \beta_5 Sex + \beta_6 APOE\epsilon 4 + \sum_{j=1}^k \beta_{pj} PC_j \\ & + \beta_7 SNP \times MCI + \beta_8 SNP \times AD \end{aligned} \quad (2)$$

The main interests in model (2) are the SNP \times MCI and SNP \times AD effects. Statistical significance of either one or both of the effects will indicate SNP-Disease association.

LD-based result clumping

We further processed the results of single-SNPs analysis based on empirical estimates of linkage disequilibrium between SNPs using PLINK's clumping procedure. This procedure is useful to assess how many independent loci are associated and to report the top SNP list in terms of a smaller number of clumps of correlated SNPs. For clumping the results of single-SNP analysis we used a p -value threshold, $p1 = 10^{-4}$, with all other parameters set to their default values.

Simultaneous (multi-SNP) imaging QTLs

Simultaneous analysis of the entire set of SNPs is possible using penalized regression methods which offer an attractive alternative to standard single-SNP testing in genetic association analysis. We used Hyperlasso [8], a Bayesian-inspired penalized maximum likelihood approach based on normal exponential gamma prior having a sharper peak at zero and heavier tails which facilitates heavily penalizing the near zero coefficients but leaving the larger effects less affected by the shrinkage process. This is considered as an advantage of Hyperlasso over other penalized regressions as it is preferable for variable selection when there are few true causal variables as is likely to be the case for most GWAS. Penalized regression requires choosing the value of the penalty parameter (λ) which controls the amount of shrinkage. For simultaneous analysis of all SNPs and SNP \times Disease interactions, we considered a regression model of the form:

$$\begin{aligned} E(y) = & \beta_0 + \beta_1 \text{Disease} + \beta_2 \text{Age} + \beta_3 \text{Sex} \\ & + \beta_4 \text{APOE}\varepsilon_4 + \sum_{j=1}^k \beta_{pj} \text{PC}_j \\ & + \sum_{i=1}^{n_s} \left\{ \beta_{si} \text{SNP}_i + \beta_{sdi} \text{SNP}_i \times \text{Disease} \right\} \end{aligned} \quad (3)$$

where the SNP and SNP \times Disease interaction effects for the i -th SNP are now denoted β_{si} and β_{sdi} respectively. Similar to the single-SNP model, the disease variable in model (3) is also considered to be quantitative (0 = Control, 1 = MCI, 2 = AD). We used a permutation procedure as discussed by Ayers and colleagues [25] to choose the penalty parameter in such a way that the false positive rate is controlled at 1×10^{-8} , the genome-wide significance threshold. Calculation of this threshold is further discussed in the 'Genome-wide significance threshold' section. Regression coefficients with smaller effects are driven down to zero by the shrinkage process and any non-zero terms can be considered statistically significant. We focused on the SNP \times Disease interaction terms, i.e., statistical significance of the β_{sd} terms to identify susceptibility loci for AD.

Genome-wide significance threshold

Typical single phenotype GWAS such as the Wellcome Trust Case Control Consortium (WTCCC)

paper [25] uses 5×10^{-7} as a p -value cut-off for genome-wide statistical significance. When there are multiple phenotypes involved in the association analysis, this threshold should be adjusted accordingly which will lead to a more stringent p -value cut-off. Although we have 109 phenotypes, they are not fully independent of each other and the effective number of independent phenotypes should be used for adjusting the multiple correction threshold. Using the Matrix Spectral Decomposition (<http://gump.qimr.edu.au/general/daleN/matSpD/>) based on a method by Nyholt [26], we found the effective number of independent phenotypes to be ≈ 60 . We therefore set our p -value threshold for genome-wide statistical significance at $(5 \times 10^{-7})/60 \approx 1 \times 10^{-8}$.

RESULTS

Single-SNP imaging QTLs

We initially performed single-SNP analysis for testing SNP \times Disease interaction for each of the 109 phenotypes at 504,874 SNPs (model 1 – disease coded 0, 1, 2 for Ctrl, MCI, AD). By single-SNP analysis, we refer to models where a single SNP was tested for SNP \times Disease interaction in each model adjusting for the covariates and significant principal components. The list of significant SNPs is shown in Table 2. We identified the SNP rs1256429, intronic to HOMER2, as the top hit. The marker was found to be significantly associated ($p = 8.7 \times 10^{-10}$) with the volume measure of the right thalamus and the disease status. Regression effects of this marker on the regional volume of the right thalamus are significantly different ($p = 8.7 \times 10^{-10}$) between the disease groups (Table 4). Negative coefficients for the Disease and SNP \times Disease interaction effects implies that the progression of AD results in a significant shrinkage of the volume of this brain region, and the atrophy is greater with every extra dose of the minor allele in that genetic marker (SNP).

Table 2
List of SNPs significant at $p < 3 \times 10^{-8}$ based on the single SNP model. Several markers in chromosomes 15 and 19 show significant SN P \times Disease associations.

