Common variants in ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease

Paul Hollingworth1,109, Denise Harold1,109, Rebecca Sims1,109, Amy Gerrish1,109, Jean-Charles Lambert2,3,4,109, Minerva M Carraquillo5,109, Richard Abraham1, Marian L Hamshere1, Jaspreet Singh Pahwa1, Valentina Moskvin1, Kimberley Dowzell1, Nicola Jones1, Alexandra Stretton1, Charlene Thomas1, Alex Richards1, Dobril Ivanov1, Caroline Widdowson1, Jade Chapman1, Simon Lovestone6,7, John Powell7, Petroula Proitsi1, Michelle K Lupton7, Carol Brayne8, David C Rubinsztain9, Michael Gill10, Brian Lawlor10, Aoibhinn Lynch10, Kristelle S Brown11, Peter A Passmore12, David Craig12, Bernadette McGuinness12, Stephen Todd12, Clive Holmes13, David Mann14, A David Smith15, Helen Beaumont15, Donald Wardens5, Gordon Wilcock16, Seth Love17, Patrick G Kehoe17, Nigel M Hooper18, Emma R. L. C. Vardy14,18,19, John Hardy20,21, Simon Mead22, Nick C Fox22, Martin Rossor22, John Collinge22, Wolfgang Maier23,24, Frank Jessen23, Britta Schürmann23,26, Eckart Rüther24,25,26, Reiner Heun23,27, Heike Kölsch23, Hendrik van den Bussche28, Isabella Heuser29, Johannes Kornhuber30, Jens Willfang31, Martin Dichgans32,33, Lutz Fröhlich34, Harald Hampel35, Michael Hül36, John Gallagher36, Dan Rujescu35, Ina Giegling35, Alison M Goate37,38,39, John S K Kauwe40, Carlos Cruchaga37, Petra Nowotny37, John C Morris38, Kevin Mayo37, Kristel Sleegers41,42, Karolien Betbens41,42, Sebastiaan Engelborghs41,43, Peter P Deyn41,43, Christine Van Broeckhoven41,42, Gill Livingstone41,44, Nicholas J Bass44, Hugh Gurling44, Andrew McQuillin44, Rhian Gwilliam45, Panagiotis Deloukas45, Ammar Al-Chalabi46, Christopher E Shaw46, Magda Tsalaki47, Andrew B Singleton48, Rita Guerreiro48, Thomas W Mühleisen49,50, Markus M Nöthen48,50, Susanne Moebus51, Karl-Heinz Jöckel51, Norman Klop52, H-Erich Wichmann52,53,54, V Shane Pankratz55, Sigrid B Sando56,57, Jan O Aasly56,57, Maria Barckikos58, Zbigiew K Wszolek59, Dennis W Dickson5, Neill R Graff-Radford5,59, Ronald C Petersen60,61, the Alzheimer's Disease Neuroimaging Initiative62, Cornelia M van Duijn63,64, Monique MB Breteler63,64, M Arfan Ikram65,66, Anaït L

Contributors.1,109 These authors contributed equally to this work.

Author Contributions J. Williams and P.H. conceived the study and designed the manuscript. J. Williams, D. Harold, P.H., and D.H. organized the manuscript. M.O.D. was responsible for drafting the manuscript. E. C. Vardy, S. Moebus, and J. S. K. Kauwe provided the manuscript. J. Williams, D. Harold, and P.H. took primary responsibility for drafting the manuscript assisted by R.S., A.G., R.A., M.O. and D.H. All authors discussed the results and approved the manuscript.

Competing financial interests The authors have applied for a patent based on the results of this research.

