*Human Molecular Genetics, 2010, Vol. 19, No. 16* **3295–3301**
*doi:10.1093/hmg/ddq221*
*Advance Access published on June 9, 2010*

# Association of *CR1*, *CLU* and *PICALM* with Alzheimer's disease in a cohort of clinically characterized and neuropathologically verified individuals

**Jason J. Corneveaux[1,2,†], Amanda J. Myers[3,4,5,6,†], April N. Allen[1,2], Jeremy J. Pruzin[1,2], Manuel Ramirez[3,4,5], Anzhelika Engel[3,4,5], Michael A. Nalls[7], Kewei Chen[2,12], Wendy Lee[2,12], Kendria Chewning[1,2], Stephen E. Villa[1,2], Hunsar B. Meechoovet[1,2], Jill D. Gerber[1,2], Danielle Frost[1,2], Hollie L. Benson[1,2], Sean O'Reilly[1,2], Lori B. Chibnik[9], Joshua M. Shulman[9], Andrew B. Singleton[7], David W. Craig[1,2], Kendall R. Van Keuren-Jensen[1,2], Travis Dunckley[1,2], David A. Bennett[8], Philip L. De Jager[9], Christopher Heward[10], John Hardy[11], Eric M. Reiman[1,2,12] and Matthew J. Huentelman[1,2,*]**

[1]Neurogenomics Division, The Translational Genomics Research Institute (TGen), 445 N Fifth Street, Phoenix, AZ 85004, USA, [2]The Arizona Alzheimer's Consortium, 901 E. Willetta Street, Phoenix, AZ 85006, USA, [3]Laboratory of Functional Neurogenomics, [4]Department of Psychiatry and Behavioral Sciences, Division of Neuroscience, [5]Department of Human Genetics and Genomics, University of Miami, Miller School of Medicine, Rosenstiel Medical Science Building, Room 6152A, 1600 NW 10th Avenue, Miami, FL 33136, USA, [6]Johnnie B. Byrd Sr. Alzheimer's Center and Research Institute, 4001 E. Fletcher Avenue, Tampa, FL 33613, USA, [7]Molecular Genetics Section, Laboratory of Neurogenetics, Intramural Research Program, National Institute on Aging NIH, Bethesda, MD 20892, USA, [8]Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, IL 60612, USA, [9]Program in Translational NeuroPsychiatric Genomics, Department of Neurology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA, [10]Kronos Science, 2222 E. Highland, Suite 220, Phoenix, AZ 85016, USA, [11]Department of Molecular Neuroscience and Reta Lila Research Laboratories, UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK and [12]Banner Alzheimer's Institute, 901 E. Willetta St., Phoenix, AZ 85006, USA

Received April 13, 2010; Revised May 19, 2010; Accepted May 25, 2010

**In this study, we assess 34 of the most replicated genetic associations for Alzheimer's disease (AD) using data generated on Affymetrix SNP 6.0 arrays and imputed at over 5.7 million markers from a unique cohort of over 1600 neuropathologically defined AD cases and controls (1019 cases and 591 controls). Testing the top genes from the AlzGene meta-analysis, we confirm the well-known association with *APOE* single nucleotide polymorphisms (SNPs), the *CLU, PICALM* and *CR1* SNPs recently implicated in unusually large data sets, and previously implicated *CST3* and *ACE* SNPs. In the cases of *CLU, PICALM* and *CR1*, as well as in *APOE*, the odds ratios we find are slightly larger than those previously reported in clinical samples, consistent with what we believe to be more accurate classification of disease in the clinically characterized and neuropathologically confirmed AD cases and controls.**

*To whom correspondence should be addressed. Tel: +1 6023438730; Fax: +1 6023438844; Email: mhuentelman@tgen.org
†The authors wish it to be known that in their opinion these individuals contributed equally to the described work.

*© The Author 2010. Published by Oxford University Press. All rights reserved.*
*For Permissions, please email: journals.permissions@oxfordjournals.org*

## INTRODUCTION

Alzheimer's disease (AD) is the most common form of dementia. Estimates suggest that ~10% of individuals over the age of 65 and almost half over the age of 85 have Alzheimer's dementia (1). As human life expectancy continues to rise, the total number of afflicted individuals is expected to also increase dramatically. Chronological age is the most significant risk factor for AD. However, twin studies suggest that genetics plays a crucial role in altering an individual's risk for developing the disease contributing between 37 and 78% of the variation in the age of onset (2).

An early onset form of AD (EOAD) typically develops before the age of 65 and is often linked to autosomal dominant mutations in one of three known genes (*PSEN1*, *PSEN2* and *APP*) (3). The more common late onset form (LOAD) of the disease is more complex in nature with proposed combined genetic and environmental risk factors. The most well-replicated genetic association for LOAD is the ε4 variant of the APOE gene (4,5); however, it has been estimated that variation at the *APOE* locus may account for 20% or less of LOAD risk (6–10). With the advent and ongoing improvement of genome-scanning technologies, the search for the remaining genetic risk factors for LOAD is still ongoing. Recently, two large studies provided some of the strongest evidence to date for three additional loci (*CLU*, *PICALM* and *CR1*) (11,12).