Phenotype	SNP	CHR	p -value	Annotation
Right Thalamus Proper	rs1256429	15	8.7×10^{-10}	HOMER2
Left Lateral	rs735273	19	6.2×10^{-09}	TM6SF2
Orbitofrontal	rs6909	19	6.7×10^{-09}	GATAD2A

We found another two markers that passed the genome-wide significance threshold for the test of SNP \times Disease interaction. Regression analysis of the cortical thickness measure of left lateral orbitofrontal cortex identified two markers TM6SF2 (rs735273; flanking, $p = 6.2 \times 10^{-9}$) and GATAD2A (rs6909; flanking, $p = 6.7 \times 10^{-9}$) located in chromosome 19 showing significant interaction with the disease.

With the categorical coding of the disease variable in model (2), test for SNP \times (AD versus Ctrl) effect led to similar results to that of continuous disease coding. The marker rs1256429 was again found to show significant ($p = 6.7 \times 10^{-10}$) interaction with the (AD versus Ctrl) effect for the volume measure of the right thalamus. Markers rs735273 and rs6909 (flanking TM6SF2 and GATAD2A, respectively) were also found to be significant for the phenotype left lateral orbitofrontal cortex, but with a slightly less significant p -values ($p = 1.058 \times 10^{-8}$ and $p = 1.064 \times 10^{-8}$, respectively). None of the SNP \times (MCI versus Ctrl) interaction effects, however, passed the significance threshold. The Manhattan plots are shown in Figs. 1 and 2. The full results for LD-based clump procedure are given in the Supplementary file clumped analysis.csv. Table 5 lists selected SNPs and their p -values from the LD-based clumped results which were previously identified as associated with AD in the GWAS literature extracted from the NHGRI GWAS catalogue [27].

Simultaneous (multi-SNP) imaging QTLs

For all 109 phenotypes, Hyperlasso selected around 500 SNP \times Disease interaction terms that passed the penalization process, but most of the effects were very close to zero. We therefore performed another filtering of the significant SNP \times Disease interaction terms based on the magnitude of the penalized regres-

sion coefficients (β_{sd}). We consider SNPs as potential risk markers for AD that were found to have significant SNP \times Disease interaction effects and have penalized regression effects of magnitude 0.15 or more ($|\beta_{sd}| \geq 0.15$). We have reported both negative and positive interactions, but negative coefficients may be more interesting as the negative effects correspond to atrophy of the brain regions and are likely to be linked to AD. The resulting SNP list and the corresponding phenotypes are summarized in Table 3 and Fig. 3. Within this list, the four markers (rs1256429, rs2724509, rs2829841, and rs10770042 on chromosomes 15, 3, 21, and 11, respectively) relating to the genes HOMER2 (intronic), EOMES (flanking), JAM2 (intronic), and WEE1 (coding) appear to have interesting links with AD. These markers showed significant SNP \times Disease interaction effects for the phenotypes corresponding to right thalamus, left isthmus cingulate, brain stem and right putamen respectively. Tests of Disease \times APOE ϵ 2 and Disease \times APOE ϵ 3 interaction effects, based on the ten brain imaging measures showing association with AD in the single-SNP and simultaneous multi-SNP analyses, did not lead to any statistically significant interaction effect after correction for multiple testing (see Supplementary Table 1). This study therefore does not verify APOE ϵ 2 and APOE ϵ 3 as protective genetic markers for AD.

Previously associated AD GWAS findings

A considerable number of genetic markers for AD have been reported in the literature based on imaging-QTL studies as well as various case-control GWAS. The AlzGene database [28] (<http://www.alzgene.org>) ranked those genes in order of meta-analysis p -values and assessed for their epidemiological credibility. The top 10 genes with meta-analysis p -values < 0.00001

Table 3
List of SNPs selected by Hyperlasso. Listed SNPs can be regarded as significant at the genome-wide level ($p < 1 \times 10^{-8}$). The SNP \times Disease interaction effects with absolute coefficient values ≥ 0.15 are tabulated. Main interest is in the negative coefficients as negative effects correspond to shrinkage of the brain regions and are likely to be linked to AD

Phenotype	SNP	CHR	Coeff.	Annotation
Left Isthmus cingulate	rs2724509	3	-0.240	EOMES
Brain Stem	rs2829841	21	-0.236	JAM2
Right Putamen	rs10770042	11	-0.184	WEE1
Right Putamen	rs802101	7	-0.183	LOC645617
Right Isthmus cingulate	rs11118607	1	-0.162	MOSC1
Left Thalamus Proper	rs10129316	14	-0.156	C14orf83
Right Thalamus Proper	rs1256429	15	-0.152	HOMER2
Right Thickavg Lingual	rs10196975	2	-0.152	FLJ16124
Right Thickavg Lingual	rs2527083	8	-0.152	LOC392180
Left Lateral Ventricle	rs1265143	8	0.19	RRM2B
CSF	rs12462609	19	0.17	CACNA1A

