A large scale multivariate parallel ICA method reveals novel imaging–genetic relationships for Alzheimer's disease in the ADNI cohort ☆

Shashwath A. Meda [a,b,*,1], Balaji Narayanan [a,1], Jingyu Liu [c,d,1], Nora I. Perrone-Bizzozero [d,e,1], Michael C. Stevens [a,1], Vince D. Calhoun [c,d,1], David C. Glahn [a,g,h,1], Li Shen [f,1], Shannon L. Risacher [f,1], Andrew J. Saykin [f,1], Godfrey D. Pearlson [a,g,h,1]

[a] Olin Neuropsychiatric Research Center, Hartford Hospital/IOL, Hartford, CT 06106, USA
[b] Center for Human Genetics and Research, Vanderbilt University, Nashville, TN 37203, USA
[c] Department of ECE, University of New Mexico, Albuquerque, NM 87106, USA
[d] Mind Research Network, Albuquerque, NM 87106, USA
[e] Department of Neurosciences, University of New Mexico, Albuquerque, NM 87106, USA
[f] Department of Radiology and Imaging Sciences, Indiana University School of Medicine, Indianapolis, IN 46202, USA
[g] Department of Psychiatry, Yale University, New Haven, CT 06511, USA
[h] Department of Neurobiology, Yale University, New Haven, CT 06511, USA

ARTICLE INFO

ABSTRACT

The underlying genetic etiology of late onset Alzheimer's disease (LOAD) remains largely unknown, likely due to its polygenic architecture and a lack of sophisticated analytic methods to evaluate complex geno-type–phenotype models. The aim of the current study was to overcome these limitations in a bi-multivariate fashion by linking intermediate magnetic resonance imaging (MRI) phenotypes with a genome-wide sample of common single nucleotide polymorphism (SNP) variants. We compared associations between 94 different brain regions of interest derived from structural MRI scans and 533,872 genome-wide SNPs using a novel multivariate statistical procedure, parallel-independent component analysis, in a large, national multi-center subject cohort. The study included 209 elderly healthy controls, 367 subjects with amnestic mild cognitive impairment and 181 with mild, early-stage LOAD, all of them Caucasian adults, from the Alzheimer's Disease Neuroimaging Initiative cohort. Imaging was performed on comparable 1.5 T scanners at over 50 sites in the USA/Canada. Four primary "genetic components" were associated significantly with a single structural network including all regions involved neuropathologically in LOAD. Pathway analysis suggested that each component included several genes already known to contribute to LOAD risk (e.g. APOE4) or involved in pathologic processes contributing to the disorder, including inflammation, diabetes, obesity and cardiovascular disease. In addition significant novel genes identified included ZNF673, VPS13, SLC9A7, ATP5G2 and SHROOM2. Unlike conventional analyses, this multivariate approach identified distinct groups of genes that are plausibly linked in physiologic pathways, perhaps epistatically. Further, the study ex-emplifies the value of this novel approach to explore large-scale data sets involving high-dimensional gene and endophenotype data.

© 2012 Elsevier Inc. All rights reserved.

Introduction

Late onset Alzheimer's disease (LOAD), the commonest cause of late-life dementia (Bekris et al., 2010) has high heritability (Gatz et al., 2006a, 2006b). However, its etiopathology, pathogenesis and major risk genes are only partly known, mainly due to its genetic complexity and heterogeneity. The "amyloid hypothesis" seems insufficient to fully explain LOAD etiology and alternative hypotheses continue to be advanced (Pimplikar et al., 2010).

To date only one gene of major effect, apolipoprotein E ε4 (APOE4), replicates as significantly influencing LOAD risk (Strittmatter et al., 1993), but does not account for all genetic variability, suggesting the in-terplay of multiple, mostly unidentified susceptibility loci of smaller

effect size acting multiplicatively under a common disease variant model (Eccles and Tapper, 2010) and/or with environmental factors (Traynor and Singleton, 2010). Recent high-throughput genome wide association studies (GWAS) (Grupe et al., 2007; Harold et al., 2009; Seshadri et al., 2010; van Es and van den Berg, 2009) have identified and replicated in addition to APOE4, other genes such as BIN1, CLU, ABCA7, CR1, PICALM, MS4A6A, CD33, MSA4E and CD2AP, all of which (apart from APOE) have modest effect sizes and cumulatively account for only 35% of the population attributable risk (Ku et al., 2011; Naj et al., 2011). However, if LOAD risk is mediated in part by common polymorphisms individually conferring low disease risk, acting in concert, typical univariate GWAS might not have enough power to consistently detect these effects unless they utilize very large sample sizes. This might be an inherent problem as obtaining such large sample sizes are usually quite difficult. Also more importantly univariate studies do not take into account the effect of multiple genes at once. This is important because major LOAD risk factors include obesity, cerebrovascular disease and diabetes, all disorders with significant genetic underpinnings (Profenno et al., 2010), suggesting causative genes might belong to common biological pathways shared by these conditions. To circumvent some of these issues, multivariate analyses have been suggested as an approach to identify important genetic factors in LOAD (Gandhi and Wood, 2010).

MRI captures robust phenotypic neuroanatomical LOAD biomarkers, most consistently implicating posterior cingulate and entorhinal cortices, hippocampus and other medial temporal structures (Jack et al., 2010a, 2010b; Smith, 2010; Villain et al., 2010) corresponding to sites of early, severe LOAD-related neuropathology. Imaging genetics attempts to bridge genetic variations with phenotypic trait markers, relating genotypic variations to underlying biological disease etiologies and increasing statistical power, thereby requiring smaller sample sizes (Potkin et al., 2009). However, such strategies require tools to simultaneously accommodate thousands of data points per feature set (e.g. ~10^5 voxels from imaging data and up to 10^6 SNPs from genetic data), posing a major statistical challenge. Often, large scale studies are performed in a univariate fashion that significantly limits either one or both feature sets. However, these techniques can curtail the usefulness of multidimensional data to identify potentially informative relationships. Conventional voxel-wise analyses are computationally time consuming on a genome-wide scale and ineffectively capture cumulative effect spread over multiple genes. Prior analyses (Biffi et al., 2010; Potkin et al., 2009; Shen et al., 2010) on the multi-site MRI/genetic ADNI dataset used massively univariate approaches: GWAS, that confirmed the risk status of APOE4 and identified TOMM40 (Shen et al., 2010) and hypothesis-driven analyses using pre-selected known affected brain regions plus GWAS, that reinforced the status of promising individual genes of interest (Biffi et al., 2010). However, no analyses have evaluated the premise that genetic determinants are not randomly distributed among relevant biological pathways but instead grouped together among specific biological processes, nor have they detected predicted groups of common, interactive risk polymorphisms.

Parallel independent component analysis (Para-ICA) a novel multivariate data-driven, hypothesis-free statistical technique, extends ICA to analyze multiple modalities simultaneously (Calhoun et al., 2009). Para-ICA identifies simultaneously clusters of associated, likely interacting genes related to either: (a) functional brain networks, (b) related structural brain regions, or (c) physiologic processes e.g. EEG patterns or other potential endophenotypes and shows their relationships (Calhoun et al., 2009). Beginning with two modalities (here, SNP's and MRI-derived regional brain volume/thickness), we sought to discover underlying factors from both modalities and their connections. Similar to conventional ICA analyses, extracted structural MRI components are maximally independent within modality and loading coefficients represent variation among individuals. Networks or components extracted from genetic data are groups of interacting SNP loci, contributing with varying degrees to a genetic process affecting a downstream biological function, i.e. linear SNP combinations highly associated with related phenotypes. To date, this technique has been used mainly in schizophrenia and healthy controls to find genes responsible for brain structure and function using MRI and EEG patterns (Jagannathan et al., 2010; Liu et al., 2008; Meda et al., 2010). However, subject and SNP numbers in those studies were typically small.

Genetic and structural MRI data from the ADNI sample provide an ideal test bed to explore LOAD and to validate application of Para-ICA to larger datasets. The subject number (>800) and large genotypic dataset (>600,000 SNPs) allow for examination of feasibility of scaling up this technique where some valid results are published in this dataset from conventional, hypothesis-driven analyses (Biffi et al., 2010). Because many LOAD risk genes remain to be discovered, the technique can simultaneously be used to identify novel risk genes, as it identifies clusters of related, interacting SNPs.

We had the following goals: 1) to evaluate whether Para-ICA could be scaled up to deal with larger populations and many more SNPs than previously analyzed; 2) to identify new risk genes for LOAD and their corresponding endophenotypes and 3) to explore the different LOAD-mediating biological interactive pathways in which the identified risk genes might participate. We hypothesized that the method might identify previously unknown LOAD risk genes, as well as known candidate genes. We hypothesized that identified genes would group into LOAD-associated physiologic pathways and processes.

Materials and methods

We evaluated associations between two data modalities, structural MRI (sMRI), (regional brain volumes and cortical thicknesses), and genome-wide genotypic data (SNPs), to reveal multivariate relationships between structural brain regions and SNP's that differed between healthy controls, MCI and AD subjects.

Alzheimer's Disease Neuroimaging Initiative (ADNI) study

Data used in the preparation of this article were obtained from the ADNI database (adni.loni.ucla.edu). ADNI results from efforts of many co-investigators from a broad range of academic institutions and private corporations, with subjects recruited from over 50 sites across the U.S. and Canada. For up-to-date information, see www.adni-info.org. The Principal Investigator of this initiative is Michael W. Weiner, MD, VA Medical Center and University of California-San Francisco.

Study participants

Data derived from the ADNI database on 818 subjects included baseline 1.5 T MRI scans, Illumina SNP genotyping data, APOE genotype status and demographic information. We limited analyses to European-American ADNI subjects (classified initially into respective ethnic groups based on self-report and validated subsequently using genetic markers) to prevent confluence of population stratification effects on data, yielding a total of 209 HC (mean/SD age = 76.05/4.94; 113 males) with no past history of neurological or psychiatric disorder, 367 subjects with MCI (mean/SD age = 74.95/7.37; 239 males) and 181 subjects with clinically-assessed AD (mean/SD age = 75.57/7.48; 100 males) for analysis.

Single nucleotide polymorphism (SNP)-genotype

Sample collection and single nucleotide polymorphism (SNP) genotyping for more than 620,000 target SNPs across the whole genome was completed on all ADNI participants as described in Saykin et al. (2010) and Shen et al. (2010).

Quality control/pre-processing

Prior to Para-ICA, genotyped SNP's underwent two pre-processing stages. First, quality control parameters were employed to discard data unsuitable for further analysis. Samples (both subjects and SNP's) were checked for missing data and those with missing call rates > 5% were excluded. Remaining samples were imputed for missing values (< 1%) by replacing data with the corresponding major genotype. Following this, all uninformative SNP's (constant variance) were excluded. SNP's were then checked for minor allele frequency (MAF); rare SNP variants with MAFs < 0.01 were excluded. Highly correlated SNP's (r > 0.95) (in block sizes of 100 kb) were removed. Finally, SNPs (in controls only) were checked for Hardy–Weinberg equilibrium set at a threshold of $p < 1E{-}7$. QC'd SNP's (N = 533,872) were then carried over to the next processing stage. The above analyses were performed using custom scripts in Matlab 7.0 (www.mathworks.com).