Published in final edited form as: Nat Genet. 2011 May ; 43(5): 429–435. doi:10.1038/ng.803.
1Medical Research Council (MRC) Centre for Neuropsychiatric Genetics and Genomics, Department of Psychology and Clinical Neurosciences, School of Medicine, University of Cambridge, Cambridge, UK. 2Inserm U744, F-59019 Lille, France. 3Institut Pasteur de Lille, F-59019, Lille, France. 4Université de Lille Nord de France, F-59000 Lille, France. 5Department of Neuroscience, Mayo Clinic College of Medicine, Jacksonville, Florida, USA. 6National Institute for Health Research Biomedical Research Centre for Mental Health at the South London and Maudsley National Health Service Foundation Trust and Institute of Psychiatry, Kings College, London, UK. 7Department of Neuroscience, Institute of Psychiatry, Kings College, London, UK. 8Institute of Public Health, University of Cambridge, Cambridge, UK. 9Cambridge Institute for Medical Research, University of Cambridge, Cambridge, UK. 10Mercer’s Institute for Research on Aging, St. James Hospital and Trinity College, Dublin, Ireland. 11Institute of Genetics, Queen’s Medical Centre, University of Nottingham, Nottingham, UK. 12Ageing Group, Centre for Public Health, School of Medicine, Dentistry and Biomedical Sciences, Queen’s University Belfast, UK. 13Division of Clinical Neurosciences, School of Medicine, University of Southampton, Southampton, UK. 14Neurodegeneration and Mental Health Research Group, School of Community Based Medicine, University of Manchester, Hope Hospital, Stott Lane, Salford, Manchester, UK. 15Oxford Project to Investigate Memory and Ageing (OPTIMA), University of Oxford, John Radcliffe Hospital, Oxford, UK. 16Nuffield Department of Clinical Neurosciences, University of Oxford, Headington, Oxford, UK. 17Dementia Research Group, University of Bristol Institute of Clinical Neurosciences, Frenchay Hospital, Bristol, UK. 18Institute of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, LS2 9JT, UK. 19Cerebral Function Unit, Salford Royal NHS Trust, Stott Lane, Salford, M6 8HD, UK. 20Department of Molecular Neuroscience, Institute of Neurology, University College London, London, UK. 21Reta Lila Weston Laboratories, Institute of Neurology, London, UK. 22Department of Neurodegenerative Disease, UCL Institute of Neurology, London, UK. 23Department of Psychiatry, University of Bonn, Bonn, Germany. 24German Centre for Neurodegenerative Diseases, Bonn, Bonn, Germany. 25Institute for Molecular Psychiatry, University of Bonn, Bonn, Germany. 26Department of Psychiatry, University of Göttingen, Germany. 27Department of Psychiatry, Royal Derby Hospital, Derby, DE2 2WQ, UK. 28Institute of Primary Medical Care, University Medical Center Hamburg-Eppendorf, Germany. 29Department of Psychiatry, Charité Berlin, Berlin, Germany. 30Department of Psychiatry and Psychotherapy, University of Erlangen-Nuremberg, Germany. 31Landschaftsverband Rheinland-Hospital Essen, Department of Psychiatry and Psychotherapy, University Duisburg-Essen, Essen, Germany. 32Department of Neurology, Klinikum der Universität München, Munich,
Germany. 33Institute for Stroke and Dementia Research, Klinikum der Universität München, Munich, Germany. 34Department of Geriatric Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany. 35Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, Johann Wolfgang Goethe-University, Frankfurt, Germany. 36Department of Primary Care and Public Health, School of Medicine, Cardiff University, Cardiff, UK. 37Department of Psychiatry, Washington University School of Medicine, St Louis, Missouri, USA. 38Department of Neurology, Washington University School of Medicine, St Louis, Missouri, USA. 39Department of Genetics, Washington University School of Medicine, St Louis, Missouri, USA. 40Department of Biology, Brigham Young University, Provo, Utah, USA. 41Institute Born-Bunge and University of Antwerp, Antwerp, Belgium. 42Neurodegenerative Brain Diseases group, Department of Molecular Genetics, VIB, Antwerp, Belgium. 43Memory Clinic and Department of Neurology, ZNA Middelheim, Antwerp, Belgium. 44Department of Mental Health Sciences, University College London, UK. 45The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK. 46MRC Centre for Neurodegeneration Research, Department of Clinical Neuroscience, King’s College London, Institute of Psychiatry, London, UK. 47Third Department of Neurology, Aristotle University of Thessaloniki, Thessaloniki, Greece. 48Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, Maryland, USA. 49Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany. 50Institute of Human Genetics, University of Bonn, Bonn, Germany. 51Institute for Medical Informatics, Biometry and Epidemiology, University Hospital of Essen, University Duisburg-Essen, Essen, Germany. 52Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany. 53Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany. 54Klinikum Grosshadern, Munich, Germany. 55Division of Biomedical Statistics and Informatics, Mayo Clinic and Mayo Foundation, Rochester, Minnesota, USA. 56Department of Neurology, St. Olav’s Hospital, Edvard Griegs Gate 8, 7006 Trondheim, Norway. 57Department of Neuroscience, Norwegian University of Science and Technology, NTNU, 7491 Trondheim, Norway. 58Department of Neurodegenerative Disorders, Medical Research Centre, Polish Academy of Sciences, Warsaw, Poland. 59Department of Neurology, Mayo Clinic College of Medicine, Jacksonville, FL 32224, USA. 60Department of Neurology, Mayo Clinic College of Medicine, Rochester, MN 55905, USA. 61Mayo Alzheimer Disease Research Center, Mayo Clinic College of Medicine, Rochester, MN 55905, USA. 62Data used in the preparation of this article were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. ADNI investigators include (complete listing available at www.loni.ucla.edu/ADNI/Collaboration/ADNI_Authorship_list.pdf). 63Department of Epidemiology, Erasmus MC University Medical Center, Rotterdam, The Netherlands. 64Netherlands Consortium for Healthy Aging, The Netherlands. 65Departments of Neurology and Biostatistics, Boston University School of Medicine, Boston, Massachusetts, USA. 66The National Heart Lung and Blood Institute’s Framingham Heart Study, Framingham, Massachusetts, USA. 67Department of Epidemiology, University of Washington, Seattle, Washington, USA. 68Department of Neurology, The Alzheimer’s Disease Research Center, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA. 69Department of Psychiatry, The Alzheimer’s Disease Research Center, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA. 70Neuroepidemiology Section, Laboratory of Epidemiology, Demography and Biometry (LJL), National Institute on Aging, Washington DC, USA. 71Department of Neurology, Boston University School of Medicine, Boston, Massachusetts, USA. 72Inserm U888, Hôpital La Colombière, Montpellier, France. 73Inserm U614, Faculté de Médecine-Pharmacie de Rouen, Rouen, France. 74UMR 894, Inserm Faculté de Médecine, Université Paris Descartes, Paris, France. 75Inserm U897, Victor Segalen University, Bordeaux, France. 76Inserm U708, Paris, France. 77Centre
Abstract