To aggregate the majority of genetic data related to AD, Bertram *et al*. (13) created and currently maintain a database of association studies performed for the disease termed AlzGene. The goal of AlzGene is to provide guidance to the field regarding the rank-ordered importance of purported AD-associated genetic polymorphisms. In this study, we used whole genome single nucleotide polymorphism (SNP) data to assess the top associations within AlzGene in our unique collection of clinically characterized and neuropathologically defined cases and controls. We previously reported on findings from this collection (14–16); however, this new study includes additional individuals, was performed on a higher density SNP array and leveraged genome-wide imputation using the current 1000 Genomes data set. We tested the SNPs from the newest AlzGene data freeze (February 1, 2010) and analyzed only those 38 SNPs from the 'Top Results' listing of genes with Overall Grades of either A, B or C. Using imputation, we were able to genotype 34 out of the 38 SNPs. Six of the tested SNPs demonstrated association with AD risk ($P < 0.05$) including the recently identified *CLU*, *CR1* and *PICALM* SNPs. We also replicate the associations with rs6907175 (*LOC651924*), rs1800764 (*ACE*) and rs1064039 (*CST3*).

## RESULTS

This study analyzed genome-wide association data from 1610 clinically and neuropathologically well-characterized expired brain donors, including 1019 cases (652 females, 367 males) with a clinical diagnosis of dementia and neuropathologically confirmed AD (Braak stage V or VI) and 591 cognitively normal persons (285 females, 306 males) without neuropathologic AD (Braak stage <III). Age at death for controls was 80.7 ± 8.7 and 82.0 ± 7.7 (mean ± SD). *APOE* genotypes are shown in Table 1. The ORs observed for *APOE* in this postmortem cohort are significantly higher than those in studies of clinically characterized subjects. For instance, in this larger neuropathological sample, as well as in our previous neuropathological sample (14), comparing to *ε3* homozygotes, *ε4* homozygotes had an OR of about 20, compared with an OR of about 13 in at the AlzGene site.

Assessment of the other significant SNPs from the AlzGene list is shown in Table 2. Six loci reach nominal significance in this analysis: *CR1, LOC651924, CLU, PICALM, ACE* and *CST3*. We note that we have previously reported on *ACE* in a proportion of this data set (16) and further replicated that association, but none of the other loci have been previously assessed in this data set. In all the cases, the direction of the association in this study is the same as is reported on AlzGene. It is notable that three of the five other confirmed loci are those identified in the two recent large genome-wide association studies (11,12) and this study therefore supports the application of this approach to late onset AD.

Interpretation of the analysis of associations we report here in the presence or absence of the *APOE ε4* allele is hampered by the reduction power due to reduction of the number of cases and controls in each group and by the increase in the amount of multiple testing. We therefore carried out only exploratory analysis in *APOE ε4+* cases and controls, in *APOE ε4−* cases and controls and found no significant associations (data not shown). We additionally conducted an epistatic analysis in which *APOE ε4* was included as a covariate (Supplementary Material, Table S1). These analyses suggested that *CCR2* (rs1799864: A allele) showed an epistatic interaction with *APOE ε4* (nominal $P$-value = 0.024) and was apparently a risk factor only in the presence of an *APOE* ε4 allele.

## DISCUSSION

In this study, we present data from a large neuropathologically verified cohort for the top reported genetic risk factors for late onset AD. We leveraged neuropathological phenotyping and strict sample quality control, which we argue provides additional power to detect subtle associations due to improved classification of AD cases and controls (14). As shown in Table 1, the odds ratios observed for the *APOE* epsilon variants in this postmortem cohort are markedly higher than those in studies of clinically characterized subjects. This is likely due in part to the fact that *APOE* is more strongly related to neuropathologic phenotypes than to clinical phenotypes (17) and the study design reduced both the numbers of false positive and false negative cases and controls. Additionally, even though the initial studies often suffer from the 'winner's curse' and subject to inflated odds ratios, the odds ratios we observed in this study are equal to or marginally greater than the odds ratios observed in the original studies, although the 95% confidence intervals overlap. Again, we suspect that the marginally higher odds ratios are a result of better case and control definition.

We recently sequenced the *CLU* locus and showed that, in contrast to the case of *APOE*, there is no common coding

**Table 1.** *APOE* genotypes

| APOE | Cases (%) | Controls (%) | OR (95% CI) | AlzGene cases (%) | AlzGene controls (%) | AlzGene OR (95% CI) |
|---|---|---|---|---|---|---|
| ε2ε2 | 4 (0.4) | 19 (3.2) | 0.24 (0.08–0.7) | 7 (0.3) | 44 (0.9) | 0.5 (0.22–1.1) |
| ε2ε3 | 32 (3.1) | 68 (12) | 0.5 (0.3–0.9) | 132 (4.9) | 602 (12) | 0.7 (0.6–0.8) |
| ε2ε4 | 37 (3.6) | 13 (2.2) | 3.2 (1.7–6.2) | 75 (2.8) | 106 (2.1) | 2.3 (1.6–3) |
| ε3ε3 | 329 (32) | 379 (65) | 1.0 | 969 (36) | 3039 (60) | 1.0 |
| ε3ε4 | 455 (45) | 95 (16) | 5.5 (4.2–7.2) | 1210 (45) | 1122 (22) | 3.4 (3–3.8) |
| ε4ε4 | 161 (16) | 9 (1.5) | 20.6 (10.4–41) | 407 (15) | 99 (2) | 12.9 (10.2–16.2) |
| Total | 1018 | 583 | $n = 1601$ | 2712 | 5100 | $n = 7812$ |

Apoliprotein E genotyping from this neuropathological series compared with the AlzGene (February 2010 freeze) data from Caucasian clinical series. APOE ε3ε3 used as neutral reference for contingency table.