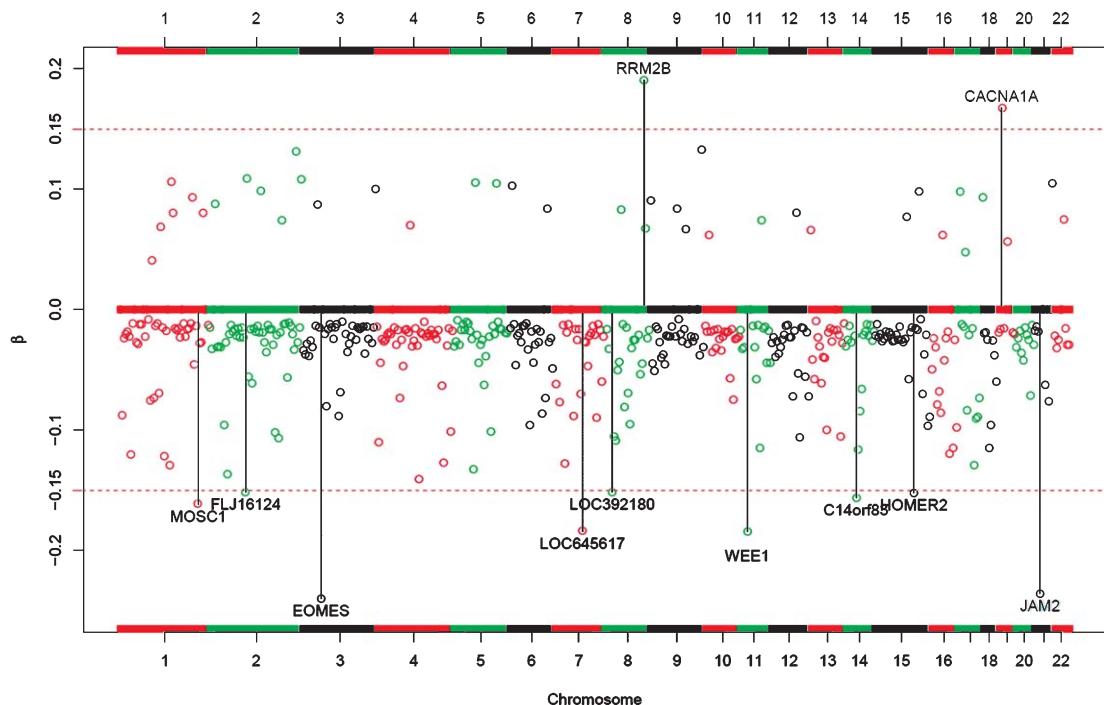


Fig. 3. Plot of the SNPs selected by Hyperlasso analysis. The horizontal axis shows the chromosome numbers - SNPs are plotted in order of their positions within chromosomes. The vertical axis represents the penalized regression coefficients (β_{sd}) for the SNP \times Disease interactions.

Table 4

Right thalamus-parameter estimates. The regression coefficients of interest for the SNP marker rs1256429 on the right thalamus region from the single-SNP model

Variable	β	SE(β)	t	p -value
SN P (rs1256429)	0.2469	0.066	3.738	0.0002
Disease	-0.0034	0.042	-0.082	0.9348
SN P \times Disease	-0.3415	0.055	-6.190	8.7×10^{-10}

are: APOE, BIN1, CLU, ABCA7, CR1, PICALM, MS4A6A, CD33, MS4A4E, and CD2AP. Although the current study did not identify any marker from the top 10 genes of the AlzGene, the study replicated two previous hits from the literature, namely ARPP-21 and DISC1 originally identified in our previous work [7] and [30], respectively (see full Hyperlasso results—Supplementary document). In a recent meta-

analysis of 74,046 subjects, Lambert et al. [29] identified 11 new susceptibility loci for AD: HLA-DRB5-DRB1, SORL1, PTK2B, SLC24A4/RIN3, ZCWPW1, CELF1, NME8, FERMT2, CASS4, INPP5D, and MEF2C. Two of the new genes, PTK2B (rs755951) and SLC24A4 (rs8019291), were identified in our study at the significance thresholds $p = 3.95 \times 10^{-6}$ and $p = 4.29 \times 10^{-6}$, respectively. These genes were associated with the cortical thicknesses of left frontal pole and left inferior temporal regions. Our current study identified relatively more markers as possible susceptibility loci for AD than our previous study. This is not surprising as this study is more extensive in terms of statistical analysis, data size as well as the number of imaging phenotypes and the coverage of the brain region considered.

Table 5

List of SNPs from the LD-based clump procedure, which were also identified as associated with AD in the previous GWAS literature. Previous GWAS hits in the LD-based clumped analysis.