We sought to identify new susceptibility loci for Alzheimer’s disease (AD) through a staged association study (GERAD+) and by testing suggestive loci reported by the Alzheimer’s Disease Genetic Consortium (ADGC). First, we undertook a combined analysis of four genome-wide association datasets (Stage 1) and identified 10 novel variants with $P \leq 1 \times 10^{-5}$. These were tested for association in an independent sample (Stage 2). Three SNPs at two loci replicated and showed evidence for association in a further sample (Stage 3). Meta-analyses of all data provide compelling evidence that $ABCA7$ (meta-$P=4.5 \times 10^{-17}$; including ADGC meta-$P=5.0 \times 10^{-21}$) and the $MS4A$ gene cluster (rs610932, meta-$P=1.8 \times 10^{-14}$; including ADGC meta-$P=1.2 \times 10^{-16}$; rs670139, meta-$P=1.4 \times 10^{-9}$; including ADGC meta-$P=1.1 \times 10^{-10}$) are novel susceptibility loci for AD. Second, we observed independent evidence for association for three suggestive loci reported by the ADGC GWAS, which when combined shows genome-wide significance: $CD2AP$ (GERAD+ $P=8.0 \times 10^{-4}$; including ADGC meta-$P=8.6 \times 10^{-9}$), $CD33$ (GERAD+ $P=2.2 \times 10^{-4}$; including ADGC meta-$P=1.6 \times 10^{-9}$) and $EPHA1$ (GERAD+ $P=3.4 \times 10^{-4}$; including ADGC meta-$P=6.0 \times 10^{-10}$). These findings support five novel susceptibility genes for AD.
Alzheimer’s disease (AD) is the most common form of dementia, with both environmental and genetic factors contributing to risk. AD is genetically complex and shows heritability up to 79%\(^1\). Rare variants in three genes (APP, PSEN1 & PSEN2)\(^2\) cause disease in a minority of cases, but until recently the Apolipoprotein E gene (APOE), was the only gene known to increase disease risk for the common form of AD with late-onset. In 2009 we published a genome-wide association study (GWAS) of AD in a sample designated GERAD1 (Genetic and Environmental Risk in AD Consortium), which identified two new genome-wide significant susceptibility loci: clusterin (CLU: \(P=8.5\times10^{-10}\)) and phosphatidylinositol-binding clathrin assembly protein gene (PICALM: \(P=1.3\times10^{-9}\)). We also observed more variants with \(P\) values \(<1\times10^{-5}\) than were expected by chance (\(P=7.5\times10^{-5}\)). These included variants in the complement receptor 1 (CRI) gene, the bridging integrator 1 (BINI) gene and the membrane-spanning 4A gene cluster (MS4A gene cluster). A second independent AD GWAS by Lambert and colleagues\(^4\) using the EADI1 sample (European Alzheimer’s Disease Initiative) showed genome-wide significant evidence for association with CLU (\(P=7.5\times10^{-9}\)) and CRI (\(P=3.7\times10^{-9}\)), and support for PICALM (\(P=3\times10^{-3}\)).

Combined analysis of the GERAD1 and EADI1 data yield highly significant support for all three loci (CLU meta-\(P=6.7\times10^{-10}\), PICALM meta-\(P=6.3\times10^{-12}\), CRI meta-\(P=3.2\times10^{-12}\)). The associations in CLU, PICALM and CRI have since been replicated in several independent datasets\(^5\)–\(^8\), shown trends in another\(^9\) and relationships with neurodegenerative processes underlying disease\(^10\). In addition, members of this consortium have since reported genome-wide significant association for BIN1 (\(P=1.6\times10^{-11}\)) and support for ephrin receptor A1 (EPHA1; \(P=1.7\times10^{-6}\)).. This study sought to identify new common susceptibility variants for AD by first undertaking a three-stage association study based upon predominantly European samples (GERAD+, see Figure 1) and second, by testing these samples for loci showing suggestive evidence for association in the American Alzheimer’s Disease Genetics Consortium (ADGC) GWAS\(^12\).

The first stage of this study comprised a meta-analysis of four AD GWAS datasets (6688 cases, 13685 controls), including: GERAD1, EADI1, Translational Genomics Research Institute (TGEN1) and Alzheimer’s Disease Neuroimaging Initiative (ADNI). Single nucleotide polymorphisms (SNPs) which remained significant at \(P\leq1\times10^{-5}\) were then tested for replication in the second stage of this study, comprising 4896 cases and 4903 controls including genotyping of the GERAD2 sample and in silico replication in the deCODE and German Alzheimer’s disease Integrated Genome Research Network (AD-IG) GWAS datasets. In Stage 3, novel SNPs showing significant evidence of replication in Stage 2 were then tested for association in a sample comprising 8286 cases and 21258 controls, which included new genotyping in the EADI2 and Mayo2 samples, and in silico replication in the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) sample\(^11\). Sample descriptions and characteristics can be found in the Supplementary Note and Supplementary Table 1.

