Initial Assessment of the Pathogenic Mechanisms of the Recently Identified Alzheimer Risk Loci

Patrick Holton¹, Mina Ryten¹, Michael Nalls², Daniah Trabzuni¹,³, Michael E. Weale⁴, Dena Hernandez¹,², Helen Crehan¹, J. Raphael Gibbs¹,², Richard Mayeux⁵, Jonathan L. Haines⁶, Lindsay A. Farrer⁷, Margaret A. Pericak-Vance⁸, Gerard D. Schellenberg⁹, The Alzheimer's Disease Genetics Consortium*, Manuel Ramirez-Restrepo¹⁰,¹¹, Anzhelika Engel¹⁰,¹¹, Amanda J. Myers¹⁰,¹¹, Jason J. Corneveaux¹², Matthew J. Huettelma⁰¹², Alissa Dillman²,¹³, Mark R. Cookson², Eric M. Reiman¹²,¹₄,¹₅, Andrew Singleton², John Hardy¹,¹⁶,† and Rita Guerreiro¹

¹ Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK
² Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD
³ Department of Genetics, King Faisal Specialist Hospital and Research Centre, Riyadh, Saudi Arabia
⁴ Department of Medical & Molecular Genetics, King's College London, Guy's Hospital, London, UK
⁵ Department of Neurology, Gertrude H. Sergievsky Center and Taub Institute on Alzheimer's Disease and the Aging Brain, Columbia University, New York, NY
⁶ Department of Molecular Physiology and Biophysics and Vanderbilt Center for Human Genetics Research, Vanderbilt University, Nashville, TN
⁷ Departments of Medicine (Biomedical Genetics), Biostatistics, Ophthalmology, Epidemiology, and Neurology, Boston University Schools of Medicine and Public Health, Boston, MA
⁸ The John P. Hussman Institute for Human Genomics and Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami, Miami, FL
⁹ Department of Pathology and Laboratory Medicine, University of Pennsylvania School of Medicine, Philadelphia, PA
¹⁰ Department of Psychiatry and Behavioral Sciences, Miller School of Medicine, University of Miami, Miami, FL
¹¹ Johnnie B. Byrd Sr. Alzheimer's Center and Research Institute, Tampa, FL
¹² Neurogenomics Division, Translational Genomics Research Institute and Arizona Alzheimer’s Consortium, Phoenix, AZ
¹³ Department of Neuroscience, Karolinska Institutet, 171 77 Stockholm, Sweden
¹⁴ Banner Alzheimer’s Institute and Department of Psychiatry, University of Arizona, Phoenix, AZ
¹⁵ Department of Psychiatry, University of Arizona, Tucson, AZ
¹⁶ Reta Lila Weston Laboratories and Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK

Summary

Recent genome wide association studies have identifiedCLU, CR1, ABCA7 BIN1, PICALM and MS4A6A/MS4A6E in addition to the long established APOE, as loci for Alzheimer’s disease. We have systematically examined each of these loci to assess whether common coding variability contributes to the risk of disease. We have also assessed the regional expression of all the genes in the brain and whether there is evidence of an eQTL explaining the risk. In agreement with other studies we find that coding variability may explain the ABCA7 association, but common coding variability does not explain any of the other loci. We were not able to show that any of the loci had eQTLs within the power of this study. Furthermore the regional expression of each of the loci did not match the pattern of brain regional distribution in Alzheimer pathology. Although these results are mainly negative, they allow us to start defining more realistic alternative approaches to determine the role of all the genetic loci involved in Alzheimer’s disease.

Keywords: Alzheimer's disease, genetic risk, GWAS

*Coauthors from the Alzheimer's Disease Genetics Consortium are listed at the end of the manuscript.

¹Corresponding author: JOHN HARDY, Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK. Tel: +44 207 829 8722; Fax: +44 207 833 1016; E-mail: j.hardy@ion.ucl.ac.uk
Introduction

The recent application of genome wide association studies (GWAS) to the dissection of the risk for late onset Alzheimer’s disease (AD) has proved an outstanding success and has led to the identification of many new loci (CLU, PICALM, CR1, BIN1, MS4A6A/MS4A4E, CD33, CD2AP, ABCA7 and EPHA1) in addition to the long established apolipoprotein E locus (Harold et al., 2009; Lambert et al., 2009; Hollingworth et al., 2011; Naj et al., 2011). When such loci are identified, they simply appear as single nucleotide polymorphisms (SNPs), which have significantly different frequencies between cases and controls. It is not initially clear whether these risk SNPs are in linkage disequilibrium (LD) with coding changes or have an impact on gene expression. For all traits studied by GWAS only ~12% of the associated SNPs are located in, or occur in high LD with, protein coding regions of genes. The vast majority (~80%) of trait associated SNPs are located in intergenic regions or noncoding introns (Manolio, 2010). AD is no different: taking into account the 21 SNPs reported in the nine new loci by GWAS assessing over 1500 cases and 1500 controls (see Table 1 for details on the SNPs), 10 are located in intergenic regions; 8 in intronic regions; 1 SNP is located in the 3’UTR of MS4A4E; and 2 SNPs are located in exons (one SNP is a non-synonymous variant in ABCA7—Gly1527Ala and one synonymous variant was found as significant in PICALM). These findings clearly indicate that follow up studies should not only examine coding variability, but should also pay close attention to the potential roles of these intronic and intergenic regions in the regulation of gene expression (Myers et al., 2007; Hardy and Singleton, 2009; Manolio, 2010). In fact, for any disease associated SNP, the true variant underlying the phenotype studied may be: (1) the GWAS hit itself; (2) a known common SNP in LD with the identified GWAS hit; (3) an unknown common SNP or rare single nucleotide variant tagged by a haplotype on which the hit occurs; or (4) a linked copy number variant (Hindorff et al., 2009). In general, GWAS follow up studies rely on fine mapping of the associated locus or loci, deep re-sequencing of the associated region(s) in samples of interest (which allows the identification of all possible functional variants) and a variety of bioinformatic approaches to prioritise variants to be further studied (Stranger et al., 2011).

Confirmed functional variants underlying validated GWAS hits are still sparse in the literature, when considering all the diseases and traits studied, but each of these is extremely valuable to the respective research and clinical environments. For example, the IRF5 locus includes variants that disrupt intron splicing, decrease mRNA transcript stability and delete part of the interferon regulating factor protein (Graham et al., 2007), explaining the independent associations of this locus with three different phenotypes: systemic lupus erythmatosis (Sigurdsson et al., 2005; Graham et al., 2006), inflammatory bowel disease (Dideberg et al., 2007), and rheumatoid arthritis (Stahl et al., 2010). Similarly, allele-specific chromatin re-modelling affecting the expression of several genes in the ORMDL3 locus region (Verlaan et al., 2009) explains its association with asthma (Moffatt et al., 2007), Crohn’s disease (Barrett et al., 2008) and type 1 diabetes (Barrett et al., 2009). With this in mind we have undertaken an analysis of the recently identified AD risk loci with three components: (1) we have assessed by sequencing whether there is common coding variability in LD with the associated SNPs (2) we have assessed in a database of control human cerebral cortex samples whether the SNPs are associated with genetic variability in expression (3) we have assessed the regional distribution of expression and splicing of the genes at the risk loci to see whether this distribution is in any way consistent with the distribution of pathology in the disease.