Phenotype	SNP	CHR	p -value	Annotation
Right Middle Temporal	rs2063648	3	4.87×10^{-6}	ARPP-21
Right Inferior Parietal	rs6784162	3	2.89×10^{-5}	ARPP-21
Right Inferior Temporal	rs6784162	3	3.12×10^{-5}	ARPP-21
Right Caudal Anterior Cingulate	rs661319	10	2.53×10^{-5}	SORCS1
4th Ventricles	rs1074903	16	1.18×10^{-5}	FTO
5th Ventricles	rs2237198	6	2.47×10^{-5}	ATXN1
Left Paracentral	rs6458573	6	8.80×10^{-5}	CD2AP

DISCUSSION

Although HOMER2 was not reported to be linked with AD in the GWAS literature, neurobiological research [30] has linked HOMER2 with neurodegenerative process. This gene was associated with the volume of right thalamus, a brain region reported to be linked with AD in several studies [31–36]. The volume of right thalamus was found to be smaller in the diseased group with the difference being significantly greater for increased number of minor alleles in the associated SNP marker (rs1256429, HOMER2). One important pathologic feature of AD is the formation of neuritic plaques in the brains. Major components of these plaques are small peptides called amyloid- β (A β) that is derived from amyloid- β protein precursor (A β PP). Because of the critical role of A β in the pathogenesis of AD, understanding the cellular and molecular process underlying A β PP/A β metabolism has been of paramount importance to AD research. A recent study by Parisiadou and colleagues [30] provides evidence that expression of HOMER2 protein inhibits A β PP processing and reduces secretion of A β peptides. In addition, they decrease the levels of cell surface A β PP and inhibit maturation of A β PP and β -secretase. The effects of HOMER2 on A β PP trafficking to the cell surface could be part of the mechanism by which the expression of these proteins leads to the significant reduction of A β peptide production. A genetic mutation in HOMER2 that leads to under-expression of this protein is therefore likely to be related to AD.

EOMES (eomesodermin) is a protein-coding gene and functions as a transcriptional activator playing a crucial role during development. EOMES gene belongs to the TBR1 subfamily, which encodes *T-box brain protein-1*. This protein is believed to be involved in the differentiation and migration of neurons and plays important roles by regulating cortical development in human brain (Hsueh et al. [37]). Point mutations in the presenilin-1 gene (PS1) are thought to be causative in familial AD [38]. mRNA microarray studies (e.g., [39]) have demonstrated that PS1-deficiency is correlated with expression of EOMES and various other genes. This may support the finding from our Hyperlasso analysis that the EOMES is linked to AD.

Recent literature on AD research also linked JAM2 and WEE1 with AD. JAM2 belongs to the junctional adhesion molecule (JAM) family, and the immunoglobulin superfamily and is a protein-coding gene. The protein encoded by this gene is a type I

membrane protein that is localized at the tight junctions of both epithelial and endothelial cells. It acts as an adhesive ligand for interacting with a variety of immune cell types, and may play a role in lymphocyte homing to secondary lymphoid organs (RefSeq). A recent study (Guffanti et al. [40]) found associations between deletions of copy number variation regions and AD. The study suggested, via a functional cluster analysis annotation, that proteins containing immunoglobulin-like domains and genes coding for neural cell adhesion molecules (e.g., JAM2) may co-function in the biological network of genes putatively affected by deletions of copy number variation-regions associated with AD. Antonell and colleagues [41] performed neuropathological examination of patients with a progressive cognitive decline and observed a chromosomal 21 region duplication spanning 0.59 Mb, which comprised JAM2 and some other genes.

The WEE1 gene encodes a nuclear protein, which is a tyrosine kinase belonging to the Ser/Thr family of protein kinases. This protein catalyzes the inhibitory tyrosine phosphorylation of CDC2/cyclin B kinase, and appears to coordinate the transition between DNA replication and mitosis by protecting the nucleus from cytoplasmically activated CDC2 kinase (RefSeq, Jul 2008). It has been reported in the literature that WEE1 activity is altered in AD neurodegenerative process. Tomashevski and colleagues [42] reported WEE1 as a mitotic regulator that participates in the AD neurodegenerative process.

Hyperlasso analysis also detected another two markers (rs2063648; intronic to ARPP-21 and rs980989; flanking DISC1) showing significant association with AD at the genome-wide significance level. The gene ARPP-21 was reported to be associated with AD in our previous study [7] via another SNP (rs11129640) and the association between DISC1 and AD was reported previously (rs12044355, $p=9 \times 10^{-6}$) by Beecham and colleagues [43]. These two markers however did not pass the additional filtering criterion based on the penalized regression coefficient ($|\beta_{sd}| \geq 0.15$) that we used in this analysis. The estimated (penalized) SNP \times Disease interaction effect for these markers were $\beta_{sd} = -0.08$ and -0.03 , respectively.

CONCLUSIONS

We identified several novel markers associated with AD through our one-SNP-at-a-time and simultaneous multi-SNP Hyperlasso analysis, one of which was replicated in both analyses. The SNP rs1256429

($p = 8.7 \times 10^{-10}$, intronic to HOMER2) was found to be significantly associated with AD in one-SNP-at-a-time analysis. This SNP was also one of the top hits in the Hyperlasso analysis. The other SNPs that passed our stringent Hyperlasso penalty (to ensure statistical significance at the genome-wide level; $p = 1 \times 10^{-8}$) and the additional filtering threshold ($\beta_{sd} < -0.15$) of the penalized regression coefficients are: rs2724509 (flanking EOMES), rs2829841 (intronic to JAM2) and rs10770042 (in the coding region of WEE1).