**Table 2.** Assessment of AlzGene hits in the current data set

| Chr. | SNP | Gene | Position | Allele associated | *P*-value | OR (95% CI) | MAF cases | MAF controls (%) | MAF AlzGene cases (%) | MAF AlzGene controls (%) |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | rs1801133 | MTHFR | 11778965 | A | 0.832 | 0.98 (0.84–1.15) | 34.7 | 35.4 | 38.0 | 39.0 |
| 1 | rs4845378 | CHRNB2 | 152811275 | T | 0.827 | 1.03 (0.82–1.28) | 11.7 | 11.8 | 7.0 | 10.0 |
| 1 | rs6656401 | CR1 | 205758672 | A | **0.008** | 1.28 (1.07–1.54) | 21.1 | 17.5 | 22.0 | 18.0 |
| 2 | rs1800587 | IL1A | 113259431 | A | 0.440 | 0.94 (0.80–1.10) | 30.0 | 31.5 | 31.0 | 30.0 |
| 2 | rs1143634 | IL1B | 113306861 | A | 0.300 | 0.91 (0.77–1.08) | 22.9 | 24.7 | 28.0 | 26.0 |
| 3 | rs1049296 | TF | 134977044 | T | 0.242 | 1.12 (0.92–1.37) | 17.5 | 16.3 | 18.0 | 17.0 |
| 6 | rs760678 | NEDD9 | 11442640 | C | 0.852 | 0.99 (0.85–1.15) | 40.5 | 41.0 | 39.0 | 41.0 |
| 6 | rs3800324 | PGBD1 | 28372660 | A | 0.159 | 1.30 (0.90–1.87) | 5.0 | 3.8 | 5.0 | 4.0 |
| 6 | rs6907175 | LOC651924 | 142425775 | A | **0.027** | 1.18 (1.02–1.36) | 50.8 | 47.0 | 50.0 | 54.0 |
| 8 | rs11136000 | CLU | 27520436 | T | **0.040** | 0.86 (0.74–0.99) | 36.1 | 39.9 | 36.0 | 40.0 |
| 9 | rs11792633 | IL33 | 6238035 | T | 0.928 | 0.99 (0.85–1.17) | 29.4 | 29.9 | 35.0 | 36.0 |
| 9 | rs4878104 | DAPK1 | 89382811 | T | 0.334 | 1.08 (0.93–1.26) | 35.6 | 33.9 | 35.0 | 38.0 |
| 10 | rs2306604 | TFAM | 59818698 | G | 0.808 | 1.02 (0.88–1.18) | 43.9 | 43.6 | 41.0 | 44.0 |
| 10 | rs13500 | CH25H | 90963472 | A | 0.162 | 1.18 (0.94–1.48) | 12.4 | 11.1 | 12.0 | 10.0 |
| 10 | rs498055 | LOC439999 | 97344904 | C | 0.820 | 0.98 (0.85–1.14) | 47.8 | 48.7 | 52.0 | 48.0 |
| 10 | rs911541 | ENTPD7 | 101423382 | G | 0.929 | 0.99 (0.79–1.23) | 12.2 | 12.4 | 13.0 | 13.0 |
| 10 | rs2986017 | CALHM1 | 105208242 | A | 0.314 | 0.92 (0.78–1.08) | 24.2 | 26.0 | 28.0 | 25.0 |
| 10 | rs600879 | SORCS1 | 108913108 | T | 0.712 | 1.05 (0.82–1.34) | 9.8 | 9.3 | 11.0 | 9.0 |
| 10 | rs1903908 | hCG2039140 | 109191662 | A | 0.951 | 1.01 (0.81–1.25) | 12.5 | 12.2 | 15.0 | 12.0 |
| 11 | rs6265 | BDNF | 27636492 | T | 0.433 | 0.93 (0.78–1.11) | 19.9 | 21.1 | 22.0 | 21.0 |
| 11 | rs10793294 | GAB2 | 77674051 | C | **0.025** | 0.82 (0.69–0.98) | 19.7 | 23.9 | 20.0 | 24.0 |
| 11 | rs541458 | PICALM | 85465999 | C | **0.010** | 0.81 (0.69–0.95) | 27.7 | 32.3 | 29.0 | 31.0 |
| 11 | rs12285364 | SORL1 | 120898436 | T | 0.662 | 1.09 (0.75–1.56) | 4.4 | 4.2 | 5.0 | 4.0 |
| 14 | rs165932 | PSEN1 | 72734606 | G | 0.121 | 0.89 (0.77–1.03) | 44.5 | 47.2 | 44.0 | 45.0 |
| 14 | rs11622883 | GWA_14q32.13 | 94225529 | A | 0.339 | 1.07 (0.93–1.24) | 45.8 | 43.8 | 42.0 | 47.0 |
| 17 | rs1554948 | TNK1 | 7227050 | A | 0.569 | 0.96 (0.83–1.11) | 47.3 | 48.1 | 45.0 | 50.0 |
| 17 | rs939348 | THRA | 35485379 | T | 0.757 | 0.97 (0.82–1.15) | 25.7 | 26.1 | 29.0 | 27.0 |
| 17 | rs1800764 | ACE | 58904261 | C | **0.030** | 0.85 (0.73–0.98) | 43.5 | 47.3 | 41.0 | 47.0 |
| 19 | rs4806173 | GAPDHS | 40716765 | G | 0.509 | 0.95 (0.82–1.10) | 37.2 | 38.9 | 36.0 | 39.0 |
| 20 | rs1064039 | CST3 | 23566427 | T | **0.033** | 0.83 (0.69–0.98) | 19.8 | 23.1 | 21.0 | 19.0 |