The above pre-processed SNP's were subjected to a univariate GWAS type case–control association analysis to identify those differing significantly across the three diagnostic groups. This 1) effectively restricted the core analysis to disease-related genetic data in the current sample, 2) reduced potential noise from interacting genes with little or no relationship to the disease model, providing hypothesis-free data-driven "enrichment," and 3) improved accuracy and the linking coefficient of the Para-ICA algorithm (determined based on previous simulation results (Liu et al., 2008)). All SNP's (coded using an additive model) were entered into a mass univariate ANOVA design with diagnostic group as the independent factor, using Matlab 7.0. SNP's surviving a liberal $p < 0.05$ uncorrected threshold were then advanced to the Para-ICA multivariate association analyses. As noted, at this stage, no multiple comparison correction was performed. All significant SNP's with a $p < 0.05$ uncorrected threshold (N = 27,150) were carried forward to Para-ICA to determine genetic associations (including weak effects spread across multiple SNPs) with brain structures.

MRI structural imaging-phenotype

All subjects underwent a high-resolution 1.5T, 3D structural MRI scan (MPRAGE) as detailed in http://www.adni-info.org. We utilized recently published ADNI imaging data, analyzed in Freesurfer V4.1.0 (http://surfer.nmr.mgh.harvard.edu/fswiki) (Shen et al., 2010), thus, brain structure preparation and analysis methods are described only briefly. An automated Freesurfer Bayesian segmentation and parcellation routine extracted and labeled cortical and subcortical tissue classes (Shen et al., 2010), yielding target region volumes, cortical thicknesses and total intracranial volumes for pre-defined brain structure regions-of-interest (ROIs). Freesurfer values for two independently collected MPRAGE scans per subject were averaged to yield a single volume/cortical thickness value. Table 1 lists all imaging phenotypes (N = 94; bilateral volumes of interest and cortical thickness values). All values were normalized by Z-score transformation before entry into Para-ICA.

Genotype–phenotype associations (Para-ICA)

Para-ICA was implemented using the Fusion ICA Toolbox v2.0a; http://icatb.sourceforge.net in Matlab 7.0 to compute independent genetic/imaging networks and simultaneously identify and quantify association between the two modalities/features. This variant of ICA was designed for multimodality processing that extracts components using an entropy term based on information theory to maximize independence and enhances the interconnection by maximizing the linkage function in a joint estimation process (Calhoun et al., 2009; Liu et al., 2008). In addition, Para-ICA estimated loading parameters expressing the weight of the overall component for each subject. Overall correlation values between loading coefficients of the two sets of imaging

and genetic component(s) were calculated component-wise for the aggregate sample to identify significantly associated feature sets. Comprehensive mathematical details of the algorithm and methodology are provided in Liu et al. (2008). Data values from all three diagnostic groups were organized as a matrix of subjects by SNP's/imaging-ROI values. These genotype and phenotype data matrices were input to the Para-ICA algorithm (see diagram in Fig. 1). The number of independent estimated components for both SNP (12 components) and imaging data (8 components) was separately estimated using Akaike information criteria (AIC) (Calhoun et al., 2001). Resulting correlation values between the Para-ICA feature sets were appropriately corrected for multiple comparisons at this stage and a Bonferroni correction was applied based on $12 \times 8 = 96$ comparisons yielding a corrected p value threshold of $0.05/96 = 0.0005$. Once significant feature set associations were identified, all contributing SNPs/imaging ROIs across each significant feature/network/component were thresholded at a supra level $|Z| > 2.0$ to specifically identify dominant loadings for each individual network. SNP's or regions surpassing this threshold were deemed to be contributing significantly to the overall signal of the corresponding component/network. Subsequently, loading coefficients of significantly associated components were tested in a case–control fashion to test if they differed significantly among diagnostic groups.

Significant SNP's from each component were then batch queried against the dbSNP database (http://www.ncbi.nlm.nih.gov/projects/SNP/) to extract corresponding known gene information; genes from this query (derived for each component) were entered into the functional annotation tool, DAVID (http://david.abcc.ncifcrf.gov/) to identify enriched biological themes and visualize these genes on functional pathways e.g. KEGG and/or BioCarta. The Ingenuity pathway analysis tool (IPA; http://www.ingenuity.com/products/pathways_analysis.html) modeled and analyzed the complex biology and genetic interactions as canonical pathway models within the identified significant genetic network(s). SNPs associated with known genes were mapped to the Ingenuity Pathways Knowledgebase to delineate biological networks. Genes were also input to Funcassociate v2.0 (http://llama.mshri.on.ca/funcassociate/) to reveal significantly enriched functional attributes in each component compared against a gene ontology database. Finally, we performed standard chi-square association analyses on the top 10 Z-score-ranked genes from each Para-ICA-derived genetic network, to determine their relative association with the disease model.

Results

Initial data pre-processing with a univariate "GWAS like" analysis ($p < 0.05$ uncorrected) revealed N = 27,150 SNPs that differed significantly across groups. It confirmed SNPs from APOE (ε4; p = 6.6E-16; ε3; p = 3.6E-09) and TOMM40 (p = 7.25E-08) as the top three candidate genes, whose genotypes differed significantly across diagnostic groups, as identified in prior GWAS of the same parent dataset (Potkin et al., 2009).

Genotype–phenotype associations using Para-ICA

From the 12 SNP and 8 anatomic principal component networks identified by AIC, Para-ICA identified four different independent genetic networks significantly associated with a single structural network (following Bonferroni correction). These four networks were component/networks 1 (G1; consisting of 169 significant genes/332 SNPs), 2 (G2; 182 genes/377 SNPs), 3 (G3; 267 genes/482 SNPs) and 4 (G4; 169 genes/332 SNPs). G1 and G3 had significant loadings (Z > 2) from APOE (ε4). All four networks were significantly associated with only one structural brain network (S1) that encompassed 40 different unilateral regions (a combination of both volumes and cortical thickness surpassing a $Z = 2$ threshold). Key structural regions loading heavily in this component were entorhinal cortex and middle

Table 1
Details of all Freesurfer-derived regions of interest used as the imaging phenotype.

Rank order	Region	Abs Z-score for network S1
1	Left superior frontal**	2.82
2	Right superior frontal**	2.80
3	Left entorhinal cortex**	2.74
4	Left middle temporal**	2.72
5	Left rostral mid frontal**	2.65
6	Left caudal mid frontal**	2.64
7	Left inferior temporal**	2.58
8	Right middle temporal**	2.58
9	Right hippocampus volume**	2.57
10	Right entorhinal cortex**	2.56
11	Right rostral mid frontal**	2.54
12	Left hippocampus volume**	2.52
13	Left isthmus cingulate**	2.51
14	Right caudal mid frontal**	2.47
15	Right isthmus cingulate**	2.43
16	Right temporal pole**	2.39
17	Left temporal pole**	2.37
18	Right inferior parietal**	2.31
19	Right inferior temporal**	2.29
20	Left fusiform**	2.29
21	Left superior temporal**	2.28
22	Inferior lateral ventricle**	2.27
23	Left superior marginal**	2.24
24	Right superior marginal**	2.21
25	Left inferior parietal**	2.21
26	Right post cingulate**	2.17
27	Right precuneus**	2.16
28	Right superior temporal**	2.15
29	Right amygdala volume**	2.13
30	Right pars operculum**	2.11
31	Left precentral**	2.11
32	Left pars triangularis**	2.10
33	Left amygdala volume**	2.10
34	Left bank STS**	2.08
35	Right pars triangularis**	2.08
36	Left post cingulate**	2.07
37	Left lateral orbfrontal**	2.07
38	Left precuneus**	2.06
39	Right parahippocampus**	2.01
40	Right fusiform**	2.00
41	Right lateral orbfrontal	1.97
42	Left pars orbitalis	1.97
43	Right precentral	1.97
44	Left pars operculum	1.96
45	Left parahippocampus	1.95
46	Right bank STS	1.90
47	Left medial orbitofrontal	1.84
48	Right superior parietal	1.82
49	Right medial orbitofrontal	1.82
50	Frontal pole	1.79
51	Right ventricle DC	1.79
52	Left ventricle DC	1.78
53	Right paracentral	1.75
54	3rd ventricle	1.71
55	Left paracentral	1.68
56	Left superior parietal	1.68
57	Left frontal pole	1.61
58	Right pars orb	1.56
59	Right frontal pole	1.52
60	Right pallidum volume	1.47
61	Left post central	1.44
62	Left rostral ant cingulate	1.41
63	Right post central	1.38
64	Mid anterior corpus collosum	1.37
65	Left pallidum volume	1.32
66	Central corpus collosum	1.32
67	Right putamen volume	1.24
68	Left accumbens volume	1.20
69	Mid posterior corpus collosum	1.18
70	Right thalamus volume	1.18
71	Posterior corpus collosum	1.17
72	Left putamen volume	1.15
73	Left lingual	1.13
74	Left lateral occipital	1.11
75	Right lateral occipital	1.10

Table 1 (continued)

Rank order	Region	Abs Z-score for network S1
76	Right accumbens volume	1.06
77	Anterior corpus collosum	1.05
78	Right caudal ant cingulate	1.03
79	Right lingual	0.95
80	Left transverse temporal	0.84
81	Right rostral ant cingulate	0.75
82	Left caud ant cingulate	0.69
83	Left thalamus volume	0.68
84	Right corpus collosum	0.66
85	Left corpus collosum	0.56
86	Right cuneus	0.55
87	Right transverse temporal	0.51
88	Left Pericalcarine cortex	0.46
89	Left caudate volume	0.31
90	Right caudate volume	0.31
91	Left cuneus	0.30
92	Right cerebellar cortex	0.18
93	Right Pericalcarine cortex	0.08
94	Left cerebellar cortex	0.05

** Denotes regions with significant loadings on imaging network S1, ranked in order of Z-score loads. S1 was the only imaging network that was significantly associated with any (in this case all four) genetic networks.

temporal cortex thicknesses and amygdala and hippocampus volumes. A complete list of significant regions is highlighted with double asterisks in Table 1 and illustrated in Fig. 2. Significant genotype–phenotype correlations (between loading parameters of each modality or feature) were as follows: 1) G1–S1 (r = −0.53; p<0.0001) 2) G2–S1 (r = 0.32; p<0.0001) 3) G4–S1 (r = 0.24; p<0.0001) and G3–S1 (r = −0.14; p = 0.0001). Fig. 2 summarizes these data with the top 10 representative genes from each genetic network along with their corresponding biological functions. Fig. S1 shows correlation (scatter) plots for these associations. Testing loading coefficients of the above networks for between-group differences revealed that in addition to having significant associations they also significantly discriminated groups by baseline diagnosis (AD, MCI, or healthy control). Mean loading coefficients across each genetic/structural network are shown in Fig. S2.

