Published in final edited form as:
JAMA. 2010 May 12; 303(18): 1832–1840. doi:10.1001/jama.2010.574.

Genome-wide Analysis of Genetic Loci Associated with Alzheimer's Disease

Sudha Seshadri, MD, **Annette L. Fitzpatrick, PhD**, **M Arfan Ikram, MD, PhD**, **Anita L. DeStefano, PhD**, **Vilmundur Gudnason, MD, PhD**, **Merce Boada, MD, PhD**, **Joshua C. Bis, PhD**, **Albert V. Smith, PhD**, **Minerva M. Carassquillo, PhD**, **Jean Charles Lambert, PhD**, **Denise Harold, PhD**, **Elisabeth M. C. Schrijvers, MD**, **Reposo Ramirez-Lorca, PhD**, **Stephanie Debette, MD PhD**, **W.T. Longstreth Jr., MD**, **A. Cecile J.W. Janssens, PhD**, **V. Shane Pankratz, PhD**, **Jean François Dartigues, PhD**, **Paul Hollingworth, PhD**, **Thor Aspelund, PhD**, **Isabel Hernandez, MD**, **Alexa Beiser, PhD**, **Lewis H. Kuller, MD**, **Peter J. Koudstaal, MD, PhD**, **Dennis W. Dickson, MD**, **Christophe Tzourio, MD**, **Richard Abraham, PhD**, **Carmen Antunez, MD**, **Yangchun Du, PhD**, **Jerome I. Rotter, MD**, **Yurii S. Aulchenko, PhD**, **Tamara B. Harris, MD**, **Ronald C. Petersen, MD**, **Claudine Berr, MD, PhD**, **Michael J. Owen, MbChB, PhD**, **Jesus Lopez-Arrieta, MD**, **Badri N. Varadarajan, MS**, **James T. Becker, PhD**, **Fernando Rivadeneira, MD, PhD**, **Michael A. Nalls, PhD**, **Neill R. Graff-Radford, MD**, **Dominique Campion, MD, PhD**, **Sanford Auerbach, MD**, **Kenneth Rice, PhD**, **Albert Hofman, MD, PhD**, **Palmi V. Jonsson, MD**, **Helena Schmidt, MD, PhD**, **Mark Lathrop, PhD**, **Thomas H. Mosley, PhD**, **Rhoda Au, PhD**, **Bruce M. Psaty, MD, PhD**, **Andre G. Uitterlinden, PhD**, **Lindsay A. Farrer, PhD**, **Thomas Lumley, PhD**, **Agustin Ruiz, MD, PhD**, **Julie Williams, PhD**, **Philippe Amouyel, MD, PhD**, **Steve G. Younkin, PhD**, **Philip A. Wolf, MD**, **Lenore J. Launer, PhD**, **Oscar L. Lopez, MD**, **Cornelia M. van Duijn, PhD**, and **Monique M. B. Breteler, MD, PhD on behalf of the CHARGE, GERAD1, and EADI1 consortia**

Department of Neurology, (SS, ALD, SD, AB, YD, SA, RA, LAF, PAW) and Medicine (Genetics Program: BNV, LAF) Boston University School of Medicine, Boston, MA, USA; The National Heart Lung and Blood Institute's Framingham Heart Study (SS, ALD, SD, AB, SA, RA, PAW), Framingham, MA, USA; Departments of Epidemiology (ALF, WTL, BMP), Global Health (AF), Medicine (JCB, WTL, BMP), Neurology (WTL), Biostatistics (KR, TL), Neurology (WTL) and Health Services (BMP), University of Washington, Seattle, WA, USA; The Center for Health

Corresponding Author: Monique M.B. Breteler, MD, PhD, Professor of Neuroepidemiology, Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, PO Box 2040, 3000 CA Rotterdam, The Netherlands, Tel: +31 10 704 3489, Fax: +31 10 704 4657, m.breteler@erasmusmc.nl.

Authorship Responsibilities

Drs. Seshadri, Ikram, Destefano and Breteler had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

The funding organizations and sponsors had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review or approval of the manuscript. The final version submitted was approved without changes by the National Heart, Lung and Blood Institutes and the National Institute on Aging.

Overall meta-analyses that are presented in this paper were undertaken by Dr. Anita L. DeStefano and findings were cross-checked by Dr Cornelia M. van Duijn. Analyses specific to each study were undertaken by Drs. Albert V. Smith (for AGES), Joshua C. Bis (for CHS), Anita L. DeStefano (for FHS), M. Arfan Ikram (for the Rotterdam Study), Badri N. Varadarajan (for TGEN), Steven Younkin (for the Mayo AD GWAS), Jean Charles Lambert (for EADI1), Denise Harald (for GERAD), and Agustin Ruiz (for Fundacio Ace)

Author contributions

The following authors contributed to study design: SS, ALD, VG, JCL, WTL, TL, TH, AB, AGU, AH, BMP, PA, PAW, LJL, OLL, CMvD, MMBB; to data acquisition: SS, ALF, MAI, MB, PVJ, EMCS, RRL, PJK, FR, CB, JIR, DWD, CT, ML, JFD, RCP, JTB, LHK, MMC, IH, AB, CA, JL-A, YD, RA, BMP, SGY, LAF, NRGR, PAW; to data analysis and interpretation: SS, ALF, MAI, ALD, JCB, SD, TL, KMR, MMC, ML, LHK, SA, EMCS, SGY, JTB, ACJWJ, VSP, JCL, VG, AB, DC, THM, YD, AVS, TA, BNV, PJK, RCP, BMP, YSA, FR, TH, AR, PA, OLL, LJL, MAN, MMBB; to statistical analysis: MAI, ALD, JCB, AVS, KMR, TA, TL, LAF, YSA, BNV, AR, CMvD; to funding and supervision: SS, VG, MB, CA, PVJ, AGU, LJL, AH, BMP, TH, AR, PAW, OLL, MMBB; to drafting the manuscript: SS, MAI, CMvD, MMBB; to critical revision of the manuscript and final approval to submit: all authors; to other aspects of the research (neuropathologic diagnosis in the Mayo samples): DWD