A limitation of the current study is that, due to the unavailability of similar additional datasets, we were unable to replicate our findings in an independent validation study. An obvious future direction of research in this area would be to validate the findings in a replication study using an independent validation dataset. Collecting brain imaging and genetic data on big enough samples is challenging due to high cost of such studies. However, imaging-genetics studies may provide realistic opportunity in finding the unidentified genetic markers of AD in the near future when such studies become more affordable due to technological advances.

Nevertheless, the discussions of previous literature relevant to those genes in the previous section support our belief that the markers identified in this study are novel additions to the existing list of genetic variants associated with AD which can be validated in future replicated studies. We have also been able to replicate some earlier findings and have been able to suggest the genetic basis for some genes that have been previously identified as being linked to AD and cognitive performance such as the genes ARPP-21 [7, 44], JAM2 [44] and DISC1 [43].

ACKNOWLEDGMENTS

This study was supported by InnoMed, (Innovative Medicines in Europe) an Integrated Project funded by the European Union of the Sixth Framework program priority FP6-2004-LIFESCIHEALTH-5, Life Sciences, Genomics and Biotechnology for Health. Thanks to the Strategic Research Programme in Neuroscience at Karolinska Institutet (StratNeuro). The authors are also grateful for funding from the National Institute of Health Research (NIHR) Biomedical Research Centre for Mental Health at the South London and Maudsley NHS Foundation Trust and King's College London, Institute of Psychiatry, the National Institute of Health Research (NIHR) Biomedical Research Unit for Dementia at the South London

and Maudsley NHS Foundation Trust and King's College London, Institute of Psychiatry and Alzheimer's Research UK.

The collection and sharing of ADNI data 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; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics, N.V.; 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; 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 California Institute for Research and Education, and the study is coordinated by the 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.

Authors' disclosures available online (<http://j-alz.com/manuscript-disclosures/14-2214r1>).

SUPPLEMENTARY MATERIAL

The supplementary material is available in the electronic version of this article: <http://dx.doi.org/10.3233/JAD-142214>.

REFERENCES

- [1] Gatz M, Reynolds CA, Fratiglioni L, Johansson B, Mortimer JA, Berg S, Fiske A, Pedersen NL (2006) Role of genes and environments for explaining Alzheimer disease. *Arch Gen Psychiatry* **63**, 168-174.
- [2] Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML, Pahwa JS, Moskvina V, Dowzell K, Williams A, Jones N, Thomas C, Stretton A, Morgan AR, Lovestone S,