Pathway analysis of genetic networks (components)

Ingenuity Pathway Analysis (IPA) software (http://www.ingenuity.com/) was used to detect, visualize, and explore relevant biological networks associated with each genetic component. Top networks for G1 were cellular assembly/organization, cell morphology and development. G2 was enriched with genes related to cardiovascular disease, neurologic disease and cardiac arteriopathy. Primary networks for G4 were cell cycle, cell death and inflammatory response. G3 had an over-representation of genes related to neurological and psychological disorders. All the above networks contained known Alzheimer's-related proteins in their pathway interactions. Top dynamic networks from each genetic component are illustrated in Fig. 3. Based on known gene functions, the top five IPA canonical pathways for each gene network (sorted in terms of genotype–phenotype linkage significance) were as follows:

G1: cAMP mediated signaling, sulfur metabolism, calcium signaling, vascular NO signaling and regulation of IL-2 expression in T lymphocytes. G2: neuro-protective role of THOP1 in Alzheimer's, NOS endothelial effects, Type 2 diabetes signaling, tyrosine metabolism, CYP450. G4: cAMP-mediated signaling, cardiac beta-adrenergic signaling, synaptic long-term potentiation, molecular cancer mechanisms, NOS endothelial effects. Significantly associated non-neurologic disorders were type 2 diabetes (N = 80), coronary artery disease (N = 71), Crohn's/inflammatory bowel disease (N = 65). G3: protein kinase A signaling, cardiac beta-adrenergic signaling, cAMP-mediated signaling, amino sugars metabolism, glycosaminoglycan degradation. Significantly associated non-neurologic disorders were

Subjects sMRI images   **Subjects SNPs**

PARALLEL ICA
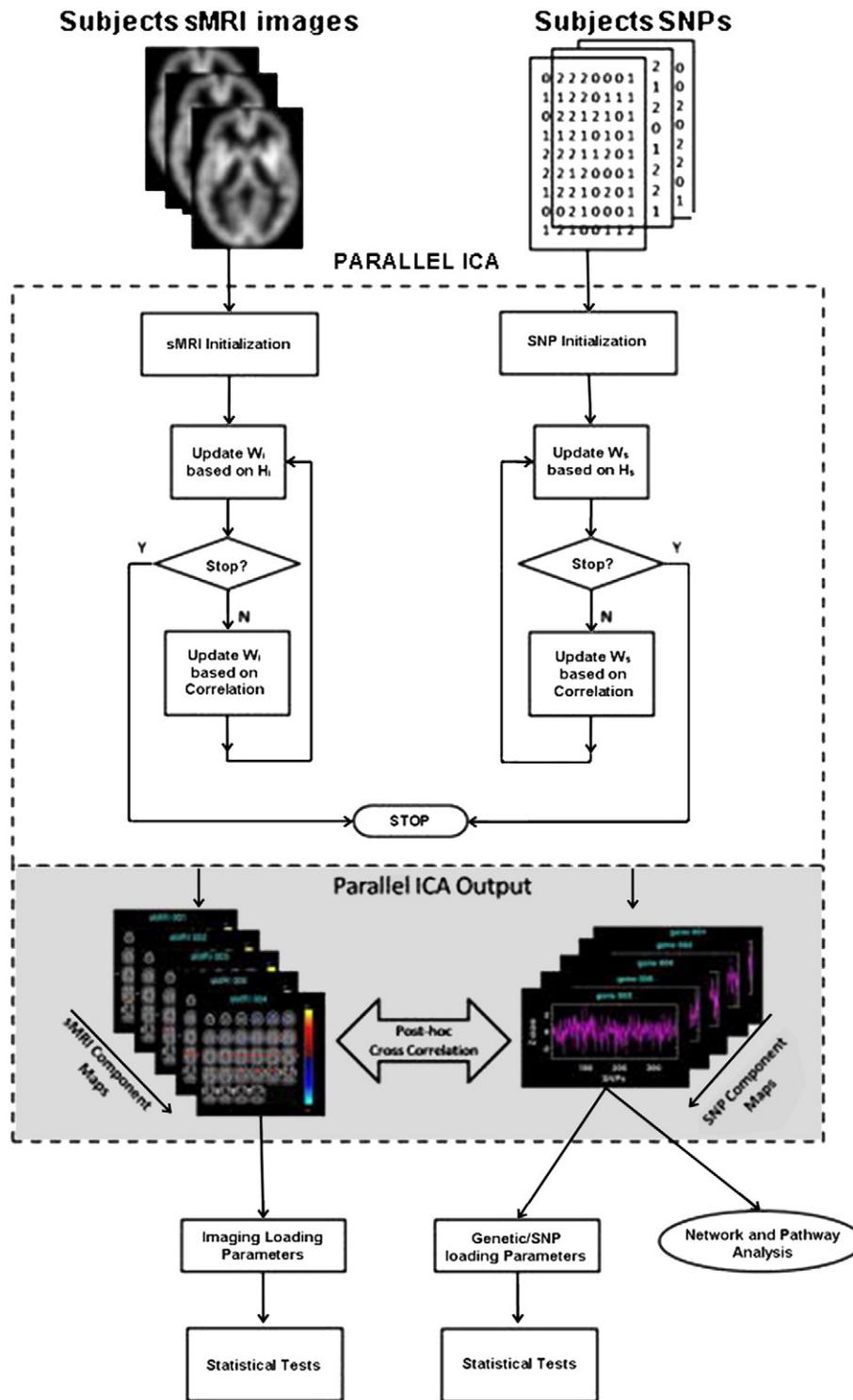
Fig. 1. Flow diagram illustrating the Para-ICA technique employed in this study.

Type 2 diabetes (N = 70), Crohn's disease/IBD (N = 57), coronary artery disease (N = 57).

The DAVID functional annotation tool revealed that the significant genes from all four genetic networks are involved in multiple biological pathways including Alzheimer's disease, adherens junction, arrhythmogenic right ventricular and dilated cardiomyopathy, axon guidance, calcium signaling, cell adhesion, ECM receptor interaction, focal adhesion and tight junction and smooth muscle contraction. Fig. 4 illustrates (marked with red stars) the significant genes in our study directly related to Alzheimer's disease on a KEGG pathway map derived from the

above tool. Genes analyzed using the Funcassociate v2.0 toolkit revealed several (N = 60) significantly over-represented attributes compared against the gene ontology database. The results presented in Table 2 are rank ordered based on adjusted p value along with the number of genes in the query, number of genes in the overall attribute and their odds ratios.

Supplementary association analysis of top genes from Para-ICA

Association analysis (to illustrate their relative disease association) of allelic frequencies for the top ten genes (based on ranked Z-scores)

COMPONENT 1 (169 significant genes/332 significant SNPs)

Gene	Z score (No of SNPS)	Function
SLC9A7/NHE7 (*G2)	7.53 (9)	Mediates electroneutral exchange of protons for Na(+) and K(+) across endomembranes. Localizes with actin. May contribute to Golgi volume & maintains function of trans-golgi network.Degrades internalized membrane proteins.
ZNF673 (*G2)	7.47 (2)	Zinc finger protein involved in transcriptional regulation. Duplication causes X linked MR.
SHROOM2	4.13 (1)	Myosin-actin binding protein involved in endothelial cell morphology changes during cell spreading
APOE (e4)	4.13 (1)	Mediates binding, internalization, and catabolism of lipoprotein particles.
SLC16A2	4 (5)	Specific thyroid hormone transporter. Stimulates cellular uptake of thyroxine (T4),triiodothyronine (T3), reverse triiodothyronine (rT3) and diiodothyronine.
COBL	3.96 (4)	Actin-related: organizes cytoskeleton during neurogenesis/morphogenesis.
MAGI3	3.83 (2)	Scaffolding protein at cell-cell junctions, regulating various cellular and signaling processes. Regulates beta-2 adrenergic receptor ER-K activation.
APOE (e3)	3.36 (1)	Mediates binding, internalization, and catabolism of lipoprotein particles.
CTNND2	3.34 (1)	Delta catenin/cadherin. Transcriptional activator when bound to ZBTB33 (By similarity). May be involved in neuronal cell adhesion, tissue morphogenesis & integrity by regulating adhesion molecules
PL-XDC2	3.26 (1)	Axon growth & guidance.

COMPONENT 2 (182 significant genes/377 significant SNPs)

Gene	Z score (No of SNPS)	Function
ZNF673 (*G1)	11.7 (2)	Zinc finger protein involved in transcriptional regulation. Duplication causes X linked MR-.
SL-C9A7 (*G1)	11.3 (8)	Mediates electroneutral exchange of protons for Na(+) and K(+) across endomembranes. Localizes with actin.May contribute to Golgi volume & maintains function of trans-golgi network. Degrades internalized membrane proteins.
CCHCR-1	8.95 (3)	Coiled coil alpha helical rod protein. Skin cell proliferation.
CFB (*G4)	7.9 (1)	Complement system Factor B, alternate pathway, .
TNXB	7.82 (3)	Mediates interactions between cells & extracellular matrix. Lipoprotein-related substrate-adhesion molecule; inhibits cell migration, accelerates collagen fibril formation.
EHMT2 (*G3,G4)	7.67 (1)	A methyltransferase.
ZKSCAN3 (ZNF263)	7.5 (1)	Zinc finger protein, transcriptional regulator. Activates transcription associated with colon cancer progression
C6orf15,PS OR-S1C1	7.43 (1)	HL-A gene.
SL-C44A4 (*G3,G4)	7.37 (1)	Na-dependent transmembrane transport protein involved in choline uptake by cholinergic neurons
ZSCAN12	7.35 (1)	Zinc finger protein involved in transcriptional regulation

r = -0.53 p<0.0001   r = 0.32 p<0.0001

Left Hemisphere - Lateral Aspect

Right Hemisphere - Medial Aspect

Subcortical Structures - Inferior Aspect

r = -0.14 p=0.001   r = 0.24 p<0.0001

COMPONENT 3 (242 significant genes/474 significant SNPs)

Gene	Z score (No of SNPs)	Function
CEP57	5.68 (4)	Centrosomal protein required for microtubule attachment to centrosomes and microtrafficing.
MTMR-2	5.4 (4)	Myotubularin related protein 2. Phosphatase acting on lipids with a phosphoinositol headgroup
ATF7	5.04 (4)	Transcription factor; regulates immunity via MAP kinase
VPS13C	4.54 (15)	Vacuolar sorting protein, ? Related to BACE.
SL-C44A4 (*G2,G4)	4.1 (1)	Na-dependent transmembrane transport protein involved in the uptake of choline by cholinergic neurons
EHMT2 (*G4,G2)	3.68 (1)	A methyltransferase.
PTCHD2	3.63 (1)	Patched domain- regulates Hedgehog.
CNTN5	3.6 (1)	A contactin; mediates cell surface interactions during nervous system development.
C2 (*G4)	3.57 (1)	Complement component C2, part of the classical complement pathway.
ATP5G2	3.53 (1)	Subunit of mitochondrial ATP synthase.