Materials and Methods

Genotyping Analysis

Samples

The 96 DNA samples selected for genotyping were previously used in a GWAS in AD (Corneveaux et al., 2010). These 96 Alzheimer disease samples were diagnosed according to the NINDS-ADRDA diagnostic criteria for Alzheimer disease, consisting of 67 females and 29 males with a mean age of 81 years (range 66–95) and mean age at onset of 71.9 years (ranging from 65 to 85 years).

SNPs studied

The GWAS SNPs studied were those found to be significantly associated with late onset AD (LOAD) by two recent studies: (Corneveaux et al., 2010; Hollingworth et al., 2011). For a complete list of SNPs analysed in the present study please refer to Table 1.

Coding SNPs were chosen based upon their reported minor allele frequency (MAF) or heterozygosity in dbSNP. For this, publicly available data in dbSNP was used and SNPs were chosen based upon the fact that they induced a coding change in the resultant protein and that they had a MAF or heterozygosity greater than 0.05 in the general population. For CR1, SNPs were excluded if they were located in highly homologous exons in order to avoid genotyping errors.

Most of the SNPs studied conformed to these specifications, however, there were some that did not and were included in the study because no better proxies were available (such as rs17259045, rs76037557, rs74727972, rs79741566, rs72973581).
Table 1 Features and minor allele frequencies of all SNPs studied here, including the GWAS hits

<table>
<thead>
<tr>
<th>SNP</th>
<th>Gene/locus</th>
<th>Chr</th>
<th>Chr position—genome build 37.1</th>
<th>Chr position—genome build 36.3</th>
<th>mRNA accession number</th>
<th>SNP class</th>
<th>Ancestral allele</th>
<th>MAF in our study</th>
<th>1000 genomes MAF</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs4844600</td>
<td>CR1</td>
<td>1</td>
<td>207679307</td>
<td>205745930</td>
<td>NM_000651.4</td>
<td>Triallelic SNP:</td>
<td>G</td>
<td>A = 0.23</td>
<td>A = 0.209/263</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>cds-synonymous</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Glu60Glu) or missense</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Glu60Asp)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs665641</td>
<td>CR1</td>
<td>1</td>
<td>207692049</td>
<td>20578672</td>
<td>NM_000651.4</td>
<td>Intronic</td>
<td>G</td>
<td>A = 0.18</td>
<td>A = 0.107/135</td>
</tr>
<tr>
<td>rs3737002</td>
<td>CR1</td>
<td>1</td>
<td>207760773</td>
<td>205827936</td>
<td>NM_000651.4</td>
<td>Missense (Thr1858Met)</td>
<td>C</td>
<td>T = 0.28</td>
<td>T = 0.268/337</td>
</tr>
<tr>
<td>rs17259045</td>
<td>CR1</td>
<td>1</td>
<td>207782707</td>
<td>205849330</td>
<td>NM_000651.4</td>
<td>Missense (Asn1990Ser)</td>
<td>A</td>
<td>G = 0.09</td>
<td>G = 0.104/110</td>
</tr>
<tr>
<td>rs17047661</td>
<td>CR1</td>
<td>1</td>
<td>207782889</td>
<td>205849512</td>
<td>NM_000651.4</td>
<td>Missense (Arg2051Gly)</td>
<td>A</td>
<td>G = 0</td>
<td>G = 0.242/219</td>
</tr>
<tr>
<td>rs8446609</td>
<td>CR1</td>
<td>1</td>
<td>207782916</td>
<td>205849539</td>
<td>NM_000651.4</td>
<td>Missense (Thr2066Ser)</td>
<td>T</td>
<td>A = 0.02</td>
<td>A = 0.01/12</td>
</tr>
<tr>
<td>rs6691117</td>
<td>CR1</td>
<td>1</td>
<td>207782931</td>
<td>205849554</td>
<td>NM_000651.4</td>
<td>Missense (Ile2065Val)</td>
<td>G</td>
<td>G = 0.22</td>
<td>G = 0.382/481</td>
</tr>
<tr>
<td>rs3818361</td>
<td>CR1</td>
<td>1</td>
<td>207784968</td>
<td>205881591</td>
<td>NM_000651.4</td>
<td>Intronic</td>
<td>C</td>
<td>A = 0.01</td>
<td>A = 0.301/379</td>
</tr>
<tr>
<td>rs3811381</td>
<td>CR1</td>
<td>1</td>
<td>207790088</td>
<td>205856711</td>
<td>NM_000651.4</td>
<td>Missense (Pro2277Arg)</td>
<td>G</td>
<td>G = 0.21</td>
<td>G = 0.165/208</td>
</tr>
<tr>
<td>rs2296160</td>
<td>CR1</td>
<td>1</td>
<td>207795320</td>
<td>205861943</td>
<td>NM_000651.4</td>
<td>Triallelic SNP:</td>
<td>C</td>
<td>A = 0.2</td>
<td>A = 0.233/293</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>missense Th2419Ala or Th2419Ser</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Chr, chromosome; AD Risk–C: this SNP was identified as a risk marker for AD in the GWAS study published by Corneveaux et al. (2010); Risk–H: this SNP was identified as a risk marker for AD in the GWAS study published by Hollingworth et al. (2011); Targeted: this SNP was identified as a common coding SNP in a gene containing an AD risk SNP from the previously mentioned GWAS; Found: this SNP was found when genotyping targeted SNPs because it was located near one of the targeted SNPs and showed up during the sequencing of the targeted SNP region.
DNA Methylation and mRNA Expression in the Human Brain

Tissue samples
Frozen samples from the frontal cerebral cortex and cerebellum were obtained from 387 Caucasian subjects without neurological disease in lifetime (Gibbs et al., 2010; Hernandez et al., 2011). Genomic DNA was extracted using phenol-chloroform and quantified on a Nanodrop1000 spectrophotometer before genotyping or bisulfite conversion for DNA methylation analysis.

CpG methylation
Bisulfite conversion of genomic DNA was performed using Zymo EZ-96 DNA Methylation kits according to the manufacturers protocol, using 1 μg of DNA input. The CpG methylation status of DNA at >27,000 sites was determined using Illumina Infinium HumanMethylation27 BeadChips (Illumina Inc., San Diego, CA, USA). Samples were included in the analysis if the threshold call rate for inclusion of was >95% in the tissue. As a second quality control, we compared reported genders with methylation levels of CpG sites on the X chromosome. After these steps, 292 samples with data at 27,465 CpG sites in the frontal cortex tissue samples and 27,419 sites in the cerebellum tissue samples were used for further analysis.

mRNA expression
Messenger RNA (mRNA) expression was analysed using Illumina HumanHT-12 v3 Expression Beadchips. Individual probes were excluded from analyses if the P value for detection was >0.01 and samples were excluded if <95% of probes were detected. Intensity values for each probe were normalised using cubic spline and transformed using log2 prior to statistical analyses. Probes were annotated using ReMOAT tool to exclude individual probes that are known to have problems in design or with ambiguous mapping. We also removed all probes that included any known SNP. After these quality control steps, data was available for 399 samples at 9814 probes from the frontal cortex, and 9587 probes in cerebellum.