In Stage 1 we identified 61 SNPs associated with AD at \(P\leq1\times10^{-5}\) following meta-analysis of 496763 SNPs in the GERAD1, TGEN1, ADNI and EADI1 (see Supplementary Table 2 and the Supplementary Note). Ten SNPs at novel loci and two at previously identified susceptibility loci that surpassed the \(P\leq1\times10^{-5}\) threshold, were selected for further analysis (see below). One SNP, rs610932 (Stage 1 \(P=1.8\times10^{-8}\)) at the MS4A (membrane spanning 4A) gene cluster, surpassed the threshold (\(P=5.0\times10^{-8}\)) for genome-wide significance. We also observed strong evidence for association at ABCA7 (ATP-binding cassette, sub-family A, member 7; rs3764650, Stage 1 \(P=2.6\times10^{-7}\)).
When selecting SNPs for testing in Stage 2, we excluded known susceptibility loci that had previously been tested in GERAD2 and limited analysis of BIN1 and CR1, which had not been tested in GERAD2, to the most significant SNPs at each locus (See Supplementary Table 2). Following pruning for linkage disequilibrium, twelve SNPs were taken forward for replication in Stage 2 (10 excluding BIN1 and CR1).

Five of the twelve SNPs tested in Stage 2 showed significant evidence for replication using a Bonferroni adjusted threshold for significance of $P=4.2 \times 10^{-3}$ (see Table 1 and Supplementary Table 3). In addition to SNPs at BIN1 and CR1, one SNP within ABCA7 (rs3764650, Stage 2 $P=1.9 \times 10^{-5}$) and two SNPs at the MS4A gene cluster (rs610932, stage 2 $P=1.6 \times 10^{-3}$; rs670139 Stage 2 $P=1.1 \times 10^{-3}$) showed evidence of replication in Stage 2. The three SNPs implicating novel risk loci were tested for association in the Stage 3 sample and showed further evidence of replication (rs3764650, Stage 3 $P=2.9 \times 10^{-7}$; rs610932, Stage 3 $P=2.1 \times 10^{-5}$; rs670139, Stage 3 $P=3.2 \times 10^{-3}$; see Table 1 and Supplementary Table 3).

We conducted an inverse variance weighted meta-analysis of data from Stages 1, 2 and 3 (See Table 1 and Supplementary Table 3). This provided strong evidence for association with rs3764650 at ABCA7 (meta-$P=4.5 \times 10^{-17}$) and two SNPs at the MS4A gene cluster: rs610932 (meta-$P=1.8 \times 10^{-10}$) and rs670139 (meta-$P=1.4 \times 10^{-5}$). When combining GERAD+ and ADGC results (after removing overlapping samples) ABCA7 has a $P$-value of $5.0 \times 10^{-21}$ (OR=1.22). The two SNPs at the MS4A gene cluster, rs610932 and rs670139, showed $P$-values of $1.2 \times 10^{-16}$ (OR=0.91) and $1.1 \times 10^{-10}$ (OR=1.08), respectively, in the combined analysis of GERAD+ and ADGC results. It is noteworthy that the most significant ADGC SNP at the MS4A locus is in LD with our top SNP (rs4938933 with rs610932 $r^2=0.62, D'=0.86$), thus both datasets may be detecting the same underlying signal.

This study also provides additional independent support for association with CR1 (Stage 2 $P=1.4 \times 10^{-5}$) and BIN1 (Stage 2 $P=3.8 \times 10^{-5}$; see Table 1 for meta-analysis.) We did not observe interaction between APOE and the novel variants identified in this study, indeed we did not find evidence of epistasis between any of the genome-wide significant variants identified to date (ABCA7, MS4A, BIN1, CR1, PICALM, CLU or APOE) (see Supplementary Table 4a). Likewise, adjusting for the presence of at least one APOE ε4 allele had little effect on the results of analysis of the three novel variants (see Supplementary Table 4b). We also found no evidence for association between these loci and age at onset of AD (rs3764650: $P=0.17$; rs670139: $P=0.38$; rs610932: $P=0.95$; rs744373: $P=0.87$; rs3818361: $P=0.58$).

This study therefore shows strong statistical support for two novel AD risk loci, which replicate over a number of independent case-control samples. The first of these is the ATP-binding cassette, sub-family A, member 7 (ABCA7) locus (Figure 2A). The associated marker is rs3764650, which is located in intron 13. This SNP was the only variant in the gene that passed our Stage 1 criterion, which is not unexpected given the low levels of linkage disequilibrium (LD) between this SNP and others included in the GWAS. However, in a preliminary attempt to identify an associated functional variant at the ABCA7 locus, we genotyped the GERAD2 sample for rs3752246, a non-synonymous SNP in exon 32 of the gene, which showed the highest LD with rs3764650 out of all HapMap ABCA7 coding variants based on $r^2$ ($r^2=0.36, D'=0.89$). This variant (which was not genotyped in Stage 1) was also associated with AD (GERAD2 $P=1\times10^{-5}$, OR=1.17). Rs3752246 encodes a glycine to alanine substitution at position 1527 of the protein (accession number NP_061985.2) which is predicted to be a benign change, and is unlikely to be the relevant functional variant. We used data from two published expression quantitative trait loci (eQTL) datasets (derived from lymphoblastoid cell lines and brain) to determine if
rs3764650 is associated with the expression of ABCA7. However, no association was observed (see Supplementary Table 5). Further work will be required to identify the causal variant(s) at this locus.