Studies, Group Health, Seattle, WA, USA (ALF, WTL, BMP); Departments of Epidemiology (MAI, ACJWJ, EMCS, YSA, AH, CMvD, MMBB), Neurology (PJK), Internal Medicine (FR, AGU), and Clinical Chemistry (AGU), from the Erasmus MC University Medical Center, Rotterdam, The Netherlands; Netherlands Consortium for Health Aging – NCHA (MAI, ACJWJ, EMCS, YSA, FR, AH, AGU, CMvD, MMBB); Departments of Biostatistics (ALD, AB, YD, LAF) and Epidemiology (LAF), Boston University School of Public Health, Boston, MA, USA; Icelandic Heart Association (VG, AVS, PVJ), Kopavogur, Iceland; University of Iceland (VG, TA, PVJ), Reykjavik, Iceland; Landspitali University Hospital (PVJ), Iceland; Memory Clinic of Fundació ACE. Institut Català de Neurociències Aplicades (MB, IH), Catalan Institute of Applied Neurosciences, and the Neurology Department (MB), Hospital G. Universitari Vall d'Hebron, Barcelona, Spain; Department of Neuroscience (MMC, DWD, NRG, SGY), Mayo Clinic College of Medicine, Jacksonville, Florida 32224, USA; Inserm U744 (PA), Institut Pasteur de Lille and Université de Lille Nord de France (JCL), Lille, France; Medical Research Council (MRC) Centre for Neuropsychiatric Genetics and Genomics (DH, PH, RA, MJO, JW) Department of Psychological Medicine and Neurology, School of Medicine, Cardiff University, Cardiff, UK; Department of Structural Genomics. Neocodex (RRL, AR), Sevilla, Spain; Division of Biomedical Statistics and Informatics (VSP), Mayo Clinic and Mayo Foundation, Rochester, Minnesota 55905, USA; Inserm U897 (JFD), Victor Segalen University, Bordeaux, France. Lille, France. Dementia Unit. University Hospital Virgen de la Arrixaca (CA), Murcia, Spain; The Departments of Neurology and Psychiatry (OL, JTB), Epidemiology (LHK), and Psychology (JTB) and The Alzheimer's Disease Research Center (OL, JTB), University of Pittsburgh School of Medicine, Pittsburgh, PA, USA, Inserm U708 (CT), and the Université Pierre et Marie Curie Paris 6, Paris, France; Alzheimer Foundation (CA), Murcia, Spain; The Medical Genetics Institute (JIR), Cedars-Sinai Medical Center, Los Angeles, CA, USA; Laboratory of Epidemiology, Demography and Biometry (TBH, LJL); Laboratory of Neurogenetics (MAN), Intramural Research Program, National Institute on Aging, Washington, DC, USA; Department of Neurology (NGR, RCP), and Mayo Alzheimer Disease Research Center (RP), Mayo Clinic, College of Medicine, Rochester, Minnesota 55905, USA; Inserm U888 (CB), Hôpital La Colombière, Montpellier, France; Memory Unit. University Hospital La Paz-Cantoblanco (JLA), Madrid, Spain; Inserm U614 (DC), Faculté de Médecine-Pharmacie de Rouen, Rouen, France; Institute of Molecular Biology and Biochemistry and University Clinic of Neurology, Department of Neurogeriatrics (HS), Medical University Graz, Austria; Centre National de Génotypage, Institut Genomique, Commissariat à l'énergie Atomique, Evry, and the Fondation Jean Dausset–Centre d'Etudes du Polymorphisme Humain (ML), Paris, France; Department of Medicine-Geriatrics (THM), University of Mississippi Medical Center, Jackson, MS, USA; Centre Hospitalier Régional Universitaire de Lille (PA), Lille, France

Abstract

Context—Genome wide association studies (GWAS) have recently identified *CLU, PICALM* and *CR1* as novel genes for late-onset Alzheimer's disease (AD).

Objective—In a three-stage analysis of new and previously published GWAS on over 35000 persons (8371 AD cases), we sought to identify and strengthen additional loci associated with AD and confirm these in an independent sample. We also examined the contribution of recently identified genes to AD risk prediction.

Design, Setting, and Participants—We identified strong genetic associations ($p<10^{-3}$) in a Stage 1 sample of 3006 AD cases and 14642 controls by combining new data from the population-based Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium (1367 AD cases (973 incident)) with previously reported results from the Translational Genomics Research Institute (TGEN) and Mayo AD GWAS. We identified 2708 single nucleotide polymorphisms (SNPs) with p-values$<10^{-3}$, and in Stage 2 pooled results for these SNPs with the European AD Initiative (2032 cases, 5328 controls) to identify ten loci with p-values$<10^{-5}$. In Stage 3, we combined data for these ten loci with data from the Genetic and Environmental Risk

in AD consortium (3333 cases, 6995 controls) to identify four SNPs with a p-value<1.7×10^{-8}. These four SNPs were replicated in an independent Spanish sample (1140 AD cases and 1209 controls).

Main outcome measure—Alzheimer's Disease.

Results—We showed genome-wide significance for two new loci: rs744373 near *BIN1* (OR: 1.13; 95%CI:1.06–1.21 per copy of the minor allele; p=1.6×10^{-11}) and rs597668 near *EXOC3L2/BLOC1S3/MARK4* (OR:1.18; 95%CI1.07–1.29; p=6.5×10^{-9}). Associations of *CLU*, *PICALM*, *BIN1* and *EXOC3L2* with AD were confirmed in the Spanish sample (p<0.05). However, *CLU* and *PICALM* did not improve incident AD prediction beyond age, sex, and *APOE* (improvement in area under receiver-operating-characteristic curve <0.003).

Conclusions—Two novel genetic loci for AD are reported that for the first time reach genome-wide statistical significance; these findings were replicated in an independent population. Two recently reported associations were also confirmed, but these loci did not improve AD risk prediction, although they implicate biological pathways that may be useful targets for potential interventions.

Keywords

genome-wide association study; genetic epidemiology; genetics; dementia; Alzheimer's disease; cohort study; meta-analysis; risk

---

It is currently estimated that one of every five persons aged 65 years will develop Alzheimer's Disease (AD) in their lifetime, and that genetic variants may play an important part in the development of the disease.1 The substantial heritability of late-onset AD2 is inadequately explained by genetic variation within the well-replicated genes (apolipiprotein E (*APOE*(RefSeq NG_007084)), presenilin-1 (*PSEN1*(RefSeq NG_007386)), presenilin-2 (*PSEN2*(RefSeq NG_007381)), and amyloid beta precursor protein (*APP*(RefSeq NM_000484)).3 Initial genome-wide association studies (GWAS) identified putative new candidate genes (GRB2-associated binding protein (*GAB2*(RefSeq NG_016171)), protocadherin 11 x-linked (*PCDH11X*(RefSeq NG_016251)), lecithin retinol acyltransferase (*LRAT*(RefSeq NG_009110)), transient receptor potential cation channel, subfamily C, member 4 associated protein (*TRPC4AP*(RefSeq NM_015638))4–6 and regions of interest (e.g. on chromosomes 14q, 10q, 12q)7–10 but no locus outside the *APOE*-region consistently reached genome-wide significance.4, 11, 12 These disappointing results are most likely explained by the modest sample size and hence limited statistical power of early studies to detect genes with small effects. Recently, two large GWAS, the UK-led Genetic and Environmental Risk in Alzheimer's Disease 1 consortium (GERAD1),13 and the European Alzheimer Disease Initiative (EADI) Stage 1,14 reported 3 new genome-wide significant loci for AD: within the *CLU* gene (GenBank AY341244) encoding clusterin (also called apolipoprotein J), near the *PICALM* gene (GenBank BC073961) encoding phosphatidylinositol binding clathrin assembly protein, and within the *CR1* (RefSeq NG_007481) gene encoding complement component (3b/4b) receptor 1.13, 14

We performed a three-stage analysis of GWAS data to identify additional loci associated with late-onset AD. Moreover, we sought to replicate genome-wide significant loci, both from the current analysis and previous reports, in an independent case-control population. Finally, we utilized two large prospective population based studies to assess the improvement in incident AD risk prediction conferred by the recently described loci.