Genotyping and imputation of control brains in epigenetic analyses
The same tissue samples were genotyped using Illumina HumanHap550 v3, Human610-Quad v1 or Human660W-Quad v1 Infinium Beadchips and shared SNPs were extracted for each sample. We excluded samples where the reported sex did not match X chromosome heterozygosity from genotype data or if the per sample genome-wide call rate was less than 95%. Individual SNPs were excluded if there was a <95% genotyping success rate per SNP, if MAF < 0.01 or if Hardy–Weinberg equilibrium (HWE) P < 1 × 10⁻⁷. Multi-dimensional scaling was used to cluster samples after merging SNPs common to CEU, JPT, CHB and YRI samples from Phase II of HapMap. Outliers >3 standard deviations from the mean component vector estimates for C1 or C2 for CEU samples were then removed, as were samples sharing greater than a proportion of 0.15 alleles. Genotypes for all European ancestry participants were imputed using MACHv1.0.16 with haplotypes derived from sequencing of 112 European ancestry samples present in the August 2009 release of phased data from the 1000 Genomes Project (available at http://www.sph.umich.edu/csg/abecasis/MACH/download/1000G-Sanger-0908.html). Data was imputed by first generating error and crossover maps as parameter estimates for the imputation on a randomly selected set of 200 samples over 100 iterations of the initial statistical model. These parameter estimates were then used to generate maximum likelihood allele dosages per SNP based on reference haplotypes for the entire study cohorts. We excluded SNPs with R² quality estimates <0.30, resulting in ~5.1 million SNPs available for analysis.

methQTL and eQTL mapping
SNPs with a fixed-effects P < 1 × 10⁻⁵ from the AD meta-analysis (Stage 1 + 2) were considered candidate quantitative trait loci (QTL; Naj et al., 2011). For each SNP, CpG sites and expression probes within ±1 MB were used for linear regression modelling using MACH2QTLv1.08. We estimated the association between the allelic dosage of each SNP against gene expression or methylation levels using linear regression.
models adjusted for covariates of gender and age at death, the first two component vectors from multi-dimensional scaling, post-mortem interval (PMI), brain bank and batch in which preparation or hybridisation were performed. SNPs with less than three minor homozygotes detected were excluded from analyses. We tested probes within 1 MB of each of 224 candidate probes in the expression datasets and 220 in the methylation datasets, resulting in 2542 associations for expression-QTLs and 11,522 associations for methylation-QTLs in the frontal cortex samples and 2395 expression-QTLs and 11,510 methylation-QTLs in the cerebellum. The resulting $P$ values were corrected for multiple testing using the Bonferroni method after removing SNPs having $r^2 > 0.5$ with SNPs in adjacent sliding windows of 50 SNPs that moved two SNPs per iteration. After these filters, the analyses used 152 mRNA probes and 603 CpG sites.

### Regional Brain Expression and Splicing Analysis

**Human post-mortem brain tissue collection and mRNA extraction**

A detailed description of the samples used in the study, tissue processing and dissection is provided in Trabzuni et al. (2011). In brief, brain and CNS tissue originating from 137 control individuals was collected by the Medical Research Council (MRC) Sudden Death Brain and Tissue Bank, Edinburgh, UK (Millar et al., 2007), and the Sun Health Research Institute (SHRI) an affiliate of Sun Health Corporation, USA (Beach et al., 2008). All samples had fully informed consent for retrieval and were authorised for ethically approved scientific investigation (Research Ethics Committee number 10/H0716/3).

Total RNA was isolated from human post-mortem brain tissues using the miRNeasy 96 kit (Qiagen), processed with the Ambion® WT Expression Kit and Affymetrix GeneChip Whole Transcript Sense Target Labelling Assay, and hybridised to the Affymetrix Exon 1.0 ST Arrays following the manufacturers’ protocols. Hybridised arrays were scanned on an Affymetrix GeneChip® Scanner 3000 7G and visually inspected for hybridisation artifacts.

**Exon array data analysis**

All arrays were pre-processed using Robust Multiarray Averaging (RMA; Irizarry et al., 2003) with quantile normalisation and GC background correction in Partek’s Genomics Suite v6.6 (Partek Incorporated, St. Louis, MO, USA). In order to filter out low expression signals, detection above background (DABG) $P$ values of exon probe sets were calculated using Affymetrix Power Tools v1.14.3 (APT, http://www.affymetrix.com/partners_programs/programs/developer/tools/powertools.affx). After re-mapping the Affymetrix probe sets onto human genome build 19 (GRCh37) as documented in the Netaffx annotation file (HuEx-1.0-st-v2 Probeset Annotations, Release 31), we restricted analysis to 174,328 probe sets that had gene annotation, contained at least three probes with unique hybridisation and had DABG $P < 0.001$ in 50% of male or female individuals. We defined an expressed gene as any gene containing $\geq 1$ exon with a median DABG $P < 0.001$ in at least 50% of male or female individuals in at least one brain region. The gene-level expression was calculated for 19,597 genes by calculating the Winsorised mean value (winsorizing the data below 10% and above 90%) of all probe set signals annotated to a single gene. Region-specific expression and splicing was investigated using Partek’s mixed-model ANOVA and alternative splice ANOVA (Partek Genomics Suite v6.6). In all types of analysis, the date of array hybridisation, brain bank and gender were included as co-factors. All $P$ values were corrected for multiple comparisons using Bonferroni correction.

### Results and Discussion

GWAS are able to identify associations between phenotypes and genetic loci. Since only tagging SNPs and not all genetic variants are assessed in these studies it is not possible, solely from GWAS results, to accurately pinpoint the associated variant(s) or even gene(s). In this study, we have chosen to study, for each associated region, the gene considered as the most likely to be associated by either having been reported as such in the original GWAS or because it was considered as the most interesting gene in the region from a functional perspective in regards to AD pathobiology.

### AD Loci

**CR1—complement component (3b/4b) receptor 1; chromosome 1q32**

GWAS have consistently identified intronic SNPs in **CR1** (rs6701713, rs1408077, rs3818361 and rs6656401) to be associated with an increased risk of AD onset (Fig. 1). The LD between these intronic SNPs and common coding variants in the same gene was investigated.

The genotyping of eight common coding SNPs in **CR1** in 96 pathologically confirmed late-onset AD cases revealed that the GWAS associated SNP rs6656401 was in LD with two coding SNPs: rs4844600 (p.Glu60Asp, NP_000642.3) and rs2296160 (p.Thr2419Ala, NP_000642.3), located in exons 2 and 44 respectively. The GWAS associated SNP rs3818361
Figure 1. Representation of the CR1 gene with the localization of the SNPs found as significant by GWAS as well as the SNPs studied here. On the left, linkage disequilibrium plot for the CR1 SNPs found to be significant in GWAS and the common coding variants genotyped in this study.
was not in LD with any of the common coding SNPs studied here (Fig. 1).

The structure of the human CR1 gene is complex. The gene is composed of tandem long homologous repeating segments that encode binding sites for C3b or C4b. Four CR1 alleles differing in the total numbers of repeating segments are known and thought to have arisen through an homologous recombination with unequal crossover mechanism. The encoded protein is made up of four structurally significant domains. These are the signal peptide, extracellular, transmembrane and cytoplasmic domains (Wong, 1990). Three of these domains are homologous in each of the four allotypes but they differ in the lengths of the extracellular domain. This region is made up of short consensus repeats (SCRs) that are also known as complement-control-protein repeats (CCPs) (Klickstein et al., 1988). These repeats are highly conserved and are characterised by the presence of three cysteine residues and one tryptophan. In addition, there is a high degree of homology between every eighth SCR, thus grouping the SCRs into sevens. Each group of seven SCRs is termed a long homologous repeat (LHR) (Klickstein et al., 1988). The smallest allotype is rare and is termed either CR1-C or CR1-F, it has three LHR regions. The most common allotype is called CR1-A or CR1-F and has four LHR regions. The CR1-B or CR1-S allotype contains five LHR regions and the very rare CR1-D allotype contains six LHR regions (Holers et al., 1987). The extracellular LHR regions contain the binding sites for the protein, including the binding sites for the complement fragments C3b and C4b, thus individuals with different CR1 alleles have different numbers of binding sites for these complement fragments and it is likely that this results in a different ability in the clearance of these fragments. There still remains some uncertainty about this process, mainly because it is currently poorly understood as to whether a decreased ability to clear complement fragments is beneficial or harmful in terms of AD aetiology. It had previously been thought that an increased ability to clear the complement fragments would be beneficial as it would decrease the activation of the complement system, generally thought of as pathogenic. However, it has been shown that individuals presenting CR1-B or CR1-S genotypes (the largest common allotype) actually have a raised risk of AD (Brouwers et al., 2012), indicating that an increased ability to clear the complement fragments C3b and C4b may be pathogenic in the situation of AD. Additionally, common variation at the CR1 locus, more specifically rs6656401 has been shown to have a broad impact on cognition. This effect was shown to be largely mediated by an individual’s amyloid plaque burden (Chibnik et al., 2011).