- Powell J, Proitsi P, Lupton MK, Brayne C, Rubinsztein DC, Gill M, Lawlor B, Lynch A, Morgan K, Brown KS, Passmore PA, Craig D, McGuinness B, Todd S, Holmes C, Mann D, Smith AD, Love S, Kehoe PG, Hardy J, Mead S, Fox N, Rossor M, Collinge J, Maier W, Jessen F, Schürmann B, Heun R, van den Bussche H, Heuser I, Kornhuber J, Wiltfang J, Dichgans M, Frölich L, Hampel H, Hüll M, Rujescu D, Goate AM, Kauwe JS, Cruchaga C, Nowotny P, Morris JC, Mayo K, Sleegers K, Bettens K, Engelborghs S, De Deyn PP, Van Broeckhoven C, Livingston G, Bass NJ, Gurling H, McQuillin A, Gwilliam R, Deloukas P, Al-Chalabi A, Shaw CE, Tsolaki M, Singleton AB, Guerreiro R, Mühlleisen TW, Nöthen MM, Moebus S, Jöckel KH, Klopp N, Wichmann HE, Carrasquillo MM, Pankratz VS, Younkin SG, Holmans PA, O'Donovan M, Owen MJ, Williams J (2009) Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat Genet* **41**, 1088-1093.
- [3] Hollingworth P, Harold D, Sims R, Gerrish A, Lambert JC, Carrasquillo MM, Abraham R, Hamshere ML, Pahwa JS, 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 MK, Brayne C, Rubinsztein DC, Gill M, Lawlor B, Lynch A, Brown KS, Passmore PA, Craig D, McGuinness B, Todd S, Holmes C, Mann D, Smith AD, Beaumont H, Warden D, Wilcock G, Love S, Kehoe PG, Hooper NM, Vardy ER, Hardy J, Mead S, Fox NC, Rossor M, Collinge J, Maier W, Jessen F, Rüther E, Schürmann B, Heun R, Kölsch H, van den Bussche H, Heuser I, Kornhuber J, Wiltfang J, Dichgans M, Frölich L, Hampel H, Gallacher J, Hüll M, Rujescu D, Giegling I, Goate AM, Kauwe JS, Cruchaga C, Nowotny P, Morris JC, Mayo K, Sleegers K, Bettens K, Engelborghs S, De Deyn PP, Van Broeckhoven C, Livingston G, Bass NJ, Gurling H, McQuillin A, Gwilliam R, Deloukas P, Al-Chalabi A, Shaw CE, Tsolaki M, Singleton AB, Guerreiro R, Mühlleisen TW, Nöthen MM, Moebus S, Jöckel KH, Klopp N, Wichmann HE, Pankratz VS, Sando SB, Asayl JO, Barcikowska M, Wszolek ZK, Dickson DW, Graff-Radford NR, Petersen RC, Alzheimer's Disease Neuroimaging Initiative, van Duijn CM, Breteler MM, Ikram MA, DeStefano AL, Fitzpatrick AL, Lopez O, Launer LJ, Seshadri S, consortium CHARGE, Berr C, Campion D, Epelbaum J, Dartigues JF, Tzourio C, Alpérovitch A, Lathrop M, EADIL consortium, Feulner TM, Friedrich P, Riehle C, Krawczak M, Schreiber S, Mayhaus M, Nicolaus S, Wagenpfeil S, Steinberg S, Stefansson H, Stefansson K, Snaedal J, Björnsson S, Jonsson PV, Chouraki V, Genier-Boley B, Hiltunen M, Soininen H, Combarros O, Zelenika D, Delepine M, Bullido MJ, 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, Bossù P, Piccardi P, Arosio B, Annoni G, Seripa D, Pilotto A, Scarpini E, Galimberti D, Brice A, Hannequin D, Licastro F, Jones L, Holmans PA, Jonsson T, Riemenschneider M, Morgan K, Younkin SG, Owen MJ, 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.
- [4] Potkin SG, Guffanti G, Lakatos A, Turner JA, Kruggel F, Fallon JH, Saykin AJ, Orro A, Lupoli S, Salvi E, Weiner M, Macciardi F, Alzheimer's Disease Neuroimaging, Initiative (2009) Hippocampal atrophy as a quantitative trait in a genome-wide association study identifying novel susceptibility genes for Alzheimer's disease. *PloS One* **4**, e6501.
- [5] Shen L, Kim S, Risacher SL, Nho K, Swaminathan S, West JD, Foroud T, Pankratz N, Moore JH, Sloan CD, Huettelman MJ, Craig DW, Dechairo BM, Potkin SG, Jack CR Jr, Weiner MW, Saykin AJ, Alzheimer's Disease Neuroimaging, Initiative (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.
- [6] Potkin SG, Turner JA, Guffanti G, Lakatos A, Torri F, Keator DB, Macciardi F (2009) Genome-wide strategies for discovering genetic influences on cognition and cognitive disorders: Methodological considerations. *Cogn Neuropsychiatry* **14**, 391-418.
- [7] Furney SJ, Simmons A, Breen G, Pedroso I, Lunnon K, Proitsi P, Hodges A, Powell J, Wahlund LO, Kloszewska I, Mecocci P, Soininen H, Tsolaki M, Vellas B, Spenger C, Lathrop M, Shen L, Kim S, Saykin AJ, Weiner MW, Lovestone S, Alzheimer's Disease Neuroimaging Initiative, AddNeuroMed, Consortium (2010) Genome-wide association with MRI atrophy measures as a quantitative trait locus for Alzheimer's disease. *Mol Psychiatry* **16**, 1130-1138.
- [8] Hoggart CJ, Whittaker JC, De Iorio M, Balding DJ (2008) Simultaneous analysis of all SNPs in genome-wide and resequencing association studies. *PLoS Genet* **4**, e1000130.
- [9] Mueller SG, Weiner MW, Thal LJ, Petersen RC, Jack CR, Jagust W, Trojanowski JQ, Toga AW, Beckett L (2005) Ways toward an early diagnosis in Alzheimer's disease: The Alzheimer's Disease Neuroimaging Initiative (ADNI). *Alzheimers Dement* **1**, 55-66.
- [10] Lovestone S, Francis P, Strandgaard K (2007) Biomarkers for disease modification trials: The innovative medicines initiative and AddNeuroMed. *J Nutr Health Aging* **11**, 359-361.
- [11] Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC (2007) PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-575.
- [12] Liu Y, Paajanen T, Zhang Y, Westman E, Wahlund LO, Simmons A, Tunnard C, Sobow T, Mecocci P, Tsolaki M, Vellas B, Muehlboeck S, Evans A, Spenger C, Lovestone S, Soininen H, AddNeuroMed, Consortium (2010) Analysis of regional MRI volumes and thicknesses as predictors of conversion from mild cognitive impairment to Alzheimer's disease. *Neurobiol Aging* **31**, 1375-1385.
- [13] Fischl B, Dale AM (2000) Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc Natl Acad Sci U S A* **97**, 11050-11055.
- [14] Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrave C, van der Kouwe A, Killiany R, Kennedy D, Klaveness S, Montillo A, Makris N, Rosen B, Dale AM (2002) Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron* **33**, 341-355.
- [15] Fischl B, Salat DH, van der Kouwe AJ, Makris N, Ségonne F, Quinn BT, Dale AM (2004) Sequence-independent segmentation of magnetic resonance images. *Neuroimage* **23**, S69-S84.
- [16] Westman E, Aguilar C, Muehlboeck J-S, Simmons A (2013) Regional magnetic resonance imaging measures for multivariate analysis in Alzheimer's disease and mild cognitive impairment. *Brain Topogr* **26**, 9-23.
- [17] Simmons A, Westman E, Muehlboeck S, Mecocci P, Vellas B, Tsolaki M, Kloszewska I, Wahlund LO, Soininen H, Lovestone S, Evans A, Spenger C, AddNeuroMed, Consortium (2009) MRI measures of Alzheimer's disease and the AddNeuroMed study. *Ann N Y Acad Sci* **1180**, 47-55.
- [18] Simmons A, Westman E, Muehlboeck S, Mecocci P, Vellas B, Tsolaki M, Kloszewska I, Wahlund LO, Soininen H,