Methods

Gene Discovery

Setting—We used a three-stage sequential analysis to identify novel loci associated with late-onset AD (Figure 1). Our initial discovery was a meta-analysis combining new GWA data from white participants in the large, population-based Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium,15 with GWA data from the Translational Genomics Research Institute (TGEN) public release database4 and the Mayo AD GWAS.5 The sample characteristics of the participants contributing to this discovery stage (stage 1) are summarized in Table 1. Next, we combined results for our most suggestive findings (SNPs with p-value<10^{-3}) with corresponding results in the EADI1 consortium (stage 2).14 Finally, in stage 3, we combined results for the most promising hits in stage 2 (selecting top SNPs from all loci that reached a p-value <10^{-5}) with data from the non-overlapping studies within the GERAD1 consortium (excluding the Mayo AD GWAS, the only overlapping study).13 All participants (or their authorized proxies) in the contributing studies gave written informed consent including for genetic analyses. Local institutional review boards approved study protocols. Details of study sample selection for the contributing studies are described in section 2 of the Supplementary material (section 1 lists commonly used abbreviations) and in Supplementary Figures 1A to 1D.

In each study, dementia was defined using the Diagnostic and Statistical Manual of Mental Disorders revised third or fourth edition (DSM-IIIR or DSM-IV) criteria.16 Among persons with dementia, all studies used the National Institute of Neurological and Communicative Disorders and Stroke and Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria to define AD, and included persons with definite (diagnosis of AD pathologically confirmed at autopsy), probable or possible AD.17

Genotyping—The individual studies in stage 1 were genotyped on different platforms as detailed in Table 1. The EADI1 used the Illumina Quad 6.0 and GERAD1 was genotyped on various Illumina chips. In each of the CHARGE cohorts and in TGEN, we used the genotype data to impute to the 2.5 million non-monomorphic, autosomal SNPs described in HapMap (CEU population). Imputations are needed when one wants to meta-analyze genome-wide association data across studies that have used different genotyping platforms, because the platforms differ in the SNPs genotyped. Imputation methods and QC filters in each sample are described in the Supplementary material (Section 3).

GWA analyses in stage 1 studies—All analyses were restricted to white persons, racial identity being self-defined by the participants (see section 2 of the online supplement for additional details). We included a few white Hispanics and adjusted for population structure. Since only one of the CHARGE studies, CHS, had a small number of African American participants (n=574 with genotyping) this racial subgroup was too small for independent analysis. Linkage disequilibrium patterns are very different in African persons and this leads to greater uncertainty in imputation, as well as the possibility of false positive associations if data from two racial groups are combined when disease risk differs by race (a phenomenon called population stratification), hence African-American participants in the CHS study were excluded from these analyses. Each study fit an additive genetic model – a 1 degree of freedom trend test – relating genotype dosage (0 to 2 copies of the minor allele) to study trait. In the CHARGE cohorts, prevalent cases were compared to controls free of dementia at the DNA draw date. Participants were excluded if they declined consent or failed genotyping. For analysis of prevalent events in the CHARGE cohorts and for the case-control data from TGEN and Mayo we used logistic regression models. For the analysis of incident events in the CHARGE cohorts, participants who were free of dementia entered the

analysis at the time of the DNA sample collection and were followed until the development of incident AD; participants were censored at death, at the time of their last follow-up examination or health status update when they were known to be free of clinical dementia, and when they developed dementia due to an alternate cause. We used Cox proportional hazards models to calculate hazard ratios with corresponding 95% confidence intervals after ensuring that assumptions of proportionality of hazards were met. In CHS, FHS, and the Rotterdam Study controls contributed one set of person-years to the prevalent analysis and a second, non-overlapping set of person-years to the incident analyses. Under the martingale property of Cox models, the two analyses are independent and their independence was confirmed in simulation studies. Primary analyses were adjusted for age and sex and any evidence of population stratification. Details of the screening for latent population substructure in each discovery sample are available in section 4 of the Supplementary material. In addition, CHS also adjusted for study site, and FHS accounted for familial relationships (by employing a Cox model with robust variance estimator clustering on pedigree to account for family relationships) and for whether the DNA had been whole genome amplified.

Meta-analyses—Our stage 1 meta-analysis combined results from nine discrete sources: incident AD in the CHS, FHS, and Rotterdam Study, prevalent AD in the AGES, CHS, FHS, and Rotterdam Study, and the TGEN and Mayo case-control studies. We used inverse-variance weighting (also known as a fixed-effects analysis) for meta-analysis applying genomic control to each study of stage 1. This approach assigns greater weight to more precise (study-specific) estimators; thus greater weight is given to studies, in which a given SNP was genotyped or more effectively imputed, and to studies with larger sample sizes. Details of meta-analyses are available in the Supplementary material (Section 5). We retained only those SNP-phenotype associations that were based on results from at least two of the nine discovery samples and where the minor allele frequency was ≥2%. For stages 2 and 3, we again used inverse-variance meta-analysis but without genomic control adjustment. We decided a priori on a genome-wide significance threshold of 1.7×10^{-8} which gives, for a three stage sequential analysis, the same control of false-positives as a single study's use of $p<5\times10^{-8}$.18 The 3 stages of meta-analyses were completed in May to August 2009.

Replication in an Independent Sample

Significant hits from stage 3 of the discovery phase were replicated in an independent Spanish case-control sample (the Fundació ACE) of 1140 AD patients (mean age 78.8±7.9years, 69.9% women) compared to 1209 general population controls (49.9±9.2years; 52.8% women).19, 20 All AD patients fulfilled DSM-IV criteria for dementia and NINCDS-ADRDA criteria for possible and probable AD.16,17 Both cases and controls were whites. Further details of the sample are provided in the Supplementary online appendix (section 6). Genotyping was undertaken using real-time polymerase chain reaction (PCR) coupled to Fluorescence Resonance Energy Transfer (FRET). Effect sizes for single markers were calculated by unconditional logistic regression analysis using SPSS v13.0. software (SPSS Inc., Chicago, IL, USA). Replication was completed in October 2009.