COMPONENT 4 (267 significant genes/482 significant SNPs)

Gene	Z score (No of SNPs)	Function
SL-C44A4 (*G2,G3)	6.46 (5)	Na-dependent transmembrane transport protein involved in choline uptake by cholinergic neurons.
CFB (*G2)	6.11 (1)	Complement Factor B, part of the alternate pathway of the complement system.
C2 (*G3)	6.03 (1)	Complement component C2, part of classical complement system. Implicated in cellular processes involving alteration of RNA secondary structure e.g. translation initiation, nuclear and mitochondrial splicing, ribosome & spliceosome assembly.
SKIV2L-	5.69 (4)	Gene close to Complement C2.
EHMT2 (*G2,G3)	5.67 (3)	A methyltransferase.
C6orf10	5.44 (1)	Unknown
TNXB	5.12 (3)	Mediates interactions between cells & extracellular matrix. Lipoprotein-related substrate-adhesion molecule. Inhibits cell migration, accelerates collagen fibril formation; ?epithelial role.
ATF6B	4.59 (2)	Transcriptional factor acting in the unfolded protein response (UPR-) pathway by activating UPR- target genes induced during ER- stress
CYP51A1	4.26 (2)	Cytochrome P450; monooxygenase catalyzing many reactions in drug metabolism, synthesis of cholesterol, steroids & other lipids
AKAP9	4.25 (8)	Scaffolding protein assembling several protein kinases & phosphatases on centrosome and Golgi apparatus. May maintain integrity of the Golgi apparatus.

Fig. 2. Significant genotype–phenotype associations identified using the hypothesis-free Para-ICA approach. Genetic factors combine effects from all risk loci, thus explaining more phenotypic variability than traditional analytic approaches. Figure shows the top ten genes, ranked by Z-scores, plus their known functions, from each of the four genetic networks that were significantly correlated with the MRI network. Numbers of significant genes and SNPs within each network is listed at top. *Indicates that the same gene appears among the top ten genes of another network. Genes in red text have been previously implicated in LOAD risk.
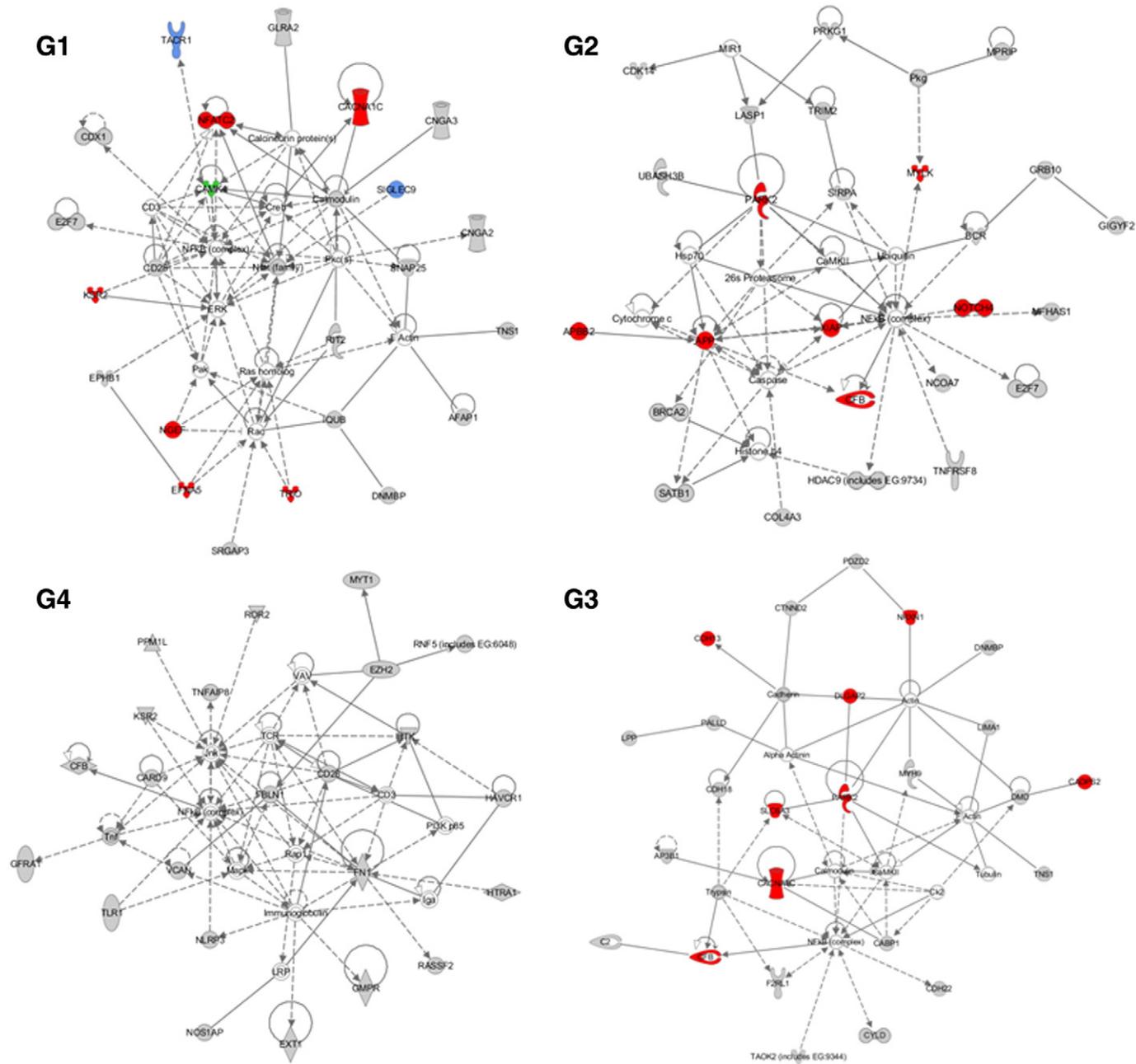
Fig. 3. Representation of the top canonical pathways derived from Ingenuity pathway software for all four significant genetic networks. G1: pathway enriched for cellular assembly/organization, cell morphology/development; G2: pathway enriched for cardiac arteriopathy, cardiovascular and neurological disease; G4: pathway enriched for cell cycle, cell death and inflammatory response; G3: enriched for genetic, neurological and psychological disorders. Key: red = known AD implicated gene; blue = involved in long-term memory, green = involved in long-term potentiation of CA1 neurons.
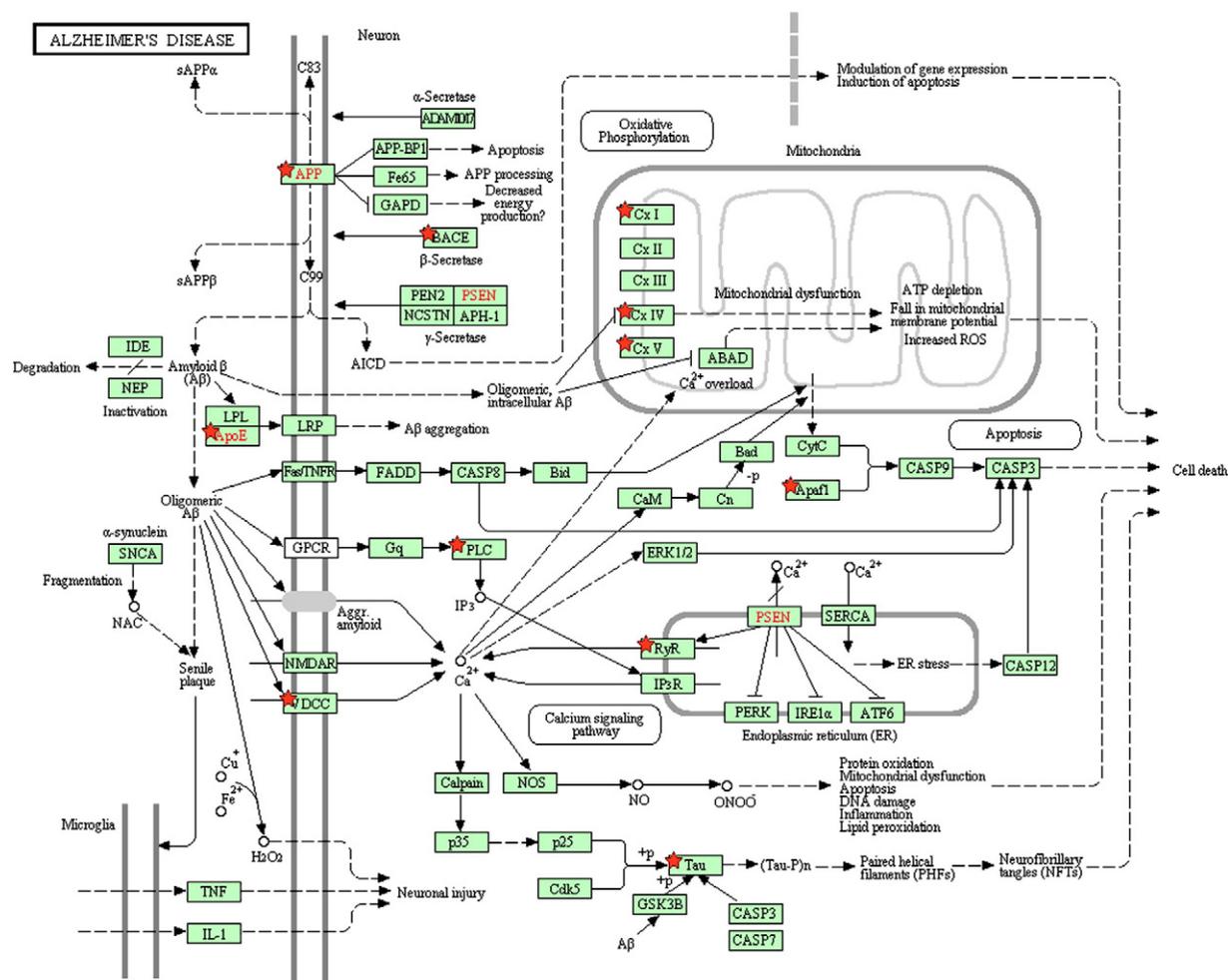
Fig. 4. Genes identified as part of the significant Para-ICA networks highlighted on a KEGG Alzheimer's pathway diagram derived from the DAVID functional association tool. Genes identified using red stars occurred in 1 or more of our 4 gene networks. Genes indicated in red text are the primary genes identified to date responsible for the early form of AD (PSEN, APP) or account for a major proportion of LOAD risk (APOE).