Uncorrected *P*-values and ORs ±95% CI were calculated logistic regression model with gender and population structure as covariates.
SNPs in bold have *P*-values less than 0.05.

variability which underlies this association (18). We interpret these data as suggesting that it is genetic variability in either resting or induced *CLU* expression which is critical to AD pathogenesis. At the *CR1* locus, we find a similar significantly associated haplotype ($P < 1 \times 10^{-4}$ after 100,000 Max (T) permutations) to that reported by Lambert *et al.* (12). Within the 100 kb haplotype, there are currently 26 missense SNPs reported in dbSNP130, and over 60 for the entire gene suggesting a possible importance of coding variability in this gene for AD risk.

With this confirmatory report, and in the absence of any current negative reports, we feel it is now appropriate to suggest that there are four loci for late onset AD: *APOE, CR1, PICALM* and *CLU*. Additionally, we believe that *ACE* should be considered very close to being additionally declared. The role of *CLU* and *APOE* in cholesterol metabolism and *CLU* and *CR1* in the complement cascade point to these biological processes as potentially important biochemical pathways involved in AD pathogenesis. There is prior mechanistic support for the role of these processes in neurodegenerative disease in general and AD (19,20). ACE is known to participate in blood pressure regulation and a link between hypertension and cognition, executive function and neurodegeneration has also been suggested (21–23). The *CST3* gene

encodes the protein cystatin C. The cystatin superfamily includes proteins that function as extracellular cysteine protease inhibitors and the biological mechanism of altered AD risk via cystatin C has been suggested to be due to altered metabolism of soluble β-amyloid species (24). It is possible that these genetic data are further suggestive of a role for the innate immune and vascular systems in AD etiology and point to the blood brain interface as the site where damage may be initiated.

## MATERIALS AND METHODS

### Clinically characterized and neuropathologically verified subjects

Our US series was obtained from 21 National Institute on Aging-supported Alzheimer's Disease Center brain banks and from the Miami Brain Bank as previously described (25,26). Additional cohorts from other brain banks in the United States, United Kingdom, and the Netherlands (see Acknowledgements) were obtained in the same manner as the original US series. Our criteria for inclusion were as follows: self-defined ethnicity of European descent (in an attempt to control for the known allele frequency differences between ethnic groups), neuropathologically confirmed AD or no neuropathology present, and age of death greater than 65. Neuropathological diagnosis was defined by board-certified neuropathologists based on the presence or absence of the clinical diagnosis of probable or possible AD, Braak and Braak staging to reflect the spatial extent of neurofibrillary tangles, and/or CERAD classification to reflect frequency of cortical neuritic plaques. Samples derived from subjects with a clinical history of stroke, cerebrovascular disease, Lewy bodies or comorbidity with any other known neurological disease were excluded. AD or control neuropathology was confirmed by plaque and tangle assessment with 45% of the entire series undergoing Braak staging (27). Of the 1019 cases, 369 were included in our previous study: of the 591 controls, 298 were included in our previous study.

Samples were de-identified before receipt, and the study met human studies institutional review board and HIPPA regulations. This work is declared not human-subjects research and is IRB exempt under regulation 45 CFR 46. See the Funding section for a list of individual sites that contributed samples to this effort.

### APOE genotyping

APOE genotyping was performed either by the method of Crook *et al*. (28) or through the use of a fluorescence-based allele-specific PCR (also called PASA; PCR Amplification of Specific Alleles) on array tape (29) by PreventionGenetics (Marshfield, WI, USA).