Second, we implicate the membrane-spanning 4A (MS4A) gene cluster (Figure 2B). The association spans an LD block of 293 kb (chr11: 59,814,287-60,107,105) and includes 6 of 16 known genes comprising the membrane-spanning 4-domains, subfamily A (MS4A). These are MS4A2, MS4A3, MS4A4A, MS4A4E, MS4A6A and MS4A6E. The associated SNPs are found in the 3’ UTR of MS4A6A (rs610932) and the intergenic region between MS4A4E and MS4A6A (rs670139). rs610932 shows nominally significant association with expression levels of MS4A6A in cerebellum and temporal cortex (0.01 < P < 0.05; see Supplementary Table 5), but not in frontal cortex, pons, or lymphoblastoid cell lines. The non-synonymous SNP that is most strongly associated with the genome-wide significant variants is rs2304933. This SNP was analyzed in Stage 1 but showed weaker evidence for association (P = 0.006) than the genome-wide significant variant at this locus in the same sample.

We also sought to follow up four additional loci showing suggestive evidence for association with AD (1 × 10^{-6} ≤ P > 5 × 10^{-8}) from the ADGC GWAS\textsuperscript{12}. These loci included CD33, EPHA1, CD2AP and ARID5B. It should be noted that evidence for suggestive association with EPHA1 and CD33 has been reported previously. Members of this collaboration were the first to report EPHA1 as showing suggestive evidence of association with AD (rs11771145, P = 1.7 × 10^{-6}; LD with ADGC SNP rs11767557: \( r^2 = 0.28, D' = 0.75 \))\textsuperscript{11}, which included GERAD1 and EADI1 samples reported on here. Similarly, Bertram and colleagues were the first to show suggestive evidence for CD33 (rs3826566, P = 4.0 × 10^{-6}; LD with ADGC SNP rs3865444: \( r^2 = 0.13, D' = 1.0 \))\textsuperscript{19}.

We combined data from the GERAD+ dataset comprising GERAD1, EADI1, deCODE and AD-IG GWAS datasets (up to 6992 cases and 13472 controls) using inverse variance meta-analysis. The TGEN1, ADNI and Mayo1 datasets were included in the ADGC discovery set and were thus excluded from these particular analyses. We observed support for association with CD2AP (rs9349407, P = 8.0 × 10^{-4}, OR = 1.11), CD33 (rs3865444, P = 2.2 × 10^{-4}, OR = 0.89) and EPHA1 (rs11767557, P = 3.4 × 10^{-4}, OR = 0.90) when these data were combined with ADGC we observed genome-wide evidence for association with AD (rs9349407, GERAD+ & ADGC meta-P = 8.6 × 10^{-9}, OR = 1.11; rs3865444, GERAD+ & ADGC meta-P = 1.6 × 10^{-5}, OR = 0.91; rs11767557, GERAD+ & ADGC meta-P = 6.0 × 10^{-10}, OR = 0.90). We observed nominally significant evidence of association with ARID5B (rs2588969, P = 3.3 × 10^{-2}, OR = 1.06), however the direction of effect was opposite to that reported by ADGC\textsuperscript{12}, and was not significant overall (GERAD+ & ADGC meta-P = 3.6 × 10^{-1}, OR = 0.99).

Taken together, these results show compelling evidence for an additional five novel AD susceptibility loci. ABCA7 encodes an ATP-binding cassette (ABC) transporter. The ABC transporter superfamily has roles in transporting a wide range of substrates across cell membranes\textsuperscript{20}. ABCA7 is highly expressed in brain, particularly in hippocampal CA1 neurons\textsuperscript{21} and in microglia\textsuperscript{22}. ABCA7 is involved in the efflux of lipids from cells to lipoprotein particles. Notably, the main lipoproteins in brain are APOE followed by CLU. Although no evidence for epistatic interactions between the three genetic loci was observed (see Supplementary Table 4a), however, this is not a prerequisite for biological interaction between these molecules. In addition, ABCA7 has been shown to regulate APP processing and inhibit \( \beta \)-amyloid secretion in cultured cells overexpressing APP\textsuperscript{23}. ABCA7
also modulates phagocytosis of apoptotic cells by macrophages mediated through the C1q complement receptor protein on the apoptotic cell surface. \textit{ABCA7} is an orthologue of \textit{C. elegans ced-7}, the product of which is known to clear apoptotic cells and the high levels of expression of \textit{ABCA7} in microglia are consistent with such a role.

The genes in the \textit{MS4A} cluster on chromosome 11 have a common genomic structure with all other members of the family, including transmembrane domains indicating that they are likely to be part of a family of cell surface proteins. \textit{MS4A2} encodes the beta subunit of high affinity IgE receptors. The remaining genes in the LD block have no known specific functions. \textit{CD33} is a member of the sialic-acid-binding immunoglobulin-like lectins (Siglec) family which are thought to promote cell-cell interactions and regulate functions of cells in the innate and adaptive immune systems. Most members of the Siglec family, including \textit{CD33}, act as endocytic receptors, mediating endocytosis through a mechanism independent of clathrin. \textit{CD2AP} (CD2-associated protein) is a scaffold/adaptor protein which associates with cortactin, a protein also involved in the regulation of receptor mediated endocytosis. It is striking that these two new susceptibility genes for AD, and the recently established susceptibility genes \textit{PICALM} and \textit{BIN1} are all implicated in cell-cell communication and transduction of molecules across the membrane. \textit{EPHA1} is a member of the ephrin receptor subfamily. Ephrins and Eph receptors are membrane bound proteins which play roles in cell and axon guidance and in synaptic development and plasticity. However \textit{EPHA1} is expressed mainly in epithelial tissues where it regulates cell morphology and motility. Additional roles in apoptosis and inflammation have also been proposed.