Using microarrays, the only brain regions in which CR1 could be detected were white matter and cerebellum and the expression levels in both regions were low. This would suggest that the role of CR1 in AD may be related to its function at a systemic level or in relation to the brain vasculature. Taken together, our results suggest that CR1 genetic variability does not act through different splicing variants or through differences in expression. The most probable scenario is that the variants found to be significant in GWAS and the coding variants found here to be in LD with these GWAS hits are tagging the structural variants known to exist in the CR1 gene.

**BIN1—bridging integrator 1; chromosome 2q14**

Three SNPs in the BIN1 locus (rs744373 (Hollingworth et al., 2011; Hu et al., 2011), rs7561528 (Hu et al., 2011; Haj et al., 2011) and rs12989701 (Hu et al., 2011) have been identified by GWAS as associated with LOAD. These three SNPs lie in a noncoding region upstream of BIN1 and downstream of CYP27C1. In this locus, BIN1 (encoding the ‘bridging integrator 1’ protein) appears to be the most likely functional candidate. Neither of the GWAS associated SNPs lie in regulatory regions, CpG islands or in microRNA target sites. Rs744373 is in a reported recombination hotspot (in HapMap) and all SNPs are in predicted transcription binding sites. There are only four non-synonymous SNPs published (NCBI SNP database, version 132) in BIN1 and for these, MAFs are described for rs112318500 and only for African populations. We attempted to genotype two SNPs in BIN1 in our cohort, rs76037557 and rs112318500, and were unable to identify the presence of the minor allele in any of the cases studied.

Although no significant eQTLs could be found for BIN1 in our samples, it is interesting to note the regional expression differences in its expression within the CNS. On the basis of the microarray results, BIN1 had the highest expression in white matter, with the mean gene expression being 5.1 times higher in white matter as compared to cerebellum (Bonferroni corrected $P < 1.0 \times 10^{-30}$). In addition, we found evidence of alternative splicing by brain region with one or more of three isoforms (NM_139343, NM_139344 and NM_139345, Bonferroni corrected alternative splicing $P < 1.0 \times 10^{-30}$) being lower in white matter as compared to all other CNS regions (Fig. 2).

Isoforms that are expressed in the central nervous system are thought to be involved in synaptic vesicle endocytosis and may interact with dynamin, synaptotagmin, endophilin and clathrin. More specifically, BIN1 has been shown to be involved in both dynamin-mediated (Wigge and McMahon, 1998) and clathrin-mediated endocytosis (CME) (Pant et al., 2009). CME is thought to mediate the internalisation of Amyloid Precursor Protein (APP; MIM 104760) from the cell surface (Nordstedt et al., 1993), after which the Aβ peptide can be cleaved from the APP. Presence of the Aβ peptide in turn inhibits CME (Kelly and Ferreira, 2007) and therefore stops excess APP from entering the cell by a mechanism of
auto-inhibition. The BIN1 protein may also have an important endocytic role being involved in synaptic vesicle recycling at the synaptic terminal (Di Paolo et al., 2002; Pant et al., 2009). Aberrant splice variants have been described to be expressed in tumor cell lines and misregulated alternative splicing of BIN1 has been shown to be associated with T tubule alterations and muscle weakness in myotonic dystrophy (Fugier et al., 2011), which leads us to hypothesise a similar situation for the CNS isoforms.

Both BIN1 and PICALM are part of endocytic pathways (Guerreiro and Hardy, 2011; Olgiati et al., 2011). The results we present here (Figs 2 and 3) show both higher expression of BIN1 and a different complement of BIN1 isoforms in white matter. Along with the different balance of isoforms for
PICALM in white matter (for more details see next section) we speculate that endocytic pathways function in a distinct manner in white matter as compared to other brain regions.

AD has historically been characterised by neuronal loss and grey matter atrophy. More recently the involvement of white matter in the disease has started to be considered (Hua et al., 2008). These results seem even more interesting in light of the association obtained by Braskie et al. (2011) between Clusterin rs11136000 and white matter microstructure in young adults. Although no replication dataset was available in this study, white matter disturbances in AD seem to be a common factor worthy of further analyses.
**PICALM**—phosphatidylinositol binding clathrin assembly protein; chromosome 11q14

There are four SNPs associated with the **PICALM** gene that have been identified in GWAS. These are rs592297 (Harold et al., 2009), rs561655 (Harold et al., 2009; Naj et al., 2011), rs541458 (Lambert et al., 2009; Corneveaux et al., 2010; Harold et al., 2009) and rs3851179 (Harold et al., 2009; Seshadri et al., 2010). The first of these SNPs is a synonymous variant (p.Gln174Gln) located in exon 5 of **PICALM** which may influence the activity of a sequence predicted to be an exon splicing enhancer. The other three SNPs are located upstream of the gene, rs561655 is found within a region that is thought to be a transcription factor binding site (Harold et al., 2009). It is possible that the change at this locus may increase or decrease the affinity of this region of the **PICALM** gene to transcription factors, thus changing the expression levels of the gene. rs541458 is located 8kb 5′ of **PICALM** and has been shown to be in LD with rs3851179, located 88.5kb 5′ of the gene (Harold et al., 2009). It is possible that the presence of both these SNPs in the 5′ region outside the gene may have an effect on the expression of the gene.

Eleven non-synonymous SNPs have been described in the **PICALM** gene (NCBI SNP database, version 132). From these, two were reported to have an established heterozygosity >0.05 (rs118027183 and rs74727972). Genotyping these SNPs in our cohort did not reveal any case with the minor allele, thus it is not likely that the association seen in the GWAS is due to common coding variability in this gene.

Similarly to **BIN1**, **PICALM** is involved in CME (Dreyling et al., 1996; Tebar et al., 1999; Yao et al., 2005) and has been shown to have a particular influence on the activity of VAMP-2 (Harel et al., 2008), a SNARE protein responsible for directing neurotransmitter vesicles to the presynaptic membrane. This role of the protein, allied to the observation of a reduced synaptic density in the brains of AD patients, suggests that the activity of this protein is potentially important in disease aetiology.

The mean gene expression of **PICALM** was found to be 1.9 times higher in white matter as compared to cerebellum (Bonferroni corrected \(P < 1.0 \times 10^{-30}\)). Two **PICALM** isoforms are detected by the exon array: full-length **PICALM** and a shorter form that lacks exon 13. Our results show that the short form of **PICALM** is expressed at lower levels in white matter as compared to the other CNS regions, in particular cerebellum (Fig. 3, Bonferroni corrected alternative splicing \(P < 1.0 \times 10^{-30}\)).