Table 2
Results derived from the Funcassociate web tool showing significantly enriched functional attributes, queried against the gene ontology database. Genes fed into the analysis comprised all significant genes from all four genetic networks identified in the study. Key: rank: position in the attribute list ranked by significance of association with the query; N: number of genes that went into the query with this attribute, X: number of genes overall in the ontology database with this attribute, LOD: logarithm (base 10) of the odds ratio, p-val: single hypothesis p-value of the association between attribute and query (based on Fisher's exact test), p-adj: adjusted p-value; fraction of 1000 null-hypothesis simulations having attributes with this single hypothesis p-value or smaller.

Rank	N	X	LOD	P	P-adj	Gene ontology attribute
1	20	100	1.847	3.50E-28	<0.001	Antigen processing and presentation of peptide antigen via MHC class I
2	20	103	1.831	6.70E-28	<0.001	Antigen processing and presentation of peptide antigen
3	23	206	1.56	3.70E-26	<0.001	MHC class I protein complex
4	23	335	1.327	2.90E-21	<0.001	MHC protein complex
5	23	377	1.271	4.10E-20	<0.001	Antigen processing and presentation
6	121	2180	0.409	6.60E-17	<0.001	Plasma membrane part
7	28	1152	0.864	5.60E-14	<0.001	Immune response
8	28	1348	0.791	2.50E-12	<0.001	Immune system process
9	160	3765	0.289	5.90E-12	<0.001	Plasma membrane
10	298	9033	0.234	1.10E-11	<0.001	Membrane
11	254	7427	0.237	3.40E-11	<0.001	Membrane part
12	7	33	1.841	8.80E-11	<0.001	MHC class I receptor activity
13	9	150	1.364	1.30E-09	<0.001	Thiolester hydrolase activity
14	8	115	1.43	3.40E-09	<0.001	Ubiquitin thiolesterase activity
15	8	118	1.418	4.20E-09	<0.001	Ubiquitin-specific protease activity
16	8	119	1.414	4.50E-09	<0.001	Small conjugating protein-specific protease activity
17	19	171	0.776	5.90E-09	<0.001	Rho protein signal transduction
18	100	6558	0.316	3.30E-08	0.002	Intrinsic to membrane
19	16	131	0.802	4.50E-08	0.002	Rho guanyl-nucleotide exchange factor activity
20	275	8319	0.183	8.10E-08	0.003	Protein binding
21	35	558	0.475	1.10E-07	0.003	Cytoskeletal protein binding
22	15	140	0.802	1.10E-07	0.003	Regulation of Rho protein signal transduction
23	11	192	0.979	1.10E-07	0.003	ATPase activity, coupled to transmembrane movement of substances
24	11	193	0.976	1.20E-07	0.003	ATPase activity, coupled to movement of substances
25	521	19305	0.204	1.30E-07	0.003	Cell part
26	521	19306	0.204	1.30E-07	0.003	Cell
27	11	197	0.967	1.40E-07	0.003	Hydrolase activity, acting on acid anhydrides, catalyzing transmembrane movement of substances
28	15	143	0.792	1.50E-07	0.003	Ras guanyl-nucleotide exchange factor activity
29	31	2607	0.533	1.50E-07	0.003	Protein complex
30	50	909	0.38	1.80E-07	0.003	Cell adhesion
31	50	909	0.38	1.80E-07	0.003	Biological adhesion
32	6	15	1.551	2.00E-07	0.003	Prostaglandin E receptor activity
33	64	1301	0.328	2.40E-07	0.003	Intrinsic to plasma membrane
34	96	6449	0.297	2.50E-07	0.004	Integral to membrane
35	11	209	0.939	2.60E-07	0.004	Primary active transmembrane transporter activity
36	11	209	0.939	2.60E-07	0.004	P–P-bond-hydrolysis-driven transmembrane transporter activity
37	63	1285	0.326	3.30E-07	0.005	Integral to plasma membrane
38	3	6	2.561	4.30E-07	0.007	Sulfonylurea receptor activity
39	27	390	0.511	6.50E-07	0.012	Proteinaceous extracellular matrix
40	6	15	1.457	6.90E-07	0.012	Cell adhesion molecule binding
41	32	1350	0.468	6.90E-07	0.012	Transmembrane transporter activity
42	17	577	0.677	7.00E-07	0.012	Active transmembrane transporter activity
43	3	4	2.603	7.90E-07	0.017	cGMP-dependent protein kinase activity
44	58	1262	0.329	7.90E-07	0.017	Calcium ion binding
45	19	258	0.622	9.10E-07	0.018	Guanyl-nucleotide exchange factor activity
46	6	19	1.398	1.00E-06	0.019	Prostaglandin receptor activity
47	27	400	0.499	1.10E-06	0.019	Extracellular matrix
48	41	3215	0.416	1.10E-06	0.019	Macromolecular complex
49	34	589	0.433	1.20E-06	0.022	Metal ion transport
50	8	246	1.082	1.20E-06	0.022	Cysteine-type peptidase activity
51	136	3975	0.212	1.30E-06	0.023	Localization
52	14	409	0.733	1.60E-06	0.027	ATPase activity, coupled
53	45	1158	0.366	1.60E-06	0.029	Ion transmembrane transporter activity
54	8	258	1.06	1.80E-06	0.029	Ubiquitin-dependent protein catabolic process
55	8	259	1.059	1.80E-06	0.029	Modification-dependent protein catabolic process
56	8	259	1.059	1.80E-06	0.029	Modification-dependent macromolecule catabolic process
57	8	260	1.057	1.90E-06	0.029	Proteolysis involved in cellular protein catabolic process
58	8	262	1.053	2.00E-06	0.032	Cellular protein catabolic process
59	15	152	0.688	2.50E-06	0.046	Calcium ion transport
60	8	42	1.057	2.60E-06	0.047	Voltage-gated calcium channel activity

from the four genetic components revealed that the genes SLC9A7, SHROOM2, ZNF673, APOE (ε3, ε4), VPS13C and ATP5G2 showed stronger effects of disease associations compared to other genes within the component. Table 3 details the allelic frequency and the associated statistical value for each of the top 10 genes from all four genetic networks; due to partial overlap (1 gene could appear in more than 1 component) totaling 32 unique genes. A weighting score was derived by normalizing

the chi-square value of each gene to the chi-square of the gene most associated with clinical disease status within each component.

Discussion

As hypothesized, we validated a scaled-up Para-ICA approach to reveal novel interactive genes and pathways for LOAD, thus

Table 3

Results of association analysis performed on the top ten genes from each genetic network sorted in descending order of their weighting scores. Weighting scores represent the relative effect of each SNPs association with disease status within a particular component, ranging from 0 to 1. A and B refer to the different allelic frequencies.

Gene	Location	Function	rs #	AD (# alleles) A	B	MCI (# alleles) A	B	HC (# alleles) A	B	Chi-Square	P-value	Weighting score
G1												
APOE3	19q13.2	N/A	E3	162	200	269	465	39	279	88	1.00E-16	1.000
APOE4	19q13.2	N/A	E4	209	153	521	273	358	60	80	1.00E-15	0.909
SHROOM2	Xp22.3	Intron	rs2405940	209	153	503	231	294	124	16.58	0.0003**	0.188
ZNF673	Xp11.3	Intron	rs7876304	103	259	250	484	172	246	14.04	0.0009**	0.160
SLC9A7	Xp11.23	Intron	rs1883255	99	263	241	493	165	253	13	0.0015**	0.148
SLC16A2	Xq13.2	Intron	rs479640	89	273	252	482	138	280	11.15	0.0038	0.127
MAGI3	1p13.2	Intron	rs7552954	154	208	373	361	226	192	10.97	0.0041	0.125
COBL	7p12.2	Intron	rs1437490	195	167	374	360	255	163	10.9	0.0043	0.124
PLXDC2	10p12.31	Intron	rs2358839	102	260	265	469	150	268	7.54	0.0229	0.086
CTNND2	5p15.2	Intron	rs757459	215	147	388	346	210	208	6.93	0.0311	0.079
G2												
ZNF673	Xp11.3	Intron	rs7876304	103	259	250	484	172	246	14.04	0.0009**	1.000
SLC9A7	Xp11.3	Intron	rs1883255	99	263	241	493	165	253	13	0.0015**	0.926
TNXB	6p21.3	Intron	rs2239689	133	229	213	521	114	304	9.47	0.0088	0.675
ZSCAN12	6p22.1	Mis-sense	rs1361385	260	102	497	237	258	160	9.24	0.0098	0.658
EHMT2	6p21.3	Intron	rs2844457	146	216	236	498	131	287	8.9	0.0116	0.634
ZKSCAN3	6p22.1	Intron	rs213228	258	104	491	243	258	160	8.03	0.018	0.572
SLC44A4	6p21.3	Intron	rs2736428	139	223	227	507	127	291	7.41	0.0246	0.528
CFB	6p21.3	Intron	rs4151657	223	139	509	225	286	132	6.96	0.0307	0.496
C6orf15	6p21.3	Intron	rs3130977	259	103	504	230	264	154	6.7	0.0349	0.477
CCHCR1	6p21.3	Intron	rs746647	256	106	490	244	261	157	6	0.0496	0.427
G3												
ATP5G2	12q13.13	Intron	rs1800634	111	251	166	568	134	284	14.97	0.0006**	1.000
VPS13C	15q22.2	Intron	rs1981916	190	172	449	285	216	202	12.84	0.0016**	0.858
ATF7	12q13.13	Intron	rs784568	189	173	337	397	230	188	9.82	0.0073	0.656
CEP57	11q21	Intron	rs3017756	154	208	316	418	215	203	8.95	0.0113	0.598
EHMT2	6p21.3	Intron	rs2844457	146	216	236	498	131	287	8.9	0.0116	0.595
MTMR2	11q21	Intron	rS193364	225	137	437	297	219	199	8.66	0.0132	0.578
PTCHD2	1p36.22	Intron	rs2379135	160	202	259	475	165	253	8.32	0.0156	0.556
CNTN5	11q22.1	Intron	rs10892901	134	228	314	420	146	272	7.86	0.0196	0.525
SLC44A4	6p21.32	Intron	rs2736428	139	223	227	507	127	291	7.41	0.0246	0.495
C2	6p21.3	Intron	rs2734335	204	158	355	379	211	207	6.22	0.0445	0.415
G4												
SKIV2L	6p21.3	Intron	rs592229	227	135	388	346	239	179	9.69	0.0078	0.999
TNXB	6P21.3	Intron	rs2239689	133	229	213	521	114	304	9.47	0.0088	0.976
EHMT2	6p21.3	Intron	rS159445	268	94	486	248	277	141	7.71	0.0211	0.795
AKAP9	7q21.2	Intron	rs756647	234	128	461	273	234	184	7.39	0.0248	0.762
SLC44A4	6p21.3	Intron	rs2242665	227	135	397	337	238	180	7.34	0.0254	0.757
CFB	6p21.3	Intron	rs4151657	223	139	509	225	286	132	6.96	0.0307	0.718
ATF6B	6p12.3	Intron	rs2228628	242	120	526	208	315	103	6.9	0.0316	0.711
C6orf10	6p21.3	Intron	rs485774	203	159	455	279	272	146	6.81	0.0331	0.702
CYP51A1	7q21.2	Intron	rs2301559	233	129	454	280	234	184	6.34	0.0419	0.654
C2	6p21.3	Intron	rs2734335	204	158	355	379	211	207	6.22	0.0445	0.641