### Genome-wide SNP genotyping

Genomic DNA samples were analyzed on the Genome-Wide Human SNP 6.0 Array (Affymetrix, Inc., Santa Clara, CA, USA) according to the manufacturer's protocols (Affymetrix Genome-Wide Human SNP Nsp/Sty 6.0 User Guide; Rev. 1, 2007). Before the initiation of the assay, 50 ng of genomic DNA from each sample was examined qualitatively on a 1% Tris-acetate-EDTA agarose gel for visual signs of degradation. Any degraded DNA samples were excluded from further analysis (~3%). Samples were quantitated by OD Spectrometry and diluted to 50 ng/μl in reduced EDTA TE buffer (10 mM Tris–HCl, 0.1 mM EDTA, pH 8.0). Two hundred and fifty nanogram of DNA was then aliquoted into two 96-well reaction plates and digested in either Sty or Nsp restriction enzymes (New England Biolabs, Inc. Ipswich, MA, USA) for 2 h at 37°C followed by 65°C for 20 min. Sty and Nsp digested samples were then ligated to either the Sty 1 or the Nsp 1 adaptor (Affymetrix), respectively, with T4 DNA Ligase (New England Biolabs) for 3 h at 16°C then 20 min at 70°C. The ligated samples were then diluted in molecular-grade water and subaliquoted into three (Sty) or four (Nsp) 96-well PCR plates. PCR was performed using PCR Primer 002 (Affymetrix) and Titanium *Taq* DNA Polymerase (Clontech, Mountain View, CA, USA) with the following thermal cycling parameters: (i) 94°C for 3 min, (ii) 30 cycles of 94°C for 30 s, 60°C for 30 s and 68°C for 15 s and (iii) 68°C for 7 min. Like samples for all Sty and Nsp reactions were pooled into a single deep well plate, the DNA was bound to Agencourt AMPure beads (Beckman Coulter, Inc. Berea, CA, USA), placed into MultiScreen filter plates (Millipore, Billerica, MA, USA), washed with 75% ethanol and eluted with Buffer EB (QIAGEN, Valencia, CA, USA). Purified samples were then fragmented using Fragmentation Reagent (Affymetrix) and incubated at 37°C for 35 min then at 95°C for 15 min. Fragmented samples were labeled with DNA Labeling Reagent (Affymetrix) and TdT Enzyme (New England Biolabs) at 37°C for 4 h followed by 95°C for 15 min. The samples were denatured at 95°C for 10 min and held at 49°C until they were loaded on to the arrays. The arrays were placed into the hybridization oven at 50°C and 60 rpm for 16–18 h. Arrays were then washed, stained and immediately imaged on the GeneChip Scanner 3000 (Affymetrix).

### Targeted SNP analysis

Birdsuite (30) was used to call SNP genotypes from CEL files. Initial quality control measures consisted of gender-checks and a custom SNP fingerprinting approach to identify potentially duplicated or related individuals. After removing samples based on gender-errors (0.6% of cohort) and fingerprinting overlaps (1.5% of cohort), we applied additional quality control filters using PLINK v1.07 (31) and assessed thresholds using histograms and quartile calculations. We selected samples with SNP call rates of $\geq$85%, relatedness ($F$-values $< 0.04$) yielding 1024 cases and 595 controls.

Samples were also analyzed for genetic ancestry via ADMIXTURE software v1.02 (32). Autosomal SNPs ($n =$ 8664) with call rates >99%, minor allele frequency (MAF) >0.3, pairwise $R^2 < 0.01$ that were also genotyped in the HapMap (33) phase 3 populations were selected for this analysis. Using these SNPs, model-based estimation of genetic ancestry was calculated with ADMIXTURE software v1.02 with $K = 3$. We utilize Q1 and Q2 vector solutions from ADMIXTURE as covariates in subsequent regression analysis.

Case–control SNP analysis was carried out using PLINK v1.07 entering gender and ADMIXTURE Q1 and Q2 vectors as covariates into a logistic model. Minor allele frequencies in cases and controls were calculated using the Fisher's exact test in PLINK. APOE interactions were tested entering number of APOE ε4 copies (0, 1 or 2) as covariates into a linear model using the PLINK interaction option.

### SNP imputation

Genotypes were imputed using MACH v.1.0.16 (34) for all European ancestry participants in the study based on haplotypes released from initial low coverage sequencing of 112 European ancestry samples in the 1000 genomes project (ftp://ftp.sanger.ac.uk/pub/1000genomes/REL-0908/LowCov/). European ancestry for the imputation analysis was defined as samples which were within three standard deviations from the mean C1 and C2 values of the Toscan or CEPH (TSI/CEU) samples from HapMap V3 after undertaking multidimensional scaling (MDS) in PLINK including CEU/TSI, Yorban (YRI), and Japanese and Chinese (JPT/CHB) populations from the Hapmap V3. Nine samples were excluded based on population clustering. A subset of SNPs was employed for the MDS analysis excluding SNPs in high linkage disequilibrium ($R^2 \geq 0.8$), SNPs with minor allele frequencies $\leq 0.01$, SNPs with greater than 5% of subjects not called and SNPs where Hardy–Weinberg *P*-values were $\leq 0.0000001$.