Replication of Previously Reported Associations in CHARGE sample—In secondary analyses, we also examined results for previously reported loci.5, 13, 14 For these loci, which included the recently reported loci by the EADI1 and GERAD1 consortia, we restricted our analysis to the previously unpublished CHARGE data. We did not assess the association with *PCDH11X* since we only focused on autosomal SNPs in these analyses. We did examine associations with the top 15 candidate genes listed in the Alzgene database (http://www.alzforum.org/res/com/gen/alzgene),21 as of 8/12/2009 including the *APOE/*

TOMM40/APOC1 locus and 12 genes outside that locus. Further details of SNPs selected and results for these SNPs are provided in section 7 and in eTable 3 in the Supplementary material.

Genetic Risk Prediction

We sought to estimate the impact of recently identified loci on 10-year risk prediction in the general population using the data for prospectively ascertained, incident AD in the two largest community-based cohort studies at our disposal (Rotterdam Study and CHS). In these analyses, we only included SNPs from the two loci that were shown to be genome-wide significant in previous publications, and that we replicated nominally within CHARGE, *PICALM* and *CLU* (<0.05). Moreover, the analysis was restricted to incident AD to avoid survival bias and was restricted to population-based samples, because case-control studies may overestimate the effects of the genes if cases and controls were not randomly selected from the populations in which AD risk prediction is to be applied.22 The improvement in risk prediction was investigated by comparing three sequentially incremental AD risk prediction models that first incorporated age- and sex- alone, and then added data on risk allele status at the *APOE*, and finally risk allele status at the *CLU* and *PICALM* loci. We did not assess the utility of novel loci uncovered in this paper (using CHARGE as part of the discovery sample) to avoid the risk of overestimating effects by using the same sample for gene discovery and risk prediction.22 Prediction models were constructed using Cox proportional hazards methods using the R-package survcomp. *APOE*ε4 status was included as a discrete variable (0, 1, or 2 alleles) and the other two genetic loci as dosages; all gene effects were examined using additive models. The accuracy of risk prediction for each model was assessed as the discriminative accuracy, measured by the Area under the Receiver Operating Characteristic curve (AUC). AUC theoretically ranges from 0.50 (as predictive as tossing a coin) to 1.00 (perfect prediction).

Results

The stage 1 meta-analysis had 8935 dementia-free individuals (age 72±7 years) of whom 973 developed incident AD over an average follow-up time of 8±3 years, and 2033 prevalent cases of AD who were compared to 14642 dementia-free controls. In this discovery analysis based on the CHARGE cohorts, TGEN and the Mayo GWAS, there was no evidence of spurious inflation of p-values or significant population-stratification (see Supplementary Figure 2 for the quantile-quantile plot comparing the observed and expected p-value distributions). Supplementary figure 3 illustrates the primary findings from the stage 1 meta-analysis in a Manhattan plot showing genome-wide p-values for all interrogated SNPs across the 22 autosomal chromosomes. After stage 1, 2708 SNPs had a p-value$<10^{-3}$ and were studied in stage 2. In stage 2, pooling these results with data from EADI1, 38 SNPs in ten loci had a p-value$<10^{-5}$. Finally, in stage 3, the most significant SNPs from these ten loci were meta-analysed with the non-overlapping studies from GERAD1. The findings of stages 1, 2, and 3 analyses at these 10 loci are presented in Table 2. Additional details are provided in eTable1, which shows chromosomal location, adjacent genes, sample- and stage-specific estimates of relative risks, 95% confidence intervals and p-values for each of the 38 SNPs selected in stage 2 analyses. Figures 2 and 3 are regional association plots for the two SNPs not previously reported to have reached genome-wide significance, rs744373 and rs597668 on chromosomes 2 and 19, respectively. In each Figure we show the linkage-disequilibrium (with the index SNP) and stage 1, 2 and 3 association results for the index SNP and stage 1 results for all SNPs within 200kb on either side of the index SNP at that locus, as well as gene locations and recombination rates in the region. Regional association plots for the other loci listed in Table 2 are presented as Supplemental Figures 4 to 8.

In stage 1, 11 SNPs in the *APOE/TOMM40/APOC1* region reached our pre-set threshold for genome-wide significance (see eTable 1 and Supplemental Figure 3). In stage 2, two additional loci, rs11136000 in *CLU*, and a locus (rs11771145) at chromosome 7 in the 5' upstream promoter/regulatory region of EPH receptor A1 (*EPHA1*(GenBank AH007960)) reached genome-wide significance. However, the latter became non-significant after adding GERAD1 data in stage 3, though the effect seen in GERAD1 was in the same direction in that the same allele was associated with an increased risk of AD. In stage 3, genome-wide significant evidence for association with AD was reached at the *APOE* (rs2075650; $p=1.04\times10^{-295}$), *CLU* (rs11136000; $p=1.62\times10^{-16}$) and *PICALM* (rs3851179; $p=3.16\times10^{-12}$) loci, and for two novel loci on chromosomes 2 (rs744373; $p=1.59\times10^{-11}$), and 19 (rs597668; $p=6.45\times10^{-9}$). Table 2 shows the odds ratios associated with the minor allele for each of these SNPs. Rs744373 is within 30Kb of the gene bridging integrator 1 (*BIN1*(RefSeq NG_012042)) (Figure 2), while rs597668 is within 60Kb of six genes including exocyst complex component 3-like 2 (*EXOC3L2*(RefSeq NM_138568)), biogenesis of lysosomal organelles complex-1, subunit 3 (*BLOC1S3*(RefSeq NG_008372)), and MAP/microtubule affinity-regulating kinase 4 (*MARK4*(GenBank BC071948)) (Figure 3).

Independent Replication

We replicated the four associations that reached our preset genome-wide significance threshold (1.7×10^{-8}) in an independent sample of cases and controls (see Table 3). Effect sizes in the replication cohort were similar to those observed in the discovery sample; each of these associations reached p-value <0.05.