Several methQTL associations were significant after multiple test correction. In cerebellum, these include associations with CpG sites in **APOC1**, **BCL3** and **CBLC** on chromosome 19, as well as **ME3**, **MGC34732** and **MS4A6E** on chromosome 11. In frontal cortex, six significant associations with a CpG site mapping to **APOE**, as well as one association at **APOC2** were identified. From these, only the association observed between rs10751134 in **PICALM** and a CpG site in **ME3** (a gene adjacent to **PICALM**, Fig. 4) is in moderate LD with the GWAS SNPs studied here \((\chi^2\) between rs10751134 and rs561655, rs3851179, rs541458 in HapMap populations varies between 0.5 and 0.6) and may contribute to the genome-wide association signal found for **PICALM**.

**MS4A6A/MS4A4E**—membrane-spanning 4-domains, subfamily A, members 6A and 4E; chromosome 11q12.1

There are three SNPs in the **MS4A** gene cluster that have been associated with an increased risk of LOAD. These are rs4938933 in **MS4A4A** (Hollingworth et al., 2011; Naj et al., 2011), rs670139 in **MS4A4E** (Hollingworth et al., 2011; Naj et al., 2011) and rs610932 in **MS4A4A** (Hollingworth et al., 2011). Several SNPs are described in these genes. In order to have a comprehensive view of the locus we targeted non-synonymous coding SNPs in **MS4A6A**, **MS4A4E** and **MS4A4A** (Table 1). As expected the SNPs shown to be significantly associated with AD in the previous GWAS were in LD with each other. No significant LD was observed between these AD risk SNPs and the targeted coding SNPs studied here (highest \(\chi^2 = 0.5\) between rs610932 and rs7232), indicating that non-synonymous common genetic variability in the **MS4A4** locus probably does not explain the associations established in the GWAS.

The role of the genes located in the **MS4A** cluster is so far poorly understood. The cluster is found on chromosome 11 and is made up of at least 16 genes (Liang and Tedder, 2001). It has been suggested that the proteins encoded by the cluster may be ion channels or adaptor proteins (Liang et al., 2001; Zuccolo et al., 2010). It is likely that the genes in the cluster all have a similar role due to their high homology (Liang et al., 2001). However, until more is known of their function it is impossible to speculate as to the potential role of polymorphisms in these genes in AD aetiology. Whereas **MS4A6A** was detected in all brain regions, using microarrays, we were unable to confidently detect **MS4A4A** in cerebellum, frontal cortex or occipital cortex.

**ABCA7**—ATP-binding cassette, subfamily A (ABC1), member 7; chromosome 19p13.3

Two SNPs in **ABCA7** have been associated with LOAD: rs3752246 and rs3764650 (Hollingworth et al., 2011; Naj et al., 2011). rs3752246, in exon 32 of the gene, leads to a protein change (p.Gly1527Ala) which was the only coding non-synonymous change to be identified by GWAS. In an attempt to identify an associated functional variant at the **ABCA7** locus, Hollingworth et al. chose to genotype rs3752246 in an
Initial Assessment of AD Risk Loci

**Figure 4** Association between genotypes and CpG sites in cerebellum. Here the results are shown as log-transformed P values colour-coded to match the CpG probe in the loci. The regions associated in chromosome 11 are represented: in the top panel the MS4A6E locus and in the bottom panel the ME3 region. From these, only rs10897024 in PICALM (associated with a P value of $4.65 \times 10^{-6}$ with the probe cg07560096 in ME3, represented in green) is in moderate LD with the GWAS hits for PICALM (rs561655, rs3851179 and rs541458).

Additional cohort because this was a non-synonymous SNP with the highest LD with rs3764650 of all HapMap ABCA7-coding variants based on $r^2$ values ($r^2 = 0.36$, $D^2 = 0.89$; Hollingworth et al., 2011). Although, in our study, we established a lower degree of LD ($r^2 = 0.1$ between rs3764650 and rs3752246), this was also the highest LD found between rs3764650 and all the coding SNPs studied here (Fig. 5). This suggests that the risk of AD conferred by the presence of rs3764650 is not due to the presence of multiple other common coding SNPs in the gene. rs3764650 SNP is found at position 115 of the intron between exons 13 and 14. There is no evidence to suggest that rs3764650 has an effect on the expression of the gene (Hollingworth et al., 2011). rs3752246 is predicted in silico to be a benign variant (Polyphen-2 score = 0). This variant is in moderate LD ($r^2 = 0.6$) with rs4147934, which is also predicted to be non-pathogenic. Nonetheless, rs3752246 was the only missense change to be identified by GWAS and functional studies of the real impact of these two variants at the protein level should further elucidate if any of these are the real risk associated variants in ABCA7.

ABCA7 had low, but detectable expression in all CNS regions with no striking regional differences in gene expression patterns (data not shown).

The ABCA7 gene is a member of a large family of ATP-binding cassette genes divided into the subfamilies A–G based upon sequence homology (Kim et al., 2008). The transporters of the A and G subfamilies are particularly responsible for the movement of lipids such as sterols, phospholipids and bile acids across membranes against the concentration gradient of the substrate (Schmitz et al., 2000; Schmitz and Kaminski, 2001; Kusuhara and Sugiyama, 2007). ABCA1 is known to have a role in the transport of cholesterol to lipid-free acceptors such as apoA-I and apoE. The high homology between ABCA1 and ABCA7 suggests that the two proteins should share a similar role. The high lipid content of the CNS means that lipid homeostasis is essential and so, changes in the ability to transport cholesterol, are potentially pathogenic. It has been shown that the levels of cholesterol influence the processing of the APP protein: during times of high intracellular content, the activity of $\alpha$-secretase is inhibited, whilst the activity of $\beta$-and $\gamma$-secretases is enhanced (Bodovitz and Klein, 1996; Tun et al., 2002; Kalvodova et al., 2005; Vetivel et al., 2005). The role of ABCA7 in this situation is currently unclear as there is conflicting evidence regarding its ability to transport cholesterol: work with ABCA7−/− mice has shown that cholesterol efflux is not dependent upon ABCA7 (Kim et al., 2005);
Figure 5 Representation of \textit{ABCA7} with the localisation of the SNPs found as significant by GWAS as well as the SNPs studied here. On the right, linkage disequilibrium plot for the \textit{ABCA7} SNPs found to be significant in GWAS and the common coding variants genotyped in this study.
however work with human embryonic kidney (HEK)293 cells has shown that transfection with ABCA7 cDNA leads to 1.7–2 times increase in cholesterol efflux (Chan et al., 2008). It is likely therefore that ABCA7 plays a role in cholesterol transport, but that this is a process that also involves other proteins.

CLU—clusterin; chromosome 8p21-p12
The CLU gene encodes the clusterin or apolipoprotein J protein, expressed ubiquitously but with a known higher prevalence in the brain, ovary, testis and liver (de Silva et al., 1990). Levels of clusterin are shown to be elevated in the cortex and hippocampus areas of the brains of AD patients (Oda et al., 1994). Although we found high expression of CLU throughout the control CNS, we were unable to demonstrate that this gene was more highly expressed in cortex or hippocampus relative to other CNS regions (Fig. 6). Clusterin binds to Aβ plaques in the cerebrospinal fluid, forming a complex that is able to cross the blood–brain barrier (Zlokovic, 1996). Levels of clusterin in the plasma are positively correlated with the risk of AD (Schrijvers et al., 2011). It is therefore possible that the increased risk of AD induced by the SNPs described may be due to an increased level of expression.