The imputation was undertaken in two stages: first parameters for the imputation were calibrated based on a random sample of 200 individuals from this study and over 100 iterations of the model. Once parameter estimates were constructed, maximum likelihood genotypes were imputed for SNPs in the study population based on the reference haplotypes from the 1000 Genomes Project. These haplotypes yielded a maximum of ~8.2 million SNPs possible to impute. Results from the imputation were filtered to include only high-quality imputed SNPs with a minimum quality index of 0.30 based on the squared correlation between genotyped and imputed SNPs (RSQR from MACH). All reported association statistics herein are derived from imputed genotypes. As a quality metric, we report 99.5% concordance between the subset of 12 SNPs that were both imputed and directly genotyped on the Affymetrix SNP 6.0 array.

### Data sharing

Note that genotype data for all of the imputed AlzGene SNPs are freely available at the TGen Neurogenomics Data web site [www.tgen.org/data/neurogenomics].

## SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* Online.

## ACKNOWLEDGEMENTS

We would like to dedicate this work to our colleague C.H. who died while this study was being completed. We thank the brain donors and their families for their selfless donation to the fight against this disease.

*Conflict of Interest statement*. J.H. consults for MerckSerono and for Eisai. E.M.R. is a scientific advisor to AstraZeneca, Elan, Eli Lilly, Siemens and Sygnis/Amnestix and has had research contracts from AstraZeneca and GlaxoSmithKline for research unrelated to this study. The other authors report no conflicts of interest.

## FUNDING

This work was supported by Kronos Science; The National Institute of Neurological Disorders and Stroke (R01 NS059873 to M.J.H.); The National Institute on Aging (R01 AG031581 to E.M.R.; P30 AG19610 to E.M.R.; AG034504 to A.J.M.; Z01 AG000950-06, P30AG10161 and R01AG 15819 to D.A.B.; and K01AG024079 to T.D.); The Banner Alzheimer's Foundation; The Johnnie B. Byrd Sr. Alzheimer's Disease Institute; The Medical Research Council (to J.H.); The Intramural Research Program of the National Institutes of Health; and the State of Arizona. Biomaterials were collected from several National Institute on Aging and National Alzheimer's Coordinating Center (U01 AG016976) funded sites. 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: National Institute on Aging; Ruth Seemann, John Hopkins Alzheimer's Disease Research Center (NIA grant no. AG05146); Juan C. Troncoso, MD, Dr Olga Pletnikova, University of California, Los Angeles (NIA grant no. P50 AG16570); Harry Vinters, MD, Justine Pomakian, The Kathleen Price Bryan Brain Bank, Duke University Medical Center (NIA grant no. AG05128 and NIA P30 #AG028377, NINDS grant no. NS39764, NIMH MH60451 also funded by Glaxo Smith Kline); Christine Hulette, MD, Director, John F. Ervin, Stanford University; Dikran Horoupian, MD, Ahmad Salehi, MD, PhD, New York Brain Bank, Taub Institute, Columbia University (NYBB); Jean Paul Vonsattel, MD, Katerina Mancevska, Massachusetts General Hospital; E. Tessa Hedley-Whyte, MD, Dr M.P. Frosch, Karlotta Fitch (NIH grant P50 AG005134), University of Michigan (NIH grant P50-AG08671); Dr Roger Albin, Lisa Bain, Eszter Gombosi, University of Kentucky (NIH no. AG05144): William Markesbery, MD, Sonya Anderson, Mayo Clinic, Jacksonville; Dennis W. Dickson, MD, Natalie Thomas, University Southern California; Caroll A. Miller, MD, Jenny Tang, M.S., Dimitri Diaz, Washington University, St Louis Alzheimer's Disease Research Center (NIH no. P50AG05681): Dan McKeel, MD, John C. Morris, MD, Eugene Johnson, Jr, PhD, Virginia Buckles, PhD, Deborah Carter, University of Washington, Seattle (NIH no. P50 AG05136); Thomas Montine, MD, PhD, Aimee Schantz, MD, Ann C. McKee, Carol Kubilus Sun Health Research Institute, Arizona (NIA no. P30 AG19610); Joseph Rogers, PhD, Thomas G. Beach, MD, PhD, Lucia I. Sue Emory University; Bruce H. Wainer, MD, PhD, Marla Gearing, PhD, University of Texas, Southwestern Medical School; Charles L. White, III, MD, Roger Rosenberg, Marilyn Howell, Joan Reisch, University of California, Davis;