Conditional Analyses at Chromosome 19 locus

Since rs597668 is on chromosome 19, fairly close to the *APOE* locus, we undertook conditional analyses to examine whether its association with AD was independent of *APOE*ε4. We conducted two analyses with AD (among persons with directly genotyped *APOE*ε4 status) in the CHARGE, TGEN and Mayo sample, adjusting (i) for our strongest association in the *APOE/TOMM40/APOC1* locus (rs2075650) and (ii) for the actual *APOE*ε4 SNP, rs429358. In each case, we found that the association was attenuated but a marginal signal remained when adjusting for *APOE*ε4 (OR 1.18, 95% CI 1.08–1.24, $p=3.9\times10^{-4}$ without adjustment and OR 1.17, 1.07–1.23, $p=8.7\times10^{-4}$, and OR 1.10, 1.00–1.16, p=0.05 for analyses (i) and (ii), respectively. We also examined the effect of adjusting for age, sex and presence of at least one APOEε4 allele (using a dominant genetic inheritance model) in the Spanish replication sample and here again the results were attenuated (OR 1.24, CI 1.02×1.51, p=0.03). These findings are consistent with the moderate to low level of linkage disequilibrium observed between rs597668 and SNPs within the *APOE* and *TOMM40* region ($r^2<0.01$ according to HapMap CEU data, see also Figure 3).

Replication of Previously Reported Associations in CHARGE sample

In our secondary analyses examining replication of published findings in the previously unreported CHARGE data, 6 intronic or 3' UTR SNPs in the *APOE/TOMM40/APOC1* region (rs6857, rs2075650, rs4420638, rs157582, rs6859 and rs10119) reached a genome-wide significance threshold of $<1.7\times10^{-8}$, and we replicated the top SNPs within two out of the three recently reported genetic loci associated with AD in prior GWAS: *CLU* (rs11136000, OR 0.90, CI 0.82–0.98, p=0.02) and *PICALM* (rs3851179, OR 0.90, CI: 0.83–0.99, p=0.02); see eTable 1 and the Supplementary methods for additional details. We did not find a significant association with the top *CR1* SNP (rs3818361) in the CHARGE data. However 13 SNPs within the gene showed nominal significance ($0.001<p<0.05$), as shown in eTable 2. Further, adding CHARGE and TGEN data on rs3818361 to the previously reported EADI1 and GERAD1 data – Mayo data were here included in the GERAD1 data –

showed that results now reached genome-wide significance (OR 1.15, 1.11–1.20, $p=1.04\times10^{-11}$ (Supplemental Figure 9).

Among the 54 SNPs selected from the top 12 candidate genes (outside the *APOE/TOMM40/APOC1* locus) listed in the Alzgene website, we found evidence for a nominal association of rs4362 in the angiotensin conversting enzyme (*ACE*(RefSeq)) gene and rs1784933 in the sortilin-related receptor L(DLR class A) repeats-containing (*SORL1*(RefSeq)) gene with AD (risks associated with each copy of the minor allele were 0.92, CI 0.85–0.99, p=0.03, and 1.33, CI 1.03–1.72, p=0.03, respectively; eTable 3 in Supplementary material).

Genetic Risk Prediction

We assessed the extent to which *APOE*ε4, *PICALM* and *CLU* can improve predictive models for risk of incident AD in the general population. The addition of *APOE*ε*4* carrier status to a prediction model including age and sex only, increased the AUC from 0.826 (95%CI 0.806–0.846) to 0.847 (95%CI 0.828–0.865) in the Rotterdam study and from 0.670 (95%CI 0.625–0.723) to 0.702 (95%CI 0.654–0.754) in the CHS study. Further inclusion of risk allele status for *CLU* and *PICALM* improved the AUC only minimally to 0.849 (95%CI 0.831–0.867) in the Rotterdam Study and to 0.705 (95%CI 0.654–0.751) in CHS. The corresponding Receiver Operating Characteristic curves are shown in Supplemental Figure 10.

Comment

We report results of an international three-stage genome-wide analysis to study genetic variation underlying late-onset, sporadic AD. We studied over 35,000 persons (8371 AD cases), constituting the largest sample analyzed to date. In the gene discovery phase we showed genome-wide significance for two novel loci related to AD, one on chromosome 2 and a second locus on chromosome 19 that seems independent of *APOE*. We note that *BIN1* was previously identified as showing suggestive association with AD in the recent GWAS from the GERAD1;13 our study now finds the association for the first time to be genome-wide significant, which is a major step forward. Furthermore, we replicated both these loci as well as the recently identified loci, *CLU* and *PICALM* in an independent sample. Although genetic variation at the *CLU* and *PICALM* loci did modify the risk of AD in our population-based sample, and their discovery represents a significant advance in understanding the pathophysiology of AD, these polymorphisms had a very limited impact on prediction of AD risk.

The locus on chromosome 2q14.3 is adjacent to the bridging integrator 1 (*BIN1*) gene, which is one of two amphiphysins, and is expressed most abundantly in the brain and muscle.23 Amphiphysins promote caspase-independent apoptosis and also play a critical role in neuronal membrane organization and clathrin mediated, synaptic vessel formation,24 a process disrupted by Aβ.25 Knock-out mice with decreased expression of the amphiphysins have seizures and major learning deficits.26 Altered expression of *BIN1* has been demonstrated in aging mice, in transgenic mouse models of AD and in persons with schizophrenia.27, 28

The 19q13.3 locus (rs597668), a site distal to and not in linkage disequilibrium with SNPs in the *APOE* locus, had been suspected, in an early linkage study, to harbor a gene for AD.29 There are 6 genes adjacent to this locus, two of which are part of pathways linked to Alzheimer pathology. The protein product of *BLOC1S3*, called ‘Biogenesis of lysosomal organelles complex-1, subunit 3’ is expressed in the brain, regulates endosomal to lysosomal routing,30 and has been implicated in schizophrenia.31 The second gene, *MARK4* or MAP/microtubule affinity-regulating kinase 4, is inducible, expressed only in the brain, and plays

a role in neuronal differentiation.32 MARK4 is a kinase that phosphorylates tau, is polyubiquitinated in vivo, and is a substrate of the aging-related deubiquitinating enzyme USP9X; hence it may play a role in the abnormal tau phosphorylation seen in AD.33 Little is known of the function of the gene closest to rs597668, exocyst complex component 3-like 2 gene (*EXOC3L2*), also referred to as protein 7 transactivated by hepatitis B virus X antigen (*XTP7*) gene.