Our study has generated strong statistical evidence that variants at \textit{ABCA7} and the \textit{MS4A} gene cluster confer susceptibility to AD, which replicates over a number of independent case control samples. We also provide independent support for three loci showing suggestive evidence in a companion paper: \textit{CD33}, \textit{CD2AP} and \textit{EPHA1}, which when the data are combined show genome-wide levels of significance. Finally, we provide further evidence for \textit{BIN1} and \textit{CR1} loci as susceptibility loci. What is striking about our findings is the emerging consistency in putative function of the genes identified. Five of the recently identified AD susceptibility loci \textit{CLU}, \textit{CR1}, \textit{ABCA7}, \textit{CD33} and \textit{EPHA1} have putative functions in the immune system; \textit{PICALM}, \textit{BIN1}, \textit{CD33}, \textit{CD2AP} are involved in processes at the cell membrane, including endocytosis and \textit{APOE}, \textit{CLU} and \textit{ABCA7} in lipid processing. It is conceivable that these processes would play strong roles in neurodegeneration and Aβ clearance from the brain. These findings therefore provide new impetus for focused studies aimed at understanding the pathogenesis of AD.

**Supplementary Material**

Refer to Web version on PubMed Central for supplementary material.

**Acknowledgments**

For complete acknowledgements please see the Supplementary Note. We thank the individuals and families who took part in this research and those who funded the groups who contributed to this study: Wellcome Trust; MRC (UK); ART; WAG; Mercer’s Institute for Research on Ageing; Alzheimer’s Society; Ulster Garden Villages; NI R&D Office; Royal College of Physicians; Dunhill Medical Trust; BRACE; US NIH, the Barnes Jewish Foundation; Charles and Joanne Knight Alzheimer’s Research Initiative; UCL Hospital/UCL Biomedical Centre; Lundbeck; German Federal Ministry of Education and Research Competence Network Dementia and Competence Network Degenerative Dementia; Alfred Krupp von Bohlen und Halbach-Stiftung; IRP of the NIA Department of Health and Human Services; University of Antwerp, Fund for Scientific Research-Flanders; Foundation for Alzheimer Research; Methusalem Excellence grant; Federal Science Policy Office Interuniversity Attraction Poles program; Mayo AD Research Center; NINDS; Robert and Clarice Smith Postdoctoral Fellowship and AD Research Program; Palumbo Professorship in AD Research; Carl Edward and Susan Bass Bolch Gift; Institut Pasteur de Lille; CNG; Fondation pour la Recherche Médicale Caisse; Nationale Maladie des Travailleurs Salariés, Direction...
References

1. Gatz M, et al. Role of genes and environments for explaining Alzheimer disease. Arch Gen Psychiatry. 2006; 63:168–74. [PubMed: 16461860]

2. Saunders AM, et al. Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer’s disease. Neurology. 1993; 43:1467–72. [PubMed: 8350998]

3. Harold D, et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer’s disease. Nat Genet. 2009; 41:1088–93. [PubMed: 19734902]

4. Lambert JC, et al. Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer’s disease. Nat Genet. 2009; 41:1094–9. [PubMed: 19734903]

5. Corneveaux JJ, et al. Association of CR1, CLU and PICALM with Alzheimer’s disease in a cohort of clinically characterized and neuropathologically verified individuals. Hum Mol Genet. 2010; 19:3295–301. [PubMed: 20534741]

6. Zhang Q, et al. Complement receptor 1 polymorphisms and risk of late onset Alzheimer’s disease. Brain Res. 2010 [PubMed: 20558149]

7. Carrasquillo MM, et al. Replication of CLU, CR1, and PICALM associations with alzheimer disease. Arch Neurol. 2010; 67:961–4. [PubMed: 20554627]

8. Jun G, et al. Meta-analysis Confirms CR1, CLU, and PICALM as Alzheimer Disease Risk Loci and Reveals Interactions With APOE Genotypes. Arch Neurol. 2010

9. Kamboh MI, et al. Association of CLU and PICALM variants with Alzheimer’s disease. Neurobiol Aging. 2010

10. Biffi A, et al. Genetic variation and neuroimaging measures in Alzheimer disease. Arch Neurol. 2010; 67:677–85. [PubMed: 20558387]

11. Seshadri S, et al. Genome-wide analysis of genetic loci associated with Alzheimer disease. JAMA. 2010; 303:1832–40. [PubMed: 20460622]

12. Naj AC. Common variants in MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer’s disease. Nat Genet. (In Press).