William Ellis, MD, Mary Ann Jarvis, Rush University Medical Center, Rush Alzheimer's Disease Center (NIH no. P30AG10161); David A. Bennett, MD, Julie A. Schneider, MD, MS, Karen Skish, MS, PA (ASCP)MT, Wayne T Longman, University of Miami/NPF Brain Endowment Bank; Deborah C. Mash, MD, Margaret J. Basile, Mitsuko Tanaka Oregon Health & Science University; Randy Wotljer, PhD, additional tissues include samples from the following sites; Newcastle Brain Tissue Resource (funding via the UK Medical Research Council, local NHS trusts and Newcastle University); C.M. Morris, PhD, Ian G McKeith (Tissue for this study was provided by the Newcastle Brain Tissue Resource which is funded in part by a grant from the UK Medical Research Council (G0400074), by the Newcastle NIHR Biomedical Research Centre in Ageing and Age Related Diseases to the Newcastle upon Tyne Hospitals NHS Foundation Trust and by a grant from the Alzheimer's Society and Alzheimer's Research Trust as part of the Brains for Dementia Research Project); the MRC London Brain Bank for Neurodegenerative Diseases; Simon Lovestone, MD PhD, Safa Al-Sarraj, MD, Claire Troakes, South West Dementia Brain Bank (funding via numerous sources including the Higher Education Funding Council for England (HEFCE), Alzheimer's Research Trust (ART), BRACE as well as North Bristol NHS Trust Research and Innovation Department and DeNDRoN); Seth Love, MD, Patrick Kehoe, PhD, Laura Palmer, The Netherlands Brain Bank (funding via numerous sources including Stichting MS Research, Brain Net Europe, Hersenstichting Nederland Breinbrekend Werk, International Parkinson Fonds, Internationale Stiching Alzheimer Onderzoek); Inge Huitinga, MD, Marleen Rademaker, Michiel Kooreman, Institut de Neuropatologia, Servei Anatomia Patologica, Universitat de Barcelona; Isidre Ferrer Abizanda, MD, PhD, Susana Casas Boluda.

## REFERENCES


1. Evans, D.A., Funkenstein, H.H., Albert, M.S., Scherr, P.A., Cook, N.R., Chown, M.J., Hebert, L.E., Hennekens, C.H. and Taylor, J.O. (1989) Prevalence of Alzheimer's disease in a community population of older persons. Higher than previously reported. *JAMA*, **262**, 2551–2556.
2. Meyer, J.M. and Breitner, J.C. (1998) Multiple threshold model for the onset of Alzheimer's disease in the NAS-NRC twin panel. *Am. J. Med. Genet.*, **81**, 92–97.
3. Williamson, J., Goldman, J. and Marder, K.S. (2009) Genetic aspects of Alzheimer disease. *Neurologist*, **15**, 80–86.
4. Corder, E.H., Saunders, A.M., Strittmatter, W.J., Schmechel, D.E., Gaskell, P.C., Small, G.W., Roses, A.D., Haines, J.L. and Pericak-Vance, M.A. (1993) Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science*, **261**, 921–923.
5. Strittmatter, W.J., Saunders, A.M., Schmechel, D., Pericak-Vance, M., Enghild, J., Salvesen, G.S. and Roses, A.D. (1993) Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proc. Natl Acad. Sci. USA*, **90**, 1977–1981.
6. Ashford, J.W. and Mortimer, J.A. (2002) Non-familial Alzheimer's disease is mainly due to genetic factors. *J. Alzheimers. Dis.*, **4**, 169–177.
7. Slooter, A.J., Cruts, M., Kalmijn, S., Hofman, A., Breteler, M.M., Van Broeckhoven, C. and van Duijn, C.M. (1998) Risk estimates of dementia by apolipoprotein E genotypes from a population-based incidence study: the Rotterdam Study. *Arch. Neurol.*, **55**, 964–968.
8. Daw, E.W., Payami, H., Nemens, E.J., Nochlin, D., Bird, T.D., Schellenberg, G.D. and Wijsman, E.M. (2000) The number of trait loci in late-onset Alzheimer disease. *Am. J. Hum. Genet.*, **66**, 196–204.
9. Evans, D.A., Beckett, L.A., Field, T.S., Feng, L., Albert, M.S., Bennett, D.A., Tycko, B. and Mayeux, R. (1997) Apolipoprotein E epsilon4 and incidence of Alzheimer disease in a community population of older persons. *JAMA*, **277**, 822–824.
10. Guo, Z., Fratiglioni, L., Viitanen, M., Lannfelt, L., Basun, H., Fastbom, J. and Winblad, B. (2001) Apolipoprotein E genotypes and the incidence of Alzheimer's disease among persons aged 75 years and older: variation by use of antihypertensive medication? *Am. J. Epidemiol.*, **153**, 225–231.
11. Harold, D., Abraham, R., Hollingworth, P., Sims, R., Gerrish, A., Hamshere, M.L., Pahwa, J.S., Moskvina, V., Dowzell, K., Williams, A. *et al.* (2009) Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat. Genet.*, **41**, 1088–1093.
12. Lambert, J.C., Heath, S., Even, G., Campion, D., Sleegers, K., Hiltunen, M., Combarros, O., Zelenika, D., Bullido, M.J., Tavernier, B. *et al.* (2009) Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. *Nat. Genet.*, **41**, 1094–1099.
13. Bertram, L., McQueen, M.B., Mullin, K., Blacker, D. and Tanzi, R.E. (2007) Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nat. Genet.*, **39**, 17–23.
14. Coon, K.D., Myers, A.J., Craig, D.W., Webster, J.A., Pearson, J.V., Lince, D.H., Zismann, V.L., Beach, T.G., Leung, D., Bryden, L. *et al.* (2007) A high-density whole-genome association study reveals that APOE is the major susceptibility gene for sporadic late-onset Alzheimer's disease. *J. Clin. Psychiatry*, **68**, 613–618.
15. Reiman, E.M., Webster, J.A., Myers, A.J., Hardy, J., Dunckley, T., Zismann, V.L., Joshipura, K.D., Pearson, J.V., Hu-Lince, D., Huentelman, M.J. *et al.* (2007) GAB2 alleles modify Alzheimer's risk in APOE epsilon4 carriers. *Neuron*, **54**, 713–720.
16. Webster, J.R.E., Zismann, V.L., Joshipura, J.D., Pearson, J.V., Hu-Lince, D., Huentelman, M.J., Craig, D.W., Coon, K.D., Beach, T., Roher, K.C. *et al.* (2010) Whole genome association analysis shows that ACE is a risk factor for Alzheimer's disease and fails to replicate most candidates from Meta-analysis. *Int. J. Mol. Epidemiol. Genet.*, **1**, 19–30.
17. Bennett, D.A., De Jager, P.L., Leurgans, S.E. and Schneider, J.A. (2009) Neuropathologic intermediate phenotypes enhance association to Alzheimer susceptibility alleles. *Neurology*, **72**, 1495–1503.
18. Guerreiro, R.J., Beck, J., Gibbs, J.R., Santana, I., Rossor, M.N., Schott, J.M., Nalls, M.A., Ribeiro, H., Santiago, B., Fox, N.C. *et al.* (2010) Genetic variability in CLU and its association with Alzheimer's disease. *PLoS ONE*, **5**, e9510.
19. Wray, S. and Noble, W. (2009) Linking amyloid and tau pathology in Alzheimer's disease: the role of membrane cholesterol in Abeta-mediated tau toxicity. *J. Neurosci.*, **29**, 9665–9667.
20. Michikawa, M. (2004) Neurodegenerative disorders and cholesterol. *Curr. Alzheimer. Res.*, **1**, 271–275.
21. Sink, K.M., Leng, X., Williamson, J., Kritchevsky, S.B., Yaffe, K., Kuller, L., Yasar, S., Atkinson, H., Robbins, M., Psaty, B. *et al.* (2009) Angiotensin-converting enzyme inhibitors and cognitive decline in older adults with hypertension: results from the Cardiovascular Health Study. *Arch. Intern. Med.*, **169**, 1195–1202.
22. Oveisgharan, S. and Hachinski, V. (2010) Hypertension, executive dysfunction and progression to dementia: the canadian study of health and aging. *Arch. Neurol.*, **67**, 187–192.
23. Cherbuin, N., Reglade-Meslin, C., Kumar, R., Jacomb, P., Easteal, S., Christensen, H., Sachdev, P. and Anstey, K.J. (2009) Risk factors of transition from normal cognition to mild cognitive disorder: the PATH through Life Study. *Dement. Geriatr. Cogn. Disord.*, **28**, 47–55.
24. Mi, W., Pawlik, M., Sastre, M., Jung, S.S., Radvinsky, D.S., Klein, A.M., Sommer, J., Schmidt, S.D., Nixon, R.A., Mathews, P.M. *et al.* (2007) Cystatin C inhibits amyloid-beta deposition in Alzheimer's disease mouse models. *Nat. Genet.*, **39**, 1440–1442.
25. Webster, J.A., Gibbs, J.R., Clarke, J., Ray, M., Zhang, W., Holmans, P., Rohrer, K., Zhao, A., Marlowe, L., Kaleem, M. *et al.* (2009) Genetic control of human brain transcript expression in Alzheimer disease. *Am. J. Hum. Genet.*, **84**, 445–458.
26. Myers, A.J., Gibbs, J.R., Webster, J.A., Rohrer, K., Zhao, A., Marlowe, L., Kaleem, M., Leung, D., Bryden, L., Nath, P. *et al.* (2007) A survey of genetic human cortical gene expression. *Nat. Genet.*, **39**, 1494–1499.
27. Braak, H. and Braak, E. (1995) Staging of Alzheimer's disease-related neurofibrillary changes. *Neurobiol. Aging*, **16**, 271–278 (discussion 278–284).