- Lovestone S, Evans A, Spenger C (2011) The AddNeuroMed framework for multi-centre MRI assessment of Alzheimer's disease: Experience from the first 24 months. *Int J Geriatr Psychiatry* **26**, 75-82.
- [19] Tracy CA, Widom H (1994) Level-spacing distributions and the Airy kernel. *Commun Math Phys* **159**, 151-174.
- [20] Johnstone IM (2001) On the distribution of the largest eigenvalue in principal components analysis. *Ann Statist* **29**, 295-327.
- [21] Patterson N, Price AL, Reich D (2006) Population structure and eigenanalysis. *PLoS Genet* **2**, e190.
- [22] Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* **38**, 904-909.
- [23] de-Almada BV, de-Almeida LD, Camporez D, de-Moraes MV, Morelato RL, Perrone AM, Belcavello L, Louro ID, de-Paula F (2012) Protective effect of the APOE-e3 allele in Alzheimer's disease. *Braz J Med Biol Res* **45**, 8-12.
- [24] Berge G, Sando SB, Rongve A, Aarsland D, White LR (2014) Apolipoprotein E epsilon2 genotype delays onset of dementia with Lewy bodies in a Norwegian cohort. *J Neurol Neurosurg Psychiatry* **85**, 1227-1231.
- [25] Ayers KL, Cordell HJ (2010) SNP Selection in genome-wide and candidate gene studies via penalized logistic regression. *Genet Epidemiol* **34**, 879-891.
- [26] Nyholt DR (2004) A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am J Hum Genet* **74**, 765-769.
- [27] Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA (2009) Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A* **106**, 9362-9367.
- [28] Bertram L, McQueen M, Mullin K, Blacker D, Tanzi R. (2006) The AlzGene Database. Alzheimer Research Forum, <http://www.alzgene.org>.
- [29] Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, DeStafano AL, Bis JC, Beecham GW, Grenier-Boley B, Russo G, Thorton-Wells TA, Jones N, Smith AV, Chouraki V, Thomas C, Ikram MA, Zelenika D, Vardarajan BN, Kamatani Y, Lin CF, Gerrish A, Schmidt H, Kunkle B, Dunstan ML, Ruiz A, Bihoreau MT, Choi SH, Reitz C, Pasquier F, Cruchaga C, Craig D, Amin N, Berr C, Lopez OL, De Jager PL, Deramecourt V, Johnston JA, Evans D, Lovestone S, Letenneur L, Morón FJ, Rubinsztein DC, Eiriksdottir G, Sleegers K, Goate AM, Fiévet N, Huentelman MW, Gill M, Brown K, Kamboh MI, Keller L, Barberger-Gateau P, McGinness B, Larson EB, Green R, Myers AJ, Dufouil C, Todd S, Wallon D, Love S, Rogaeva E, Gallacher J, St George-Hyslop P, Clarimon J, Lleo A, Bayer A, Tsuang DW, Yu L, Tsolaki M, Bossù P, Spalletta G, Proitsi P, Collinge J, Sorbi S, Sanchez-Garcia F, Fox NC, Hardy J, Deniz Naranjo MC, Bosco P, Clarke R, Brayne C, Galimberti D, Mancuso M, Matthews F, European Alzheimer's Disease Initiative (EADI), Genetic, Environmental Risk in Alzheimer's Disease, Alzheimer's Disease Genetic Consortium, Cohorts for Heart, Aging Research in Genomic Epidemiology, Moebus S, Mecocci P, Del Zompo M, Maier W, Hampel H, Pilotto A, Bullido M, Panza F, Cafarria P, Naemias B, Gilbert JR, Mayhaus M, Lannefelt L, Hakonarson H, Pichler S, Carrasquillo MM, Ingesson M, Beekly D, Alvarez V, Zou F, Valladares O, Younkin SG, Coto E, Hamilton-Nelson KL, Gu W, Razquin C, Pastor P, Mateo I, Owen MJ, Faber KM, Jonsson PV, Combarros O, O'Donovan MC, Cantwell LB, Soininen H, Blacker D, Mead S, Mosley TH Jr, Bennett DA, Harris TB, Fratiglioni L, Holmes C, de Brujin RF, Passmore P, Montine TJ, Bettens K, Rotter JL, Brice A, Morgan K, Foroud TM, Kukull WA, Hannequin D, Powell JF, Nalls MA, Ritchie K, Lunetta KL, Kauwe JS, Boerwinkle E, Riemschneider M, Boada M, Hiltunen M, Martin ER, Schmidt R, Rujescu D, Wang LS, Dartigues JF, Mayeux R, Tzourio C, Hofman A, Nöthen MM, Graff C, Psaty BM, Jones L, Haines JL, Holmans PA, Lathrop M, Pericak-Vance MA, Launer LJ, Farrer LA, van Duijn CM, Van Broeckhoven C, Moskina V, Seshadri S, Williams J, Schellenberg GD, Amouyel P (2013) Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* **45**, 1452-1458.