13. Reiman EM, et al. GAB2 alleles modify Alzheimer’s risk in APOE epsilon4 carriers. Neuron. 2007; 54:713–20. [PubMed: 17553421]

14. Petersen RC, et al. Alzheimer’s Disease Neuroimaging Initiative (ADNI): clinical characterization. Neurology. 2010; 74:201–9. [PubMed: 20042704]

15. Pe’er I, Yelensky R, Altshuler D, Daly MJ. Estimation of the multiple testing burden for genome-wide association studies of nearly all common variants. Genet Epidemiol. 2008; 32:381–5. [PubMed: 18348202]

16. Adzhubei IA, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010; 7:248–9. [PubMed: 20354512]

17. Stranger BE, et al. Genome-wide associations of gene expression variation in humans. PLoS Genet. 2005; 1:e78. [PubMed: 16362079]

18. Gibbs JR, et al. Abundant quantitative trait Loci exist for DNA methylation and gene expression in human brain. PLoS Genet. 2010; 6:e1000952. [PubMed: 20485568]
19. Bertram L, et al. Genome-wide association analysis reveals putative Alzheimer’s disease susceptibility loci in addition to APOE. Am J Hum Genet. 2008; 83:623–32. [PubMed: 18976728]

20. Kim WS, Weickert CS, Garner B. Role of ATP-binding cassette transporters in brain lipid transport and neurological disease. J Neurochem. 2008; 104:1145–66. [PubMed: 17973979]

21. Kim WS, et al. Abca7 null mice retain normal macrophage phosphatidylcholine and cholesterol efflux activity despite alterations in adipose mass and serum cholesterol levels. J Biol Chem. 2005; 280:3989–95. [PubMed: 15550377]

22. Kim WS, Guillemin GJ, Glaros EN, Lim CK, Garner B. Quantitation of ATP-binding cassette subfamily-A transporter gene expression in primary human brain cells. Neuroreport. 2006; 17:891–6. [PubMed: 16738483]

23. Jehle AW, et al. ATP-binding cassette transporter A7 enhances phagocytosis of apoptotic cells and associated ERK signaling in macrophages. J Cell Biol. 2006; 174:547–56. [PubMed: 16908670]

24. Liang Y, Buckley TR, Tu L, Langdon SD, Tedder TF. Structural organization of the human MS4A gene cluster on Chromosome 11q12. Immunogenetics. 2001; 53:357–68. [PubMed: 11486273]

25. Kinet JP, et al. Isolation and characterization of cDNAs coding for the beta subunit of the high-affinity receptor for immunoglobulin E. Proc Natl Acad Sci U S A. 1988; 85:6483–7. [PubMed: 2970642]

26. Crocker PR, Paulson JC, Varki A. Siglecs and their roles in the immune system. Nat Rev Immunol. 2007; 7:255–66. [PubMed: 17380156]

27. Tateno H, et al. Distinct endocytic mechanisms of CD22 (Siglec-2) and Siglec-F reflect roles in cell signaling and innate immunity. Mol Cell Biol. 2007; 27:5699–710. [PubMed: 17562860]

28. Dustin ML, et al. A novel adaptor protein orchestrates receptor patterning and cytoskeletal polarity in T-cell contacts. Cell. 1998; 94:667–77. [PubMed: 9741631]

29. Lynch DK, et al. A Cortactin-CD2-associated protein (CD2AP) complex provides a novel link between epidermal growth factor receptor endocytosis and the actin cytoskeleton. J Biol Chem. 2003; 278:21805–13. [PubMed: 12672817]

30. Martinez A, Soriano E. Functions of ephrin/Eph interactions in the development of the nervous system: emphasis on the hippocampal system. Brain Res Brain Res Rev. 2005; 49:211–26. [PubMed: 1611551]

31. Lai KO, Ip NY. Synapse development and plasticity: roles of ephrin/Eph receptor signaling. Curr Opin Neurobiol. 2009; 19:275–83. [PubMed: 19497733]

32. Coulthard MG, et al. Characterization of the EphA1 receptor tyrosine kinase: expression in epithelial tissues. Growth Factors. 2001; 18:303–17. [PubMed: 11519828]

33. Yamazaki T, et al. EphA1 interacts with integrin-linked kinase and regulates cell morphology and motility. J Cell Sci. 2009; 122:243–55. [PubMed: 19118217]

34. Duffy SL, et al. Generation and characterization of EphA1 receptor tyrosine kinase reporter knockout mice. Genesis. 2008; 46:553–61. [PubMed: 18802966]

35. Ivanov AI, Romanovsky AA. Putative dual role of ephrin-Eph receptor interactions in inflammation. IUBMB Life. 2006; 58:389–94. [PubMed: 16801213]

36. McKhann G, et al. 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; 34:939–44. [PubMed: 6610841]

37. Myers AJ, et al. A survey of genetic human cortical gene expression. Nat Genet. 2007; 39:1494–9. [PubMed: 17982457]
Figure 1.
GERAD+ study design.
* Data for rs744373 and rs3818361 in the CHARGE consortium have been presented elsewhere\textsuperscript{15}, as has data for rs381861 in the EADI2 samples\textsuperscript{4}, as such these SNPs were not included in Stage 3.
Figure 2.
Schematic of the associated variants reported in reference to (A) the ABCA7 gene and (B) chromosomal region chr1:59.81Mb-60.1Mb harboring members of the MS4A gene cluster. Chromosome positions are shown at the top of the schematics (UCSC Feb 2009). Gene schematic: horizontal arrows indicate directions of transcription, black boxes indicate gene exons/UTR. The $-\log_{10}(P)$ of the SNPs analyzed in Stage 1 are shown in chart graph. The GERAD+ Stage 1, 2 and 3 meta-analysis $P$-values for SNPs rs3764650 (ABCA7), rs610932 (MS4A6A) and rs670139 (MS4A4E) are indicated by the red lines. The D’ LD block structure of the ABCA7 gene plus surrounding region, and chr11:59.81Mb-60.1Mb according to the CEPH HapMap data, are provided at the bottom of each schematic with lines indicating where each SNP genotyped on the Illumina 610-quad chip is represented.
Figure 3.
Forest plots showing association in the different datasets for SNPs at the *ABCA7* (rs3764650) and *MS4A* (rs610932 & rs670139) loci.
Table 1