** Genes that significantly discriminate groups after correction for multiple comparisons (Bonferroni-corrected across 40 chi-square tests).

highlighting one of the primary advantages of Para-ICA which is the use of modest sample sizes compared to conventional GWAS analyses to effectively capture genotype–phenotype relationships. Dominant loading coefficients were contained in all major regions affected by LOAD pathology in the single structural component significantly associated with four different SNP/genetic networks. Seven other structural components were unassociated with other gene components. The genetic components identified included SNPs from APOE4 plus multiple other risk genes (and putatively protective SNPs, e.g. APOE2) either previously identified in LOAD risk (e.g. ATF7 in G3; Lin et al., 2006) or involved in one or more biological processes thought to contribute to LOAD pathology. Fig. 5 summarizes the involvement of these four genetic risk networks on a LOAD physiologic pathway diagram.

The most significant association was between G1 and S1. This genetic network had significant loading contributions from a total of 169 different genes (332 SNPs) and correlated negatively with brain network S1, implying increased genetic load is related to decreased brain volume/thickness within the network. This association was notable as the S1 had high loadings from APOE4 (in the top 10 gene loadings) and S1 included regions known to be affected early and severely in LOAD, including entorhinal, middle temporal and prefrontal cortices and hippocampus. More importantly, this genetic network had high loadings from several other genes (SLC9A7/NHE7, ZNF673, SHROOM2) previously unidentified in LOAD pathology. Given that they were part of the same independent genetic component as APOE, this finding both confirms APOE's established role as an important LOAD risk gene and suggests that these additional SNPs may interact with APOE to influence disease risk, supporting APOE's role as a LOAD risk factor rather than a direct cause (Guerreiro et al., 2010). The protein encoded by SLC9A7 mediates $Na+/H+$ exchange across cell surface plasma membranes (Kagami et al., 2008) cycling between the cell surface and intracellular trans-Golgi network and recycling endosomes, which are vital to APP processing (Marks and Berg, 2010). SLC9A7 co-localizes with actin, implicated in tau formation (Gallo, 2007). LOAD lymphoblasts show abnormalities modulated by sodium/hydrogen exchanger blockers (Urcelay et al., 2001). Overall, this genetic network was enriched with genes dominant in cell
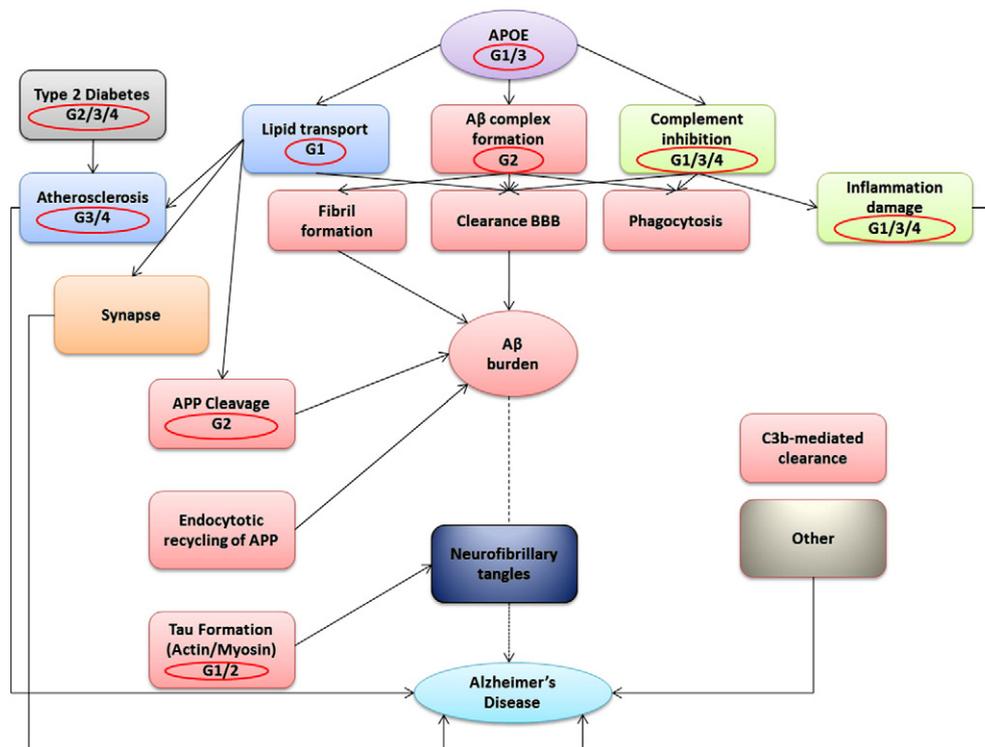
Fig. 5. A summary of different genetic components (circled in red) identified in the current study projected onto an Alzheimer's disease functional interaction pathway (modified from Sleegers K et al., TIGS 2009).

signaling pathways. Other strongly contributing genes are involved in lipid transport and tau formation (via actin/myosin binding). ZNF673 is associated with X-linked mental retardation, (Lugtenberg et al., 2006; Ramaswamy et al., 2010) and is close (~0.2 MB) to SCL9A7 on Xp11.3.

The second most significant association in genotype–phenotype correlation was G2–S1. This correlation was positive. G2 comprised 182 unique genes (377 SNPs). Some top-ranked genes from this component overlapped with those from G1, including ZNF673 and SLC9A7. These genes had a significant differential distribution among diagnostic groups, suggesting a role of actin localization and transcriptional regulation in LOAD. Other top genes from this network involved in important AD-related processes included the complement system, involved in amyloid-beta formation and inflammatory damage (van Es and van den Berg, 2009).

Network G4 correlated positively with S1. G4 contained several genes associated with risk for non-neurologic disorders, including diabetes and cardiovascular disease, both LOAD risk factors (Profenno et al., 2010). Top genes from this network, previously unidentified in the context of LOAD, belonged to the complement factor/inhibition pathway related to amyloid-beta clearance (35) or are associated with major histocompatibility class III. Additionally, AKAP9, a top 10 gene in this network, maintains neuronal Golgi integrity and is involved in LOAD pathogenesis (Stieber et al., 1996). Regarding association analysis, no top 10 gene from this network was significantly differentially distributed in the disease groups, suggesting that G4 comprises multiple SNPs of low effect acting together through diverse biological risk pathways, especially inflammation, (see Eikelenboom et al., 2006) to significantly affect LOAD-related neuropathology.

The final genotype–phenotype association was a negative correlation between G3 and S1. G3 included ATP5G2, a subunit of mitochondrial ATP-synthase, which was over-represented in the disease group. Mitochondrial ATP-synthase in entorhinal cortex is a target of oxidative stress in LOAD (Terni et al., 2010) and part of LOAD apoptosis

pathways. Several other G3 genes included dominant signaling from CNTN5, recently associated with multiple AD MRI characteristics (Biffi et al., 2010), CEP57, a microtubular/centrosomal localizer (Meunier et al., 2009), MTMR2, an endosomal regulator (Lee et al., 2010), and ATF7, associated with LOAD in Lin et al. (2006). The loading of previously identified LOAD genes and associated pathobiological pathways further supports the relevance of this genetic network.

Analyzing significant genes from all four components using DAVID and visualizing related processes on KEGG pathways revealed that genes grouped in multiple LOAD-relevant biological processes (see Fig. 4). Additional prominent processes not shown in the figure included cellular communication, cardiovascular diseases, signal transduction, calcium signaling, cell adhesion and neuronal developmental processes (e.g. axon guidance). Many such processes are implicated in LOAD pathology (e.g. neuronal calcium signaling; Kostiuk et al., 2010; LaFerla, 2002; Mattson and Chan, 2003). Semaphorin 3A, an axon-guiding membrane protein, accumulates in hippocampus in AD (Koncina et al., 2007).

Major themes deriving from the top 32 Z score-defined genes in the 4 SNP components suggest several major pathophysiological LOAD pathways, especially when such genes co-occurred within a component. From G1, APOE may relate to LOAD risk through pathways not directly linked to amyloid-beta, including actin-related mechanisms. Actin cytoskeletal changes as a path to tau formation (Gallo, 2007) are implicated across all components by SCLC987/ NHE7 (Kagami et al., 2008; Ohgaki et al., 2008), SHROOM2 and COBL (Dominguez, 2009) and microtubule-related genes including MTMR2, CEP57 and CTNND2 (Bamburg and Bloom, 2009; Meunier et al., 2009). Three such genes were present in component 1. Immune function, especially the complement system, related to amyloid-beta clearance (Guerreiro et al., 2010; Kolev et al., 2009) and expressed in cerebrovascular smooth muscle (Walker et al., 2008), is suggested by ATF7, CFB, C2, SKIV2L, C6orf10 and C6orf15 (Li et al., 2006; Veerhuis, 2011). These genes support the known role of the

complement system in LOAD pathogenesis (van Es and van den Berg, 2009), while adding new gene candidates, e.g. C2. Complement is present in dystrophic LOAD neurites, involved in immune response and linked to synaptic pruning (Hollingworth et al., 2011). Five immune related/complement genes are present in G4.

CTNND2/Delta Catenin/NPRAP is associated with GSK3-beta, hence BAP and tau (Bareiss et al., 2010). CNTN5 encodes contactin5; other contactins participate in LOAD risk (Biffi et al., 2010; Osterfield et al., 2008). The prominence of SCLC987/NHE7(and MTMR2) suggests the importance of the trans-Golgi network and recycling endosome (Lee et al., 2010). Endosomal processing of APP involving SorLA is of importance in LOAD (Lin et al., 2005; Marks and Berg, 2010; Ohgaki et al., 2008). VPS proteins are related to this process (He et al., 2005; Marks and Berg, 2010), although VPS13C has yet to be implicated. VPS13C is associated with maintenance of plasma glucose levels (Saxena et al., 2010); the related VPS26 is linked with BACE/memapsin2 (He et al., 2005). CL44A4 is involved in choline uptake (Jurgensen and Ferreira, 2010). MTMR2 has relevance to excitatory synapses (Lee et al., 2010). ZKSCAN3/ZNF263 is associated with vascular endothelial growth factor (Yang et al., 2008).