When evaluating the added value of the new AD genes in clinical risk prediction, we focused on the 2 recently reported AD genes13, 14 that were replicated in our population-based studies, *CLU* and *PICALM* and found that they minimally improved prediction of incident AD beyond age, sex and *APOE*ε4 based models; the increase in AUC was 0.002 in the Rotterdam Study and 0.003 in CHS. There are two reasons for this. First, the associations of *CLU* and *PICALM* with AD risk were markedly lower than those of age and *APOE*, and therefore a major improvement was not expected. This fits with recent insights on polygenic models that assume there are 10,000s of risk alleles, each with a small (~5% increase in relative risk) effect throughout the whole genome, rather than a discrete number of alleles with moderate effects. Such models appear to underlie the susceptibility to schizophrenia risk and a similar model may be applicable to AD.34 Second, the extent to which risk factors improve risk prediction depends on the predictive performance of the initial risk model. Added risk factors need to have stronger effects to improve a risk model with high AUC than to improve a model with lower AUC. AD risk prediction based on age, sex and *APOE* already has very high discriminative accuracy, the AUC was 0.826 in the Rotterdam Study and 0.670 in CHS, which implies that further improvements require many new variants or variants with strong effects. Whether such improvements are to be expected will depend for a large part on our ability to unravel the underlying genetic architecture and to identify and quantify environmental risk factors, including complex interactions.35 The obvious next step for genetic research in AD will be to further increase the sample size of GWAs and evaluate further genetic models.

Strengths of this study include the large sample of clinic and community-based cases and controls and the subsample of prospectively ascertained incident AD that permitted the exploration of incident risk prediction algorithms. The observed associations are unlikely to be due to population stratification since the discovery and replication samples were restricted to whites of European origin and were also investigated for latent population substructure.

The study also has limitations. Despite our large sample size, we had limited power to detect associations with small effect sizes and associations with rare variants. While all studies used accepted clinical or pathological criteria to define dementia and AD, phenotypic heterogeneity between samples may have limited our ability to detect some associations. Moreover, the controls in the Spanish replication sample were younger than the cases and their cognitive status had not been formally examined. However, whereas this could reduce our power to observe an association, it would not invalidate the associations we did observe. Further, the frequency distribution of minor and major alleles among the Spanish controls was similar to that noted in the discovery sample and in the HapMap CEU sample.

In conclusion, this meta-analysis of GWAS data from several of the largest AD GWAS studies to date confirms previously known and recently described associations (*CLU* and *PICALM*) and shows genome-wide significance and replication for two biologically plausible, novel loci on chromosomes 2 and 19.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The following authors contributed equally as first authors: Sudha Seshadri, Annette L. Fitzpatrick, M. Arfan Ikram, Anita L. DeStefano, Vilmundur Gudnason, Merce Boada.

The following authors contributed equally as last authors: Agustin Ruiz, Julie Williams, Philippe Amouyel, Steve G. Younkin, Philip A. Wolf, Lenore J. Launer, Oscar L. Lopez, Cornelia M. van Duijn, Monique M. B. Breteler.

References

1. Seshadri S, Wolf PA. Lifetime risk of stroke and dementia: current concepts, and estimates from the Framingham Study. Lancet Neurol. 2007; 6(12):1106–1114. [PubMed: 18031707]
2. Gatz M, Reynolds CA, Fratiglioni L, et al. Role of genes and environments for explaining Alzheimer disease. Arch Gen Psychiatry. 2006; 63(2):168–174. [PubMed: 16461860]
3. Ertekin-Taner N. Genetics of Alzheimer's disease: a centennial review. Neurol Clin. 2007; 25(3): 611–667. [PubMed: 17659183]
4. Reiman EM, Webster JA, Myers AJ, et al. GAB2 alleles modify Alzheimer's risk in APOE epsilon4 carriers. Neuron. 2007; 54(5):713–720. [PubMed: 17553421]
5. Carrasquillo MM, Zou F, Pankratz VS, et al. Genetic variation in PCDH11X is associated with susceptibility to late-onset Alzheimer's disease. Nat.Genet. 2009; 41(2):192–198. [PubMed: 19136949]
6. Poduslo SE, Huang R, Huang J, Smith S. Genome screen of late-onset Alzheimer's extended pedigrees identifies TRPC4AP by haplotype analysis. Am.J.Med.Genet B Neuropsychiatr.Genet. 2009; 150B(1):50–55. [PubMed: 18449908]
7. Beecham GW, Martin ER, Li YJ, et al. Genome-wide association study implicates a chromosome 12 risk locus for late-onset Alzheimer disease. Am.J.Hum.Genet. 2009; 84(1):35–43. [PubMed: 19118814]
8. Bertram L, Lange C, Mullin K, et al. Genome-wide association analysis reveals putative Alzheimer's disease susceptibility loci in addition to APOE. Am.J.Hum.Genet. 2008; 83(5):623–632. [PubMed: 18976728]
9. Myers A, Holmans P, Marshall H, et al. Susceptibility locus for Alzheimer's disease on chromosome 10. Science. 2000; 290(5500):2304–2305. [PubMed: 11125144]
10. Li H, Wetten S, Li L, et al. Candidate single-nucleotide polymorphisms from a genomewide association study of Alzheimer disease. Arch.Neurol. 2008; 65(1):45–53. [PubMed: 17998437]
11. Coon KD, Myers AJ, Craig DW, et al. A high-density whole-genome association study reveals that APOE is the major susceptibility gene for sporadic late-onset Alzheimer's disease. J Clin Psychiatry. 2007; 68(4):613–618. [PubMed: 17474819]
12. Feulner TM, Laws SM, Friedrich P, et al. Examination of the current top candidate genes for AD in a genome-wide association study. Mol.Psychiatry. 2009
13. Harold D, Abraham R, Hollingworth P, et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nat.Genet. 2009; 41(10):1088–1093. [PubMed: 19734902]
14. Lambert JC, Heath S, Even G, et al. Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. Nat.Genet. 2009; 41(10):1094–1099. [PubMed: 19734903]
15. Psaty BM, O'Donnell CJ, Gudnason V, et al. Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: Design of prospective meta-analyses of genome-wide association studies from five cohorts. Circ Cardiovasc Genet. 2009; 2:73–80. [PubMed: 20031568]
16. Association, AP. Diagnostic and statistical manual of mental disorders (DSM-IV). Washington, D.C: American Psychiatric Association; 1994.

17. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology. 1984; 34(7):939–944. [PubMed: 6610841]

18. Pe'er I, Yelensky R, Altshuler D, Daly MJ. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. Genet Epidemiol. 2008; 32(4): 381–385. [PubMed: 18348202]

19. Antunez C, Boada M, Lopez-Arrieta J, et al. GOLPH2 Gene Markers are Not Associated with Alzheimer's Disease in a Sample of the Spanish Population. J Alzheimers Dis. 2009 ahead of print.