We previously sequenced the exonic regions of CLU in 495 AD cases and 330 healthy controls and found a total of 24 variants in both cases and controls with similar frequencies between groups, indicating that common coding variability in this gene does not underlie the association seen with the intronic SNPs (Guerreiro et al., 2010). In order to determine if common variants at the CLU locus effect expression of nearby (cis) mRNA transcripts, an eQTL analysis was also performed. No significant eQTL associations were observed for the SNPs previously associated with AD, which led us to conclude that the most likely mechanism underpinning the association is either small effects of genetic variability on resting gene expression, or effects on damage induced expression of the protein. These conclusions are also supported by the absence of significant differences in the expression of the gene between several brain regions and in the gene splicing (Fig. 6). More recently, rare coding variants in CLU have been associated with the risk for AD (Bettens et al., 2012). However, this variability was found to be independent of the common association signal identified by the GWAS. Small studies have also reported a significant association between rs9331888 and alternative splicing of CLU (Szymanski et al., 2011) and blood clusterin levels (Xing et al., 2012).

Conclusion
Clearly, developing an understanding of the nature and mechanism of loci for AD (and other neurological diseases), which are identified by GWAS and are not coding changes is going to be a considerable challenge (Table 2). This study shows that simple eQTL studies in control brain tissue may not identify effects in many cases. There remain several options which are not mutually exclusive: (1) the study is underpowered (although larger than most previous studies) (2) genetic variability in splicing is an important consideration (3) genetic variability in other RNA species at the locus is important besides the obvious mRNA (4) genetic variability in damage induced expression, and not in resting expression, is the important factor.

The first three possibilities can be gradually overcome by either more samples or by improvements (for example) in sequencing technologies, which would allow transcript QTLs

Figure 6 Boxplot showing CLU expression across 12 human CNS regions, as measured using Affymetrix exon arrays in 137 neuropathologically normal individuals.
to be assessed as well as QTLs for other RNA species. Understanding QTLs in damage-induced expression is inherently difficult (Webster et al., 2009). Measuring changes in damage induced expression in tissue with changing cell populations and developing rigorous algorithms to interpret such data is problematic, but may be necessary, especially for understanding the aetiology of late onset neurodegenerative diseases.