- [30] Parisiadou L, Bethani I, Michaki V, Kroustis K, Rapti G, Efthimiopoulos S (2008) Homer2 and Homer3 interact with amyloid precursor protein and inhibit A β production. *Neurobiol Dis* **30**, 353-364.
- [31] Trollor JN, Sachdev PS, Haindl W, Brodaty H, Wen W, Walker BM (2006) A high-resolution single photon emission computed tomography study of verbal recognition memory in Alzheimer's disease. *Dement Geriatr Cogn Disord* **21**, 267-274.
- [32] Zhang HY, Wang SJ, Xing J, Liu B, Ma ZL, Yang M, Zhang ZJ, Teng GJ (2009) Detection of PCC functional connectivity characteristics in resting-state fMRI in mild Alzheimer's disease. *Behav Brain Res* **197**, 103-108.
- [33] Forster S, Vaitl A, Teipel SJ, Yakushev I, Mustafa M, la Fougere C, Rominger A, Cumming P, Bartenstein P, Hampel H, Hummel T, Buerger K, Hundt W, Steinbach S (2010) Functional representation of olfactory impairment in early Alzheimer's disease. *J Alzheimers Dis* **22**, 581-591.
- [34] Almeida OP, Garrido GJ, Alfonso H, Hulse G, Lautenschlager NT, Hankey GJ, Flicker L (2011) 24-month effect of smoking cessation on cognitive function and brain structure in later life. *Neuroimage* **55**, 1480-1489.
- [35] Dos Santos V, Thomann PA, Wustenberg T, Seidl U, Essig M, Schroder J (2011) Morphological cerebral correlates of CERAD test performance in mild cognitive impairment and Alzheimer's disease. *J Alzheimers Dis* **23**, 411-420.
- [36] Ota M, Sato N, Nakata Y, Arima K, Uno M (2012) Relationship between apathy and diffusion tensor imaging metrics of the brain in Alzheimer's disease. *Int J Geriatr Psychiatry* **27**, 722-726.
- [37] Hsueh Y-P, Wang T-F, Yang F-C, Sheng M (2000) Nuclear translocation and transcription regulation by the membrane-associated guanylate kinase CASK/LIN-2. *Nature* **404**, 298-302.
- [38] De Strooper B, Saftig P, Craessaerts K, Vanderstichele H, Guhde G, Annaert W, Von Figura K, Van Leuven F (1998) Deficiency of presenilin-1 inhibits the normal cleavage of amyloid precursor protein. *Nature* **391**, 387-390.
- [39] Mirnics ZK, Mirnics K, Terrano D, Lewis DA, Sisodia SS, Schor NF (2003) DNA microarray profiling of developing PS1-deficient mouse brain reveals complex and coregulated expression changes. *Mol Psychiatry* **8**, 863-878.
- [40] Guffanti G, Torri F, Rasmussen J, Clark AP, Lakatos A, Turner JA, Fallon JH, Saykin AJ, Weiner M, Vawter MP (2013) Increased CNV-Region deletions in mild cognitive impairment (MCI) and Alzheimer's disease (AD) subjects in the ADNI sample. *Genomics* **102**, 112-122.
- [41] Antonell A, Gelpí E, Sánchez-Valle R, Martínez R, Molinero JL, Lladó A (2012) Breakpoint sequence analysis of an A β PP locus duplication associated with autosomal dominant Alzheimer's disease and severe cerebral amyloid angiopathy. *J Alzheimers Dis* **28**, 303-308.

- [42] Tomashevski A, Husseman J, Jin L-W, Nochlin D, Vincent I (2001) Constitutive Wee1 activity in adult brain neurons with M phase-type alterations in Alzheimer neurodegeneration. *J Alzheimers Dis* **3**, 195-207.
- [43] Beecham GW, Martin ER, Li YJ, Slifer MA, Gilbert JR, Haines JL, Pericak-Vance MA (2009) Genome-wide association study implicates a chromosome 12 risk locus for late-onset Alzheimer disease. *Am J Hum Genet* **84**, 35-43.
- [44] Need AC, Attix DK, McEvoy JM, Cirulli ET, Linney KL, Hunt P, Ge D, Heinzen EL, Maia JM, Shianna KV, Weale ME, Cherkas LF, Clement G, Spector TD, Gibson G, Goldstein DB (2009) A genome-wide study of common SNPs and CNVs in cognitive performance in the CANTAB. *Hum Mol Genet* **18**, 4650-4661.