The above data suggest involvement of multiple genes influencing varied, complex pathways that might interact mutually to contribute to LOAD. Output from Para-ICA lends itself readily to functional pathway analysis and ultimately systems biology. We also identified novel putative LOAD risk genes, confirmed via testing allelic frequency distributions among diagnostic groups in standard case–control association analyses. It is notable that while none of these genes survived a standard GWAS study, they have high impact when their effect is evaluated in the context of other SNPs. In addition, the SNP components detected several genes previously unknown in the context of LOAD risk, having high Z scores, exceeding those for APOE. Several of these (e.g. SLC9A7, ZNF673, VPS13) were: (a) identified by multiple (up to 17) SNPs, (b) mediate processes plausibly associated with LOAD risk from pathway analyses and prior publications, (c) had SNPs differentially distributed among diagnostic groups and (d) are prominently expressed in brain. These results suggest validity of these novel loci as candidate LOAD risk genes.

Examining loading coefficients of the gene and structural networks revealed a stepped response pattern (see Fig. S2), with MCI values falling between those of healthy control and AD, except in G3, where they were elevated in MCI compared to AD, suggesting that this gene component may act to either protect against or hasten regional brain deterioration in MCI to influence progression rate to AD.

Our study has limitations. Although the Para-ICA method is data driven, we restricted the genetic dataset to a disease-related subset. This focused analysis might fail to uncover genes affecting LOAD pathology via other interactive pathways that may not straightforwardly show group differences. However, since we employed a liberal statistical threshold to limit the genetic dataset to disease-related genes, we were able to include numerous SNPs discarded by conventional univariate studies. Our AD/MCI-focused gene set analysis may not have detected other genetic associations to brain structure. The analysis was carried out only in European-Americans, by far the most numerous ethnicity in the dataset. Future studies could include larger mixed populations. Also, since Para-ICA identified multivariate relationships at the gene network level (comprised of linear combinations of SNPs), the directionality and effect magnitude of individual SNPs is not immediately transparent. Our supplementary association analysis to derive the top SNPs might be slightly biased, as they were already pre-selected at a liberal cut-off to be included in the multivariate analysis. Given these limitations and the novelty of our study, our results require further validation and replication in more diverse and larger independent datasets.

In conclusion, we met our major study goals by 1) confirming the feasibility of a hypothesis-blind, multivariate approach to corroborate LOAD genes associated with known pathologic mechanisms and to discover new putative disease-relevant genes that interact but fail individually to reach genome-wide significance. These data thus extend existing GWAS and hypothesis-driven analyses on the same ADNI data set (Biffi et al., 2010; Saykin et al., 2010; Shen et al., 2010). 2) The Para-ICA approach identifies genes in relatively modest sized samples that are plausibly linked collectively in known physiologic pathways, perhaps epistatically and suggests itself as a novel method for exploring other large-scale data sets involving gene and endophenotype information such as BSNIP or COGS (Calkins et al., 2007; Thaker, 2008), in psychotic disorders where the neuropathology and genetic basis are less well-defined than LOAD. Finally, 3) we identified plausible new biological pathways associated with AD neuropathology. Possible therapies resulting from our findings might include agents targeted to the complement and/or immune systems.

Conflict of interest statement

Acknowledgments

References

Bamburg, J.R., Bloom, G.S., 2009. Cytoskeletal pathologies of Alzheimer disease. Cell Motil. Cytoskeleton 66, 635–649.

Bareiss, S., Kim, K., Lu, Q., 2010. Delta-catenin/NPRAP: a new member of the glycogen synthase kinase-3beta signaling complex that promotes beta-catenin turnover in neurons. J. Neurosci. Res. 88, 2350–2363.

Bekris, L.M., Yu, C.E., Bird, T.D., Tsuang, D.W., 2010. Genetics of Alzheimer disease. J. Geriatr. Psychiatry Neurol. 23, 213–227.

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

Calkins, M.E., Dobie, D.J., Cadenhead, K.S., Olincy, A., Freedman, R., Green, M.F., Greenwood, T.A., Gur, R.E., Gur, R.C., Light, G.A., Mintz, J., Nuechterlein, K.H., Radant, A.D., Schork, N.J., Seidman, L.J., Siever, L.J., Silverman, J.M., Stone, W.S., Swerdlow, N.R., Tsuang, D.W., Tsuang, M.T., Turetsky, B.I., Braff, D.L., 2007. The Consortium on the Genetics of Endophenotypes in Schizophrenia: model recruitment, assessment, and endophenotyping methods for a multisite collaboration. Schizophr Bull 33, 33–48.

Calhoun, V.D., Adali, T., Pearlson, G.D., Pekar, J.J., 2001. A method for making group inferences from functional MRI data using independent component analysis. Hum. Brain Mapp. 14, 140–151.

Calhoun, V.D., Liu, J., Adali, T., 2009. A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. Neuroimage 45, S163–S172.

Dominguez, R., 2009. Actin filament nucleation and elongation factors—structure–function relationships. Crit. Rev. Biochem. Mol. Biol. 44, 351–366.

Eccles, D., Tapper, W., 2010. The influence of common polymorphisms on breast cancer. Cancer Treat. Res. 155, 15–32.

Eikelenboom, P., Veerhuis, R., Scheper, W., Rozemuller, A.J., van Gool, W.A., Hoozemans, J.J., 2006. The significance of neuroinflammation in understanding Alzheimer's disease. J. Neural. Transm. 113, 1685–1695.

Gallo, G., 2007. Tau is actin up in Alzheimer's disease. Nat. Cell Biol. 9, 133–134.

Gandhi, S., Wood, N.W., 2010. Genome-wide association studies: the key to unlocking neurodegeneration? Nat. Neurosci. 13, 789–794.

Gatz, M., Mortimer, J.A., Fratiglioni, L., Johansson, B., Berg, S., Reynolds, C.A., Pedersen, N.L., 2006a. Potentially modifiable risk factors for dementia in identical twins. Alzheimers Dement. 2, 110–117.

Gatz, M., Reynolds, C.A., Fratiglioni, L., Johansson, B., Mortimer, J.A., Berg, S., Fiske, A., Pedersen, N.L., 2006b. Role of genes and environments for explaining Alzheimer disease. Arch. Gen. Psychiatry 63, 168–174.

Grupe, A., Abraham, R., Li, Y., Rowland, C., Hollingworth, P., Morgan, A., Jehu, L., Segurado, R., Stone, D., Schadt, E., Karnoub, M., Nowotny, P., Tacey, K., Catanese, J., Sninsky, J., Brayne, C., Rubinsztein, D., Gill, M., Lawlor, B., Lovestone, S., Holmans, P., O'Donovan, M., Morris, J.C., Thal, L., Goate, A., Owen, M.J., Williams, J., 2007. Evidence for novel susceptibility genes for late-onset Alzheimer's disease from a genome-wide association study of putative functional variants. Hum. Mol. Genet. 16, 865–873.

Guerreiro, R.J., Gustafson, D.R., Hardy, J., 2010. The genetic architecture of Alzheimer's disease: beyond APP, PSENs and APOE. Neurobiol. Aging Epub ahead of print, http://dx.doi.org/10.1016/j.neurobiolaging.2010.03.025.

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., 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, X., Li, F., Chang, W.P., Tang, J., 2005. GGA proteins mediate the recycling pathway of memapsin 2 (BACE). J. Biol. Chem. 280, 11696–11703.

Hollingworth, P., Harold, D., Jones, L., Owen, M.J., Williams, J., 2011. Alzheimer's disease genetics: current knowledge and future challenges. Int. J. Geriatr. Psychiatry 26 (8), 793–802.

Jack Jr., C.R., Bernstein, M.A., Borowski, B.J., Gunter, J.L., Fox, N.C., Thompson, P.M., Schuff, N., Krueger, G., Killiany, R.J., Decarli, C.S., Dale, A.M., Carmichael, O.W., Tosun, D., Weiner, M.W., 2010a. Update on the magnetic resonance imaging core of the Alzheimer's disease neuroimaging initiative. Alzheimers Dement. 6, 212–220.

Jack Jr., C.R., Wiste, H.J., Vemuri, P., Weigand, S.D., Senjem, M.L., Zeng, G., Bernstein, M.A., Gunter, J.L., Pankratz, V.S., Aisen, P.S., Weiner, M.W., Petersen, R.C., Shaw, L.M., Trojanowski, J.Q., Knopman, D.S., 2010b. Brain beta-amyloid measures and magnetic resonance imaging atrophy both predict time-to-progression from mild cognitive impairment to Alzheimer's disease. Brain 133, 3336–3348.

Jagannathan, K., Calhoun, V.D., Gelernter, J., Stevens, M.C., Liu, J., Bolognani, F., Windemuth, A., Ruano, G., Assaf, M., Pearlson, G.D., 2010. Genetic associations of brain structural networks in schizophrenia: a preliminary study. Biol. Psychiatry 68, 657–666.

Jurgensen, S., Ferreira, S.T., 2010. Nicotinic receptors, amyloid-beta, and synaptic failure in Alzheimer's disease. J. Mol. Neurosci. 40, 221–229.

Kagami, T., Chen, S., Memar, P., Choi, M., Foster, L.J., Numata, M., 2008. Identification and biochemical characterization of the SLC9A7 interactome. Mol. Membr. Biol. 25, 436–447.

Kolev, M.V., Ruseva, M.M., Harris, C.L., Morgan, B.P., Donev, R.M., 2009. Implication of complement system and its regulators in Alzheimer's disease. Curr. Neuropharmacol. 7, 1–8.

Koncina, E., Roth, L., Gonthier, B., Bagnard, D., 2007. Role of semaphorins during axon growth and guidance. Adv. Exp. Med. Biol. 621, 50–64.

Kostiuk, O.P., Korol, T., Korol, S.V., Romanenko, S.V., Pinchenko, V.O., Kostiuk, P.H., 2010. Alteration of calcium signaling as one of the mechanisms of Alzheimer's disease and diabetic polyneuropathy. Fiziol. Zh. 56, 130–138.

Ku, C.S., Loy, E.Y., Pawitan, Y., Chia, K.S., 2011. The pursuit of genome-wide association studies: where are we now? J. Hum. Genet. 55, 195–206.

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

Lee, H.W., Kim, Y., Han, K., Kim, H., Kim, E., 2010. The phosphoinositide 3-phosphatase MTMR2 interacts with PSD-95 and maintains excitatory synapses by modulating endosomal traffic. J. Neurosci. 30, 5508–5518.