Results of the GERAD+ study.

| SNP     | Closest Gene | CHR | MAF  | Stage 1* | Stage 2‡ | Stage 3§ | Meta-analysis of GERAD+ Stage 1, 2 and 3§ | Meta-analysis of GERAD+ & ADGC |
|---------|--------------|-----|------|----------|----------|----------|------------------------------------------|--------------------------------|
| rs3764650 | ABCA7        | 19  | 0.10 | 2.6×10⁻⁷ | 1.22     | 1.13-1.32 | 2.9×10⁻⁷ | 1.13-1.32 | 4.5×10⁻¹⁷ | 1.23     | 1.18-1.30 | 5.0×10⁻²¹ | 1.23     | 1.17-1.28 |
| rs610932  | MS4A6A       | 11  | 0.42 | 1.8×10⁻⁸ | 0.88     | 0.85-0.92 | 6.0×10⁻⁷ | 0.91     | 0.87-0.96 | 1.8×10⁻⁴ | 0.90     | 0.87-0.92 | 1.2×10⁻¹⁶ | 0.91     | 0.88-0.93 |
| rs670139  | MS4A4E       | 11  | 0.41 | 1.0×10⁻⁵ | 1.11     | 1.06-1.16 | 1.1×10⁻³ | 1.11     | 1.04-1.19 | 3.2×10⁻⁵ | 1.06     | 1.02-1.11 | 1.1×10⁻¹⁴ | 1.09     | 1.06-1.11 |
| rs3818361 | CR1          | 1   | 0.19 | 3.2×10⁻¹² | 1.21     | 1.14-1.27 | 1.4×10⁻³ | 1.14     | 1.05-1.23 | NA       | NA       | NA       | 3.7×10⁻¹⁴ | 1.18     | 1.13-1.24 |
| rs744373  | BIN1         | 2   | 0.29 | 1.5×10⁻¹⁰ | 1.17     | 1.11-1.22 | 3.8×10⁻⁵ | 1.17     | 1.08-1.25 | NA       | NA       | NA       | 2.6×10⁻¹⁴ | 1.17     | 1.12-1.21 |

CHR=Chromosome, MAF=Minor Allele Frequency in cases and controls.

*GERAD1, EAD1, ADNI, & TGEN1 <6688 Cases, 13685 Controls.
‡GERAD2, deCODE, AD-IG: 4896 AD Cases, 4903 Controls.
§EAD2, CHARGE, Mayo2 <8286 AD Cases, 21258 Controls.

CHR=Chromosome, MAF=Minor Allele Frequency in cases and controls.
Table 2

Results of the combined analysis of the ADGC and GERAD+ consortia.

| SNP    | Closest Gene | CHR | MAF  | Linkage Disequilibrium with the top ADGC SNP at each loci | GERAD+ Consortia * | GERAD+ & ADGC Metaanalysis |
|--------|--------------|-----|------|----------------------------------------------------------|----------------------|----------------------------|
|        |              |     |      | r²  | D'   | Cases | Controls | P     | OR   | 95% CI | P     | OR   | 95% CI |
| rs9349407 † | CD2AP       | 6   | 0.29 | N/A | N/A  | 6283  | 7165     | 8.0×10^{-4} | 1.11 | 1.04-1.18 | 8.6×10^{-9} | 1.11 | 1.07-1.15 |
| rs9296559 ‡ | CD2AP       | 6   | 0.29 | 0.71 | 0.95 | 6283  | 7165     | 1.5×10^{-3} | 1.10 | 1.04-1.17 | NA    | NA    | NA    |
| rs11767557 | EPHA1       | 7   | 0.21 | N/A | N/A  | 6283  | 12935    | 3.4×10^{-4} | 0.90 | 0.85-0.95 | 6.0×10^{-10} | 0.90 | 0.86-0.93 |
| rs2588969 ‡ | ARID5B      | 10  | 0.40 | N/A | N/A  | 6283  | 7165     | 3.3×10^{-2} | 1.06‡ | 1.01-1.13 | 3.6×10^{-1} | 0.99 | 0.95-1.02 |
| rs4948288 | ARID5B      | 10  | 0.26 | 0.55 | 0.78 | 6992  | 13472    | 3.6×10^{-3} | 1.07‡ | 1.03-1.15 | NA    | NA    | NA    |
| rs3865444 § | CD33        | 19  | 0.31 | N/A | N/A  | 6283  | 7165     | 2.2×10^{-4} | 0.89 | 0.84-0.95 | 1.6×10^{-9} | 0.91 | 0.88-0.93 |

CHR=Chromosome, MAF=Minor Allele Frequency in cases and controls.

* GERAD1, EAD1, deCODE, AD-IG.
† results generated from imputed data. The results from the top genotyped SNP are also shown. See Supplementary Table 6 for full details.
‡ opposite direction of effect to that reported by Naj et al.
§ data imputed in the deCODE dataset.