20. Ramirez-Lorca R, Boada M, Saez ME, et al. GAB2 gene does not modify the risk of Alzheimer's disease in Spanish APOE 4 carriers. J Nutr Health Aging. 2009; 13(3):214–219. [PubMed: 19262956]

21. Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE. Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. Nat Genet. 2007; 39(1):17–23. [PubMed: 17192785]

22. Janssens AC, van Duijn CM. Genome-based prediction of common diseases: methodological considerations for future research. Genome Med. 2009; 1(2):20. [PubMed: 19341491]

23. Wechsler-Reya R, Sakamuro D, Zhang J, Duhadaway J, Prendergast GC. Structural analysis of the human BIN1 gene. Evidence for tissue-specific transcriptional regulation and alternate RNA splicing. J.Biol.Chem. 1997; 272(50):31453–31458. [PubMed: 9395479]

24. Wigge P, Kohler K, Vallis Y, et al. Amphiphysin heterodimers: potential role in clathrin-mediated endocytosis. Mol.Biol.Cell. 1997; 8(10):2003–2015. [PubMed: 9348539]

25. Kelly BL, Ferreira A. Beta-amyloid disrupted synaptic vesicle endocytosis in cultured hippocampal neurons. Neuroscience. 2007; 147(1):60–70. [PubMed: 17499934]

26. Di PG, Sankaranarayanan S, Wenk MR, et al. Decreased synaptic vesicle recycling efficiency and cognitive deficits in amphiphysin 1 knockout mice. Neuron. 2002; 33(5):789–804. [PubMed: 11879655]

27. English JA, Dicker P, Focking M, Dunn MJ, Cotter DR. 2-D DIGE analysis implicates cytoskeletal abnormalities in psychiatric disease. Proteomics. 2009; 9(12):3368–3382. [PubMed: 19562803]

28. Yang S, Liu T, Li S, et al. Comparative proteomic analysis of brains of naturally aging mice. Neuroscience. 2008; 154(3):1107–1120. [PubMed: 18495355]

29. Poduslo SE, Yin X. A new locus on chromosome 19 linked with late-onset Alzheimer's disease. Neuroreport. 2001; 12(17):3759–3761. [PubMed: 11726789]

30. Starcevic M, Dell'Angelica EC. Identification of snapin and three novel proteins (BLOS1, BLOS2, and BLOS3/reduced pigmentation) as subunits of biogenesis of lysosome-related organelles complex-1 (BLOC-1). J.Biol.Chem. 2004; 279(27):28393–28401. [PubMed: 15102850]

31. Morris DW, Murphy K, Kenny N, et al. Dysbindin (DTNBP1) and the biogenesis of lysosome-related organelles complex 1 (BLOC-1): main and epistatic gene effects are potential contributors to schizophrenia susceptibility. Biol.Psychiatry. 2008; 63(1):24–31. [PubMed: 17618940]

32. Moroni RF, De BS, Colapietro P, Larizza L, Beghini A. Distinct expression pattern of microtubule-associated protein/microtubule affinity-regulating kinase 4 in differentiated neurons. Neuroscience. 2006; 143(1):83–94. [PubMed: 16973293]

33. Trinczek B, Brajenovic M, Ebneth A, Drewes G. MARK4 is a novel microtubule-associated proteins/microtubule affinity-regulating kinase that binds to the cellular microtubule network and to centrosomes. J.Biol.Chem. 2004; 279(7):5915–5923. [PubMed: 14594945]

34. Purcell SM, Wray NR, Stone JL, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature. 2009; 460(7256):748–752. [PubMed: 19571811]

35. Janssens AC, van Duijn CM. Genome-based prediction of common diseases: advances and prospects. Hum Mol Genet. 2008; 17(R2):R166–R173. [PubMed: 18852206]

Figure 1.
Figure showing the three-stage approach and the various studies included in the different stages.

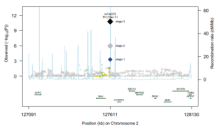

Figure 2.

Regional association plot for novel loci that were significantly associated ($p<5\times10^{-8}$) with AD in stage 3 analyses (rs744373 near *BIN1*, rs597668 near *BLOC1S3* and *MARK4*). Each data marker represents the statistical significance (p-value) of each SNP plotted on the $-\log_{10}$ scale against its chromosomal position (NCBI build 36).The blue diamonds show stage 1 p-values for the sentinel (top) SNP at each locus, whereas the grey and black diamonds show the p-values for this same SNP following stage 2 and stage 3 meta-analyses, respectively. P-values from stage 1 for additional SNPs at that locus are color- and size-coded according to the strength of their linkage disequilibrium with the top SNP as follows: $r^2<0.2$ white; $0.2<r^2<0.5$ yellow; $0.5<r^2<0.8$ orange; $r^2>0.8$ red. The fine scale recombination rate is shown by the blue line which shows the average frequency with which recombination occurs (exchange of genetic material between maternal and paternal chromosomes during meiosis) at that site. Genes located in the region shown (on either strand of the chromosome) are shown as green lines with Human Genome Organization (HUGO) gene nomenclature committee gene symbols, the length of the green line represents the size/extent of the gene and the arrow the direction in which transcription of mRNA occurs.

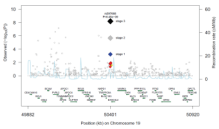

Figure 3.

Regional association plot for novel loci that were significantly associated ($p<5\times10^{-8}$) with AD in stage 3 analyses (rs744373 near *BIN1*, rs597668 near *BLOC1S3* and *MARK4*). Each data marker represents the statistical significance (p-value) of each SNP plotted on the $-\log_{10}$ scale against its chromosomal position (NCBI build 36).The blue diamonds show stage 1 p-values for the sentinel (top) SNP at each locus, whereas the grey and black diamonds show the p-values for this same SNP following stage 2 and stage 3 meta-analyses, respectively. P-values from stage 1 for additional SNPs at that locus are color- and size-coded according to the strength of their linkage disequilibrium with the top SNP as follows: $r^2<0.2$ white; $0.2<r^2<0.5$ yellow; $0.5<r^2<0.8$ orange; $r^2>0.8$ red. The fine scale recombination rate is shown by the blue line which shows the average frequency with which recombination occurs (exchange of genetic material between maternal and paternal chromosomes during meiosis) at that site. Genes located in the region shown (on either strand of the chromosome) are shown as green lines with Human Genome Organization (HUGO) gene nomenclature committee gene symbols, the length of the green line represents the size/extent of the gene and the arrow the direction in which transcription of mRNA occurs.

Table 1

Characteristics of studies in stage 1 of the analysis.