28. Crook, R., Hardy, J. and Duff, K. (1994) Single-day apolipoprotein E genotyping. *J. Neurosci. Methods*, **53**, 125–127.
29. Hawkins, J.R., Khripin, Y., Valdes, A.M. and Weaver, T.A. (2002) Miniaturized sealed-tube allele-specific PCR. *Hum. Mutat.*, **19**, 543–553.
30. Korn, J.M., Kuruvilla, F.G., McCarroll, S.A., Wysoker, A., Nemesh, J., Cawley, S., Hubbell, E., Veitch, J., Collins, P.J., Darvishi, K. *et al.* (2008) Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nat. Genet.*, **40**, 1253–1260.
31. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J. *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, **81**, 559–575.
32. Alexander, D., Novembre, J. and Lange, K. (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.*, **19**, 1655–1664.
33. Frazer, K.A., Ballinger, D.G., Cox, D.R., Hinds, D.A., Stuve, L.L., Gibbs, R.A., Belmont, J.W., Boudreau, A., Hardenbol, P., Leal, S.M. *et al.* (2007) A second generation human haplotype map of over 3.1 million SNPs. *Nature*, **449**, 851–861.
34. Li, Y., Willer, C., Sanna, S. and Abecasis, G. (2009) Genotype imputation. *Annu. Rev. Genomics Hum. Genet.*, **10**, 387–406.