**Alzheimer’s Disease Genetics Consortium coauthors and affiliations**

Liana G. Apostolova, MD1, Steven E. Arnold, MD2, Clinton T. Baldwin, PhD3, Robert Barber, PhD4, Michael M. Barmada, PhD3, Thomas G. Beach, MD, PhD6, Gary W. Beecham, PhD7,8, Duane Beekly, BS8, David A. Bennett, MD9,10,11, Eileen H. Bigio, MD12, Thomas D. Bird, MD3, Deborah Blacker, MD, PhD14,15, Bradley F. Boeve, MD16, James D. Bowen, MD17, Adam Boxer, MD, PhD18, James R. Burke, MD, PhD19, Jacqueline Buros, BA1, Joseph D. Buxbaum, PhD20,21,22, Nigel J. Cairns, PhD, FRCP, Patent23, Laura B. Cantwell, MPH24, Chuanhui Cao, PhD25, Chris S. Carlson, PhD26, Regina M. Carney, MD27, Minerva M. Carrasquillo, PhD28, Steven L. Carroll, MD, PhD29, Helena Chiu, MD30, David G. Clark, MD31, Carl W. Cotman, PhD32, Paul K. Crane, MD, MPH33, Elizabeth A. Crocco, MD34, Carlos Cruchaga, PhD35, Jeffrey L. Cummings, MD1, Philip L. De Jager, MD, PhD36,37, Charles DeCarli, MD38, Steven T. DeKosky, MD39, F. Yestin Demirci, MD39, Ramon Diaz-Arrastia, MD, PhD40, Malcolm Dick, PhD32, Dennis W. Dickson, MD28, Ranjan Duara, MD41, William G. Ellis, MD42, Nilüfer Ertekin-Taner, MD, PhD28,43, Denis Evans, MD44, Kelley M. Faber, MS45, Kenneth B. Fallon, MD29, Martin R. Farlow, MD46, Steven Ferris, PhD47, Tatiana M. Foroud, PhD48, Matthew P. Frosch, MD, PhD48, Douglas R. Galasko, MD49, Mary Ganguli, MD50, Marla Gearing, PhD51,52, Daniel H. Geschwind, MD, PhD53, Bernardino Ghetti, MD54, John R. Gilbert, PhD7,8, Sid Gilman, MD, FRCPath55, Bruno Giordani, PhD56, Jonathan D. Glass, MD57, Alison M. Goate, DPhil58, Neill R. Graff-Radford, MD28,43, Robert C. Green, MD, MPH59, John H. Growdon, MD60, Hakon Hakonarson, MD, PhD60, Ronald L. Hamilton, MD61, Lindy E. Harrell, MD, PhD61, Elizabeth Head, PhD62, Lawrence S. Honig, MD, PhD63, Christine M. Hulette, MD64, Bradley T. Hyman, MD, PhD65, Gail P. Jarvik, MD, PhD65,66, Gregory A. Jicha, MD, PhD67, Lee–Way Jin, MD, PhD62, Gyunghun Jun, PhD3,68,69, M. Ilyas Kamboh, PhD67,70, Jason Karlawish, MD71, Anna Karydas, BA18, John S.K. Kauwe, PhD72, Jeffrey A. Kaye, MD73,74, Ronald Kim, MD75, Edward H. Koo, MD76,77, Patricia Kramer, PhD78,79, Walter A. Kukull, PhD79, James J. Lah, MD, PhD80, Eric B. Larson, MD MPH81,82, Allan I. Levey, MD PhD83, Andrew P. Lieberman, MD PhD81, Oscar L. Lopez, MD70, Kathryn L. Lunetta, PhD84, Wendy J. Mack, PhD82, Daniel C. Marson, JD PhD31, Eden R. Martin, PhD7,8, Frank Martinuik, PhD93, Deborah C. Marsh, PhD84, Eliezer Masliah, MD94,85, Wayne C. McCormick, MD, MPH83, Susan M. McCurry, PhD86, Andrew N. McDavid, BA87, Ann C. McKeever, MD76,77, Marcell Mesulam, M.D87,88, Bruce L. Miller, MD19, Carol A. Miller, MD89, Joshua W. Miller, PhD82, Thomas J. Montine, MD PhD90, John C. Morris, MD23,91, Adam C. Naj, PhD7, Petra Nowotny, PhD35, Joseph E. Parisi, MD92,93, Elaine Pescind, MD94, Ronald C. Petersen, MD PhD16, Wayne W. Poon, PhD32, Huntington Potter, PhD25, Joseph F. Quinn, MD73, Ashok Raj, MD25, Ruchita A. Rajbhandary, MPH47, Murray Raskind, MD94, Barry Reisberg, MD47,55, Christiane Reitz, MD, PhD63,96,97, John M. Ringman, MD1, Erik D. Roberson, MD, PhD31, Ekaterina Rogaeva, PhD98, Roger N. Rosenberg, MD40, Mary Sano, PhD21, Andrew J. Saykin, PsyD45,99, Julie A. Schneider, MD100,10, Lon S. Schneider, MD10,101, William W. Seeley, MD18, Michael L. Shelanski, MD, PhD102, Charles D. Smith, MD67, Joshua A. Sonnen, MD30, Salvatore Spina, MD34, Peter St George-Hyslop, MD, FRCP98,103, Robert A. Stern, PhD97, Rudolph E. Tanzi, PhD95, John Q. Trojanowski, MD, PhD24, Juan C. Troncoso, MD104, Debby W. Tsuang, MD94, Otto Valladares, MS24, Viviana M. Van Deelen, MD, PhD24, Badri N. Vardarajan, MS3, Harry V. Vinters, MD110, Jean Paul Vonsattel, MD106, Li-San Wang, PhD24, Sandra Weintraub, PhD97,88, Kathleen A. Welsh-Bohmer, PhD39,107, Jennifer Williamson, MS48, Randall L. Wolter, MD, PhD108, Clinton B. Wright, MD, MS109, Steven G. Younkin, MD, PhD28, 1Department of Neurology, University of California Los Angeles, Los Angeles, California; 2Department of Psychiatry, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania; 3Department of Medicine (Genetics Program), Boston University, Boston, Massachusetts; 4Department of Pharmacology and Neuroscience,
University of North Texas Health Science Center, Fort Worth, Texas; 2Department of Human Genetics, University of Pittsburgh, Pittsburgh, Pennsylvania; 3Civin Laboratory for Neuropathology, Banner Sun Health Research Institute, Phoenix, Arizona; 4The John P. Hussman Institute for Human Genomics, University of Miami, Miami, Florida; 5Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami, Miami, Florida; 6National Alzheimer's Coordinating Center, University of Washington, Seattle, Washington; 7Department of Neurological Sciences, Rush University Medical Center, Chicago, Illinois; 8Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, Illinois; 9Department of Pathology, Northwestern University, Chicago, Illinois; 10Department of Neurology, University of Washington, Seattle, Washington; 11Department of Neurology, Harvard School of Public Health, Boston, Massachusetts; 12Department of Psychiatry, Massachusetts General Hospital/Harvard Medical School, Boston, Massachusetts; 13Department of Neurology, Mayo Clinic, Rochester, Minnesota; 14Department of Pathology, University of California San Francisco, San Francisco, California; 15Department of Medicine, Duke University, Durham, North Carolina; 16Department of Neuroscience, Mount Sinai School of Medicine, New York, New York; 17Department of Psychiatry, Mount Sinai School of Medicine, New York, New York; 18Departments of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, New York; 19Department of Pathology and Immunology, Washington University, St. Louis, Missouri; 20Department of Pathology and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania; 21Byrd Alzheimer Institute, University of Southern Florida Health, Tampa, Florida; 22Fred Hutchinson Cancer Research Center, Seattle, Washington; 23Department of Psychiatry, Vanderbilt University, Nashville, Tennessee; 24Department of Neuroscience, Mayo Clinic, Jacksonville, Florida; 25Department of Pathology, University of Alabama at Birmingham, Birmingham, Alabama; 26Department of Neurology, University of Southern California, Los Angeles, California; 27Department of Neurology, University of Alabama at Birmingham, Birmingham, Alabama; 28Institute for Memory Impairments and Neurological Disorders, University of California Irvine, Irvine, California; 29Department of Medicine, University of Washington, Seattle, Washington; 30Department of Psychiatry and Behavioral Sciences, Miller School of Medicine, University of Miami, Miami, Florida; 31Department of Psychiatry and Hope Center Program on Protein Aggregation and Neurodegeneration, Washington University School of Medicine, St. Louis, Missouri; 32Program in Translational Neuropsychiatric Genomics, Institute for the Neurosciences, Department of Neurology & Psychiatry, Brigham and Women's Hospital, and Harvard Medical School, Boston, Massachusetts; 33Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts; 34Department of Neurology, University of California Davis, Sacramento, California; 35University of Virginia School of Medicine, Charlottesville, Virginia; 36Department of Neurology, University of Texas Southwestern, Dallas, Texas; 37Wien Center for Alzheimer's Disease and Memory Disorders, Mount Sinai Medical Center, Miami Beach, Florida; 38Department of Pathology and Laboratory Medicine, University of California Davis, Sacramento, California; 39Department of Neurology, Mayo Clinic, Jacksonville, Florida; 40Rush Institute for Healthy Aging, Department of Internal Medicine, Rush University Medical Center, Chicago, Illinois; 41Department of Medical and Molecular Genetics, Indiana University, Indianapolis, Indiana; 42Department of Neurology, Indiana University, Indianapolis, Indiana; 43Department of Psychiatry, New York University, New York, New York; 44C.S. Kubik Laboratory for Neuropathology, Massachusetts General Hospital, Charlestown, Massachusetts; 45Department of Neurosciences, University of California San Diego, La Jolla, California; 46Department of Psychiatry, University of Pittsburgh, Pittsburgh, Pennsylvania; 47Department of Pathology and Laboratory Medicine, Emory University, Atlanta, Georgia; 48Emory Alzheimer's Disease Center, Emory University, Atlanta, Georgia; 49Neurogenetics Program, University of California Los Angeles, Los Angeles, California; 50Department of Pathology and Laboratory Medicine, Indiana University, Indianapolis, Indiana; 51Department of Neurology, University of Michigan, Ann Arbor, Michigan; 52Department of Psychiatry, University of Michigan, Ann Arbor, Michigan; 53Department of Neurology, Emory University, Atlanta, Georgia; 54Division of Genetics, Department of Medicine and Partners Center for Personalized Genetic Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts; 55Department of Neurology, Massachusetts General Hospital/Harvard Medical School, Boston, Massachusetts; 56Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania; 57Department of Pathology (Neuropathology), University of Pittsburgh, Pittsburgh, Pennsylvania; 58Department of Biostatistics, University of Pittsburgh, Pittsburgh, Pennsylvania; 59Department of Pathology, Duke University, Durham, North Carolina; 60Department of Genome Sciences, University of Washington, Seattle, Washington; 61Department of Medicine (Medical Genetics), University of Washington, Seattle, Washington; 62Taub Institute on Alzheimer's Disease and the Aging Brain, Department of Neurology, Columbia University, New York, New York; 63Department of Pathology, Duke University, Durham, North Carolina; 64Department of Neuroscience, University of California San Diego, La Jolla, California; 65Department of Neurology, University of Alabama at Birmingham, Birmingham, Alabama; 66Department of Psychiatry, Emory University, Atlanta, Georgia; 67Department of Pathology, University of Michigan, Ann Arbor, Michigan; 68Department of Neurology, University of Southern California, Los Angeles, California; 69Department of Neurology, University of Virginia School of Medicine, Charlottesville, Virginia; 70Department of Biochemistry, University of Virginia, Charlottesville, Virginia; 71Department of Neurology, Universit...
Ophthalmology, Boston University, Boston, Massachusetts; University of Pittsburgh Alzheimer's Disease Research Center, Pittsburgh, Pennsylvania; Department of Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania; Department of Biology, Brigham Young University, Provo, Utah; Department of Neurology, Oregon Health & Science University, Portland, Oregon; Department of Neurology, Portland Veterans Affairs Medical Center, Portland, Oregon; Department of Pathology and Laboratory Medicine, University of California Irvine, Irvine, California; Department of Neurology, Boston University, Boston, Massachusetts; Department of Pathology, Boston University, Boston, Massachusetts; Department of Molecular & Medical Genetics, Oregon Health & Science University, Portland, Oregon; Department of Epidemiology, University of Washington, Seattle, Washington; Group Health Research Institute, Group Health, Seattle, Washington; Department of Pathology, University of Michigan, Ann Arbor, Michigan; Department of Preventive Medicine, University of Southern California, Los Angeles, California; Department of Medicine – Pulmonary, New York University, New York, New York; Department of Neurology, University of Miami, Miami, Florida; Department of Pathology, University of California San Diego, La Jolla, California; School of Nursing Northwest Research Group on Aging, University of Washington, Seattle, Washington; Alzheimer's Disease Center, Northwestern University, Chicago, Illinois; Cognitive Neurology, Northwestern University, Chicago, Illinois; Department of Pathology, University of Southern California, Los Angeles, California; Department of Pathology, University of Washington, Seattle, Washington; Department of Neurology, Washington University, St. Louis, Missouri; Department of Anatomic Pathology, Mayo Clinic, Rochester, Minnesota; Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota; Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, Washington; Alzheimer's Disease Center, New York University, New York, New York; Gertrude H. Sergievsky Center, Columbia University, New York, New York; Department of Neurology, Columbia University, New York, New York; Tanz Centre for Research in Neurodegenerative Disease, University of Toronto, Toronto, Ontario; Department of Radiology and Imaging Sciences, Indiana University, Indianapolis, Indiana; Department of Pathology (Neuropathology), Rush University Medical Center, Chicago, Illinois; Department of Psychiatry, University of Southern California, Los Angeles, California; Department of Pathology, Columbia University, New York, New York; Cambridge Institute for Medical Research and Department of Clinical Neurosciences, University of Cambridge, Cambridge, United Kingdom; Department of Pathology, Johns Hopkins University, Baltimore, Maryland; Department of Pathology & Laboratory Medicine, University of California Los Angeles, Los Angeles, California; Taub Institute on Alzheimer's Disease and the Aging Brain, Department of Pathology, Columbia University, New York, New York; Department of Psychiatry & Behavioral Sciences, Duke University, Durham, North Carolina; Department of Pathology, Oregon Health & Science University, Portland, Oregon; Evelyn F. McKnight Brain Institute, Department of Neurology, Miller School of Medicine, University of Miami, Miami, Florida.