	CHS		FHS		Rotterdam		AGES		TGEN		Mayo	
Study design	Cohort		Cohort		Cohort		Cohort		Case-control		Case-control	
Genotype platform	Illumina HumanCNV370-Duo®		Affymetrix GeneChip® Human Mapping 500K Array Set + 50K Gene Focused Panel®		Illumina Infinium HumanHap550-chip v3.0®		Illumina HumanCNV370-Duo®		Affymetrix GeneChip® Human Mapping 500K Array Set,		Illumina Human-Hap300v2-Duo BeadChips	
Prevalence studies	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls
N*	93	2429	52	2091	171	5700	78	2684	829	536	810	1202
Women (%)	49 (53)	1506 (62)	42 (81)	1192 (57)	128 (75)	3347 (59)	39 (50)	1557 (58)	431 (52)	338 (63)	462 (57)	601 (50)
Age	80±6	75±5	87±6	76±7	84±9	69±9	81±5	76±5	81±10	80±7	73±4	74±5
APOE e4 +ve (%) †	35 (38)	583 (24)	20 (38)	418 (20)	62 (36)	1549 (28)	38 (49)	725 (27)	481 (58)	107 (20)	535 (66)	337 (28)
Incidence studies												
Cohort at risk*	2429		806		5700		-		-		-	
Women, %	1506 (62)		484 (60)		3347 (59)		-		-		-	
Ages at start (and at incident dementia)	75±5 (82±5)		82±6 (88±5)		69±9 (82±7)		-		-		-	
Incident AD cases	435		76		462		-		-		-	
Mean follow-up (years)	6.8±3.6		4.8±3.0		9.3±3.2		-		-		-	
APOE e4 +ve, %†	632 (26)		153 (19)		1549 (28)		-		-		-	

In the prevalence studies, cases were those persons who suffered from AD at time of DNA draw. Controls were those that were free of any dementia. In the incidence studies, cases were those persons from the cohort at risk who developed dementia during the follow-up. Persons who developed another type of dementia were censored at the date of onset.

Data are means (SD), unless otherwise indicated. AGES=Age, Gene/Environment Susceptibility-Reykjavik Study, CHS=Cardiovascular Health Study, FHS=Framingham Heart Study, TGEN=Translational Genomics Research Institute.

*Includes only those genotyped persons who also provided consent for these analyses and had high-quality genotyping (met QC-criteria), details are in the Supplement. In the FHS only Original cohort participants were included in incident analyses.

†Among those with *APOE* genotyping available

Table 2

Genetic loci at which SNPs are associated with AD at $p<10^{-5}$ in the stage 2 meta-analysis, and which were further meta-analyzed in stage 3.

Top SNP[*]	*Chr:Position*	*Additional SNPs*[**]	*Nearest Gene*[†]	*Minor Allele*[††]	*MAF*	*Stage 1 meta-analysis*		*Stage 2 meta-analysis*		*Stage 3 meta-analysis*	
						Meta odds ratio[§]	Meta pvalue	Meta odds ratio[§]	Meta pvalue	Meta odds ratio[§]	Meta pvalue
rs2075650	19:50087459	18	*APOE (RefSeq NG_007084)*	G	13.7	2.23 (2.04–2.44)	3.18×10^{-68}	2.61 (2.45–2.80)	4.67×10^{-172}	2.53 (2.41–2.66)	1.04×10^{-295}
rs11136000	8:27520436		*CLU (GenBank AY341244)*	T	39.2	0.89 (0.83–0.94)	4.98×10^{-4}	0.85 (0.81–0.90)	1.49×10^{-9}	0.85 (0.82–0.88)	1.62×10^{-16}
rs3851179	11:85546288		*PICALM (GenBank BC073961)*	T	37.1	0.86 (0.81–0.92)	1.22×10^{-5}	0.89 (0.84–0.93)	2.81×10^{-6}	0.87 (0.84–0.91)	3.16×10^{-12}
rs744373	2:127611085		*BIN1 (RefSeq NG_012042)*	G	29.1	1.13 (1.06–1.21)	4.93×10^{-4}	1.14 (1.08–1.20)	1.02×10^{-6}	1.15 (1.11–1.20)	1.59×10^{-11}
rs597668	19:50400728	1	*EXOC3L2 (RefSeq NM_138568)*	C	15.4	1.18 (1.07–1.29)	5.91×10^{-4}	1.18 (1.10–1.26)	2.16×10^{-6}	1.17 (1.11–1.23)	6.45×10^{-9}
rs11771145	7:142820884		*EPHA1 (GenBank AH007960)*	A	34.7	0.87 (0.81–0.94)	2.14×10^{-4}	0.86 (0.81–0.90)	1.32×10^{-8}	0.91 (0.87–0.94)	1.70×10^{-6}
rs2043948	14:74142801		*LTBP2 (RefSeq NM_000428)*	T	7.7	1.25 (1.10–1.42)	6.96×10^{-4}	1.27 (1.16–1.39)	4.44×10^{-7}	1.13 (1.06–1.22)	4.46×10^{-4}
rs2825544	21:19662423		*PRSS7 (RefSeq NG_012207)*	C	34.6	113 (1.06–1.21)	2.55×10^{-4}	1.14 (1.08–1.20)	4.85×10^{-7}	1.09 (1.05–1.13)	2.10×10^{-5}
rs7527934	1:14231011	9	*PRDM2 (RefSeq NM_012231)*	G	25.7	0.86 (0.79–0.93)	3.50×10^{-4}	0.87 (0.82–0.92)	5.87×10^{-6}	*0.97 (0.91–1.03)*	-
rs4296166	14:32022118		*AKAP6 (RefSeq NM_004274)*	A	47.8	1.14 (1.07–1.21)	8.36×10^{-5}	1.12 (1.07–1.18)	4.08×10^{-6}	*0.98 (0.89–1.08)*	-

MAF=Minor allele frequency

[*] At each locus, the SNP with lowest p-value was selected for stage 3 meta-analysis.

[**] Number of additional SNPs at the locus with $p<10^{-5}$

[†] Column shows the Human Gene Organization (HUGO) Gene Nomenclature System symbols for the gene located closest to each SNP. Standardized gene annotations for all SNP results were derived programmatically from the UCSC Genome Browser RefSeq gene track (hg18).

[††] Alleles were coded on the forward strand of the genome.

[§] The minor allele was taken as coded allele. The odds-ratios represent the relative increase of disease risk per increase of one copy of the minor allele.

Table 3

Replication of genome-wide significant results from discovery sample in an independent Spanish (Fundació ACE) sample.

Gene	SNP	MAF (cases/controls)	OR	95% CI	P value
CLU	rs11136000	0.36/0.39	0.82	0.77–0.99	0.03
PICALM	rs3851179	0.30/0.34	0.84	0.74–0.95	0.007
BIN1	rs744373	0.30/0.27	1.17	1.03–1.33	0.02
EXOC3L2	rs597668	0.13/0.11	1.26	1.05–1.51	0.01

JAMA. Author manuscript; available in PMC 2011 May 12.