Li, D., Parks, S.B., Kushner, J.D., Nauman, D., Burgess, D., Ludwigsen, S., Partain, J., Nixon, R.R., Allen, C.N., Irwin, R.P., Jakobs, P.M., Litt, M., Hershberger, R.E., 2006. Mutations of presenilin genes in dilated cardiomyopathy and heart failure. Am. J. Hum. Genet. 79, 1030–1039.

Lin, P.J., Williams, W.P., Luu, Y., Molday, R.S., Orlowski, J., Numata, M., 2005. Secretory carrier membrane proteins interact and regulate trafficking of the organellar (Na+, K+)/H+ exchanger NHE7. J. Cell Sci. 118, 1885–1897.

Lin, P.I., Martin, E.R., Browning-Large, C.A., Schmechel, D.E., Welsh-Bohmer, K.A., Doraiswamy, P.M., Gilbert, J.R., Haines, J.L., Pericak-Vance, M.A., 2006. Parsing the genetic heterogeneity of chromosome 12q susceptibility genes for Alzheimer disease by family-based association analysis. Neurogenetics 7, 157–165.

Liu, J., Demirci, O., Calhoun, V.D., 2008. A parallel independent component analysis approach to investigate genomic influence on brain function. IEEE Signal Process. Lett. 15, 413–416.

Lugtenberg, D., Yntema, H.G., Banning, M.J., Oudakker, A.R., Firth, H.V., Willatt, L., Raynaud, M., Kleefstra, T., Fryns, J.P., Ropers, H.H., Chelly, J., Moraine, C., Gecz, J., van Reeuwijk, J., Nabuurs, S.B., de Vries, B.B., Hamel, B.C., de Brouwer, A.P., van Bokhoven, H., 2006. ZNF674: a new kruppel-associated box-containing zinc-finger gene involved in nonsyndromic X-linked mental retardation. Am. J. Hum. Genet. 78, 265–278.

Marks, N., Berg, M.J., 2010. BACE and gamma-secretase characterization and their sorting as therapeutic targets to reduce amyloidogenesis. Neurochem. Res. 35, 181–210.

Mattson, M.P., Chan, S.L., 2003. Neuronal and glial calcium signaling in Alzheimer's disease. Cell Calcium 34, 385–397.

Meda, S.A., Jagannathan, K., Gelernter, J., Calhoun, V.D., Liu, J., Stevens, M.C., Pearlson, G.D., 2010. A pilot multivariate parallel ICA study to investigate differential linkage between neural networks and genetic profiles in schizophrenia. Neuroimage 53, 1007–1015.

Meunier, S., Navarro, M.G., Bossard, C., Laurell, H., Touriol, C., Lacazette, E., Prats, H., 2009. Pivotal role of translokin/CEP57 in the unconventional secretion versus nuclear translocation of FGF2. Traffic 10, 1765–1772.

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.

Ohgaki, R., Fukura, N., Matsushita, M., Mitsui, K., Kanazawa, H., 2008. Cell surface levels of organellar Na+/H+ exchanger isoform 6 are regulated by interaction with RACK1. J. Biol. Chem. 283, 4417–4429.

Osterfield, M., Egelund, R., Young, L.M., Flanagan, J.G., 2008. Interaction of amyloid precursor protein with contactins and NgCAM in the retinotectal system. Development 135, 1189–1199.

Pimplikar, S.W., Nixon, R.A., Robakis, N.K., Shen, J., Tsai, L.H., 2010. Amyloid-independent mechanisms in Alzheimer's disease pathogenesis. J. Neurosci. 30, 14946–14954.

Potkin, S.G., Guffanti, G., Lakatos, A., Turner, J.A., Kruggel, F., Fallon, J.H., Saykin, A.J., Orro, A., Lupoli, S., Salvi, E., Weiner, M., Macciardi, F., 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.

Profenno, L.A., Porsteinsson, A.P., Faraone, S.V., 2010. Meta-analysis of Alzheimer's disease risk with obesity, diabetes, and related disorders. Biol. Psychiatry 67, 505–512.

Ramaswamy, V., Castillo, M., Bolduc, F.V., 2010. Developmental disability: duplication of zinc finger transcription factors 673 and 674. Pediatr. Neurol. 43, 209–212.

Saxena, R., Hivert, M.F., Langenberg, C., Tanaka, T., Pankow, J.S., Vollenweider, P., Lyssenko, V., Bouatia-Naji, N., Dupuis, J., Jackson, A.U., Kao, W.H., Li, M., Glazer, N.L., Manning, A.K., Luan, J., Stringham, H.M., Prokopenko, I., Johnson, T., Grarup, N., Boesgaard, T.W., Lecoeur, C., Shrader, P., O'Connell, J., Ingelsson, E., Couper, D.J., Rice, K., Song, K., Andreasen, C.H., Dina, C., Kottgen, A., Le Bacquer, O., Pattou, F., Taneera, J., Steinthorsdottir, V., Rybin, D., Ardlie, K., Sampson, M., Qi, L., van Hoek, M., Weedon, M.N., Aulchenko, Y.S., Voight, B.F., Grallert, H., Balkau, B., Bergman, R.N., Bielinski, S.J., Bonnefond, A., Bonnycastle, L.L., Borch-Johnsen, K., Bottcher, Y., Brunner, E., Buchanan, T.A., Bumpstead, S.J., Cavalcanti-Proenca, C., Charpentier, G., Chen, Y.D., Chines, P.S., Collins, F.S., Cornelis, M., G.J.C., Delplanque, J., Doney, A., Egan, J.M., Erdos, M.R., Firmann, M., Forouhi, N.G., Fox, C.S., Goodarzi, M.O., Graessler, J., Hingorani, A., Isomaa, B., Jorgensen, T., Kivimaki, M., Kovacs, P., Krohn, K., Kumari, M., Lauritzen, T., Levy-

Marchal, C., Mayor, V., McAteer, J.B., Meyre, D., Mitchell, B.D., Mohlke, K.L., Morken, M.A., Narisu, N., Palmer, C.N., Pakyz, R., Pascoe, L., Payne, F., Pearson, D., Rathmann, W., Sandbaek, A., Sayer, A.A., Scott, L.J., Sharp, S.J., Sijbrands, E., Singleton, A., Siscovick, D.S., Smith, N.L., Sparso, T., Swift, A.J., Syddall, H., Thorleifsson, G., Tonjes, A., Tuomi, T., Tuomilehto, J., Valle, T.T., Waeber, G., Walley, A., Waterworth, D.M., Zeggini, E., Zhao, J.H., Illig, T., Wichmann, H.E., Wilson, J.F., van Duijn, C., Hu, F.B., Morris, A.D., Frayling, T.M., Hattersley, A.T., Thorsteinsdottir, U., Stefansson, K., Nilsson, P., Syvanen, A.C., Shuldiner, A.R., Walker, M., Bornstein, S.R., Schwarz, P., Williams, G.H., Nathan, D.M., Kuusisto, J., Laakso, M., Cooper, C., Marmot, M., Ferrucci, L., Mooser, V., Stumvoll, M., Loos, R.J., Altshuler, D., Psaty, B.M., Rotter, J.I., Boerwinkle, E., Hansen, T., Pedersen, O., Florez, J.C., McCarthy, M.I., Boehnke, M., Barroso, I., Sladek, R., Froguel, P., Meigs, J.B., Groop, L., Wareham, N.J., Watanabe, R.M., 2010. Genetic variation in GIPR influences the glucose and insulin responses to an oral glucose challenge. Nat. Genet. 42, 142–148.

Saykin, A.J., Shen, L., Foroud, T.M., Potkin, S.G., Swaminathan, S., Kim, S., Risacher, S.L., Nho, K., Huentelman, M.J., Craig, D.W., Thompson, P.M., Stein, J.L., Moore, J.H., Farrer, L.A., Green, R.C., Bertram, L., Jack Jr., C.R., Weiner, M.W., 2010. Alzheimer's Disease Neuroimaging Initiative biomarkers as quantitative phenotypes: genetics core aims, progress, and plans. Alzheimers Dement. 6, 265–273.

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., 2010. Genome-wide analysis of genetic loci associated with Alzheimer disease. JAMA 303, 1832–1840.

Shen, L., Kim, S., Risacher, S.L., Nho, K., Swaminathan, S., West, J.D., Foroud, T., Pankratz, N., Moore, J.H., Sloan, C.D., Huentelman, M.J., Craig, D.W., Dechairo, B.M., Potkin, S.G., Jack Jr., C.R., Weiner, M.W., Saykin, A.J., 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.

Smith, C.D., 2010. Neuroimaging through the course of Alzheimer's disease. J. Alzheimers Dis. 19, 273–290.

Stieber, A., Mourelatos, Z., Gonatas, N.K., 1996. In Alzheimer's disease the Golgi apparatus of a population of neurons without neurofibrillary tangles is fragmented and atrophic. Am. J. Pathol. 148, 415–426.

Strittmatter, W.J., Saunders, A.M., Schmechel, D., Pericak-Vance, M., Enghild, J., Salvesen, G.S., Roses, A.D., 1993. Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. Proc. Natl. Acad. Sci. U. S. A. 90, 1977–1981.

Terni, B., Boada, J., Portero-Otin, M., Pamplona, R., Ferrer, I., 2010. Mitochondrial ATP-synthase in the entorhinal cortex is a target of oxidative stress at stages I/II of Alzheimer's disease pathology. Brain Pathol. 20, 222–233.

Thaker, G., 2008. Psychosis endophenotypes in schizophrenia and bipolar disorder. Schizophr Bull 34, 720–721.

Traynor, B.J., Singleton, A.B., 2010. Nature versus nurture: death of a dogma, and the road ahead. Neuron 68, 196–200.

Urcelay, E., Ibarreta, D., Parrilla, R., Ayuso, M.S., Martin-Requero, A., 2001. Enhanced proliferation of lymphoblasts from patients with Alzheimer dementia associated with calmodulin-dependent activation of the Na+/H+ exchanger. Neurobiol. Dis. 8, 289–298.

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

Veerhuis, R., 2011. Histological and direct evidence for the role of complement in the neuroinflammation of AD. Curr. Alzheimer Res. 8 (1), 34–58.

Villain, N., Chetelat, G., Desgranges, B., Eustache, F., 2010. Neuroimaging in Alzheimer's disease: a synthesis and a contribution to the understanding of physiopathological mechanisms. Biol Aujourdhui 204, 145–158.

Walker, D.G., Dalsing-Hernandez, J.E., Lue, L.F., 2008. Human postmortem brain-derived cerebrovascular smooth muscle cells express all genes of the classical complement pathway: a potential mechanism for vascular damage in cerebral amyloid angiopathy and Alzheimer's disease. Microvasc. Res. 75, 411–419.

Yang, L., Zhang, L., Wu, Q., Boyd, D.D., 2008. Unbiased screening for transcriptional targets of ZKSCAN3 identifies integrin beta 4 and vascular endothelial growth factor as downstream targets. J. Biol. Chem. 283, 35295–35304.