Acknowledgements

The authors have no conflicts of interest regarding this study. This work was supported in part by the Intramural Research Program of the National Institute of Health, National Institute on Aging (Z01-AG000950–10), and utilized the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, MD (http://biowulf.nih.gov). Genotyping of the TGEN2 cohort was supported by Kronos Science; The National Institute on Aging [R01 AG034504, R01 AG031581, P30 AG19610, Z01 AG000950–06, P30AG10161, R01AG15819]; The National Institute of Neurological Disorders and Stroke [R01 NS059873] The Banner Alzheimer’s Foundation; The Johnnie B. Byrd Sr. Alzheimer’s Disease Institute; The Medical Research Council; and the state of Arizona. Many data and biomaterials were collected from several National Institute on Aging (NIA) and National Alzheimer’s Coordinating Center (NACC, grant #U01 AG016976) funded sites. This work was also partially supported by the MRC through the MRC Sudden Death Brain Bank. MR was supported by an MRC Training Fellowship and DT was supported by King Faisal Hospital.

We thank the brain donors and their families for their selfless donation to the fight against this disease. Amanda J. Myers, PhD (University of Miami, Department of Psychiatry) and John A. Hardy, PhD (Reta Lila Weston Institute, University College London) collected and prepared the series. Marcelle Morrison-Bogorad, PhD, Tony Phelps, PhD and Walter Kukull, PhD are thanked for helping to co-ordinate this collection. The directors, pathologist and technicians involved include: University of Michigan (NIH grant P50-AG08671): Dr. Roger Albin, Lisa Bain, Eszter Gombosi, The Netherlands Brain Bank (funding via numerous sources including Stichting MS Research, Brain Net Europe, Hersenstichting Nederland Breinbrekend Werk, International Parkinson Fonds, Internationale Stichting Alzheimer Onderzoek): Inge Huittinga, MD, Marleen Rademaker, Michiel Kooreman. We
thank Colin Smith and Robert Walker at the Sudden Death Brain Bank for their help.

The National Institutes of Health, National Institute on Aging (NIH-NIA) supported the ADGC through the following grants: ADGC, U01 AG032984, RC2 AG036528; NACC, U01 AG016976; NCRAD, U24 AG021886; NIA LOAD, U24 AG026395, U24 AG026390; Banner Sun Health Research Institute P30 AG019610; Boston University, P30 AG013846, U01 AG10483, R01 CA129769, R01 MH080295, R01 AG017173, R01 AG025259, R01AG33193; Columbia University, P50 AG008702, R37 AG015473; Duke University, P30 AG028377, AG05128; Emory University, AG025688; Group Health Research Institute, U01 AG06781, U01 HG004610; Indiana University, P30 AG10133; Johns Hopkins University, P50 AG005146, R01 AG020688; Massachusetts General Hospital, P50 AG005134; Mayo Clinic, P50 AG016574; Mount Sinai School of Medicine, P50 AG005138, P01 AG02219; New York University, P30 AG00851, MO1RR00096, and UL1 RR029893; Northwestern University, P30 AG013854; Oregon Health & Science University, P30 AG008017, R01 AG026916; Rush University, P30 AG010161, R01 AG019085, R01 AG15819, R01 AG17917, R01 AG30146; TGen, R01 NS059873; University of Alabama at Birmingham, P50 AG016582, UL1RR02777; University of Arizona, R01 AG031581; University of California, Davis, P30 AG010129; University of California, Irvine, P50 AG016573, P50, P50 AG016575, P50 AG016576, P50 AG016577; University of California, Los Angeles, P50 AG016570; University of California, San Diego, P50 AG005131; University of California, San Francisco, P50 AG023501, P01 AG019724; University of Kentucky, P30 AG028383, AG05144; University of Michigan, P50 AG008671; University of Pennsylvania, P30 AG010124; University of Pittsburgh, P50 AG005133, AG030653; University of Southern California, P50 AG005142; University of Texas Southwestern, P30 AG012300; University of Miami, R01 AG027944, AG010491, AG027944, AG021547, AG019757; University of Washington, P50 AG005136; Vanderbilt University, R01 AG019085; and Washington University, P50 AG005681, P01 AG03991. The Kathleen Price Bryan Brain Bank at Duke University Medical Center is funded by NINDS grant # NS39764, NIMH MH60451 and by Glaxo Smith Kline. We thank Drs. D. Stephen Snyder and Marilyn Miller from NIA who are ex-officio ADGC members. Support was also from the Alzheimer's Association (LAF, IIRG-08–89720; MP-V, IIRG-05–14147) and the US Department of Veterans Affairs Administration, Office of Research and Development, Biomedical Laboratory Research Program. P.S.G.-H. is supported by Wellcome Trust, Howard Hughes Medical Institute, and the Canadian Institute of Health Research.

ADNI Funding for ADNI is through the Northern California Institute for Research and Education by grants from Abbott, AstraZeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson and Johnson, Eli Lilly and Co., Medpace, Inc., Merck and Co., Inc., Novartis AG, Pfizer Inc, F Hoffman-La Roche, Schering-Plough, Synarc, Inc., Alzheimer's Association, Alzheimer's Drug Discovery Foundation, the Dana Foundation, and by the National Institute of Biomedical Imaging and Bioengineering and NIA grants U01 AG024904, RC2 AG036535, K01 AG030514.

References


P. Holton et al.


**Supporting Information**

Additional supporting information may be found in the online version of this article:

**Table S1** SNPs found associated with LOAD by GWAS assessing >1500 cases and >1500 controls.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organised for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

**Received:** 19 January 2012  
**Accepted:** 5 March 2012

© 2013 Blackwell Publishing Ltd/University College London  
105