

Analysis of Copy Number Variation in Alzheimer's Disease in a Cohort of Clinically Characterized and Neuropathologically Verified Individuals

Shanker Swaminathan^{1,2}, Matthew J. Huentelman^{3,4}, Jason J. Corneveaux^{3,4}, Amanda J. Myers^{5,6}, Kelley M. Faber², Tatiana Foroud^{1,2,7}, Richard Mayeux⁸, Li Shen^{1,7}, Sungeun Kim^{1,7}, Mari Turk^{3,4}, John Hardy⁹, Eric M. Reiman^{3,4,10}, Andrew J. Saykin^{1,2,7*}, for the Alzheimer's Disease Neuroimaging Initiative (ADNI) and the NIA-LOAD/NCRAD Family Study Group

1 Center for Neuroimaging, Department of Radiology and Imaging Sciences, Indiana University School of Medicine, Indianapolis, Indiana, United States of America, **2** Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, Indiana, United States of America, **3** Neurogenomics Division, The Translational Genomics Research Institute (TGen), Phoenix, Arizona, United States of America, **4** The Arizona Alzheimer's Consortium, Phoenix, Arizona, United States of America, **5** Departments of Psychiatry and Behavioral Sciences, and Human Genetics and Genomics, University of Miami, Miller School of Medicine, Miami, Florida, United States of America, **6** Johnnie B. Byrd Sr. Alzheimer's Center and Research Institute, Tampa, Florida, United States of America, **7** Center for Computational Biology and Bioinformatics, Indiana University School of Medicine, Indianapolis, Indiana, United States of America, **8** The Gertrude H. Sergievsky Center, The Taub Institute for Research on Alzheimer's Disease and the Aging Brain, and the Department of Neurology, Columbia University College of Physicians and Surgeons, New York, New York, United States of America, **9** Department of Molecular Neuroscience and Reta Lila Research Laboratories, University College London Institute of Neurology, London, United Kingdom, **10** Banner Alzheimer's Institute, Phoenix, Arizona, United States of America

Abstract

Copy number variations (CNVs) are genomic regions that have added (duplications) or deleted (deletions) genetic material. They may overlap genes affecting their function and have been shown to be associated with disease. We previously investigated the role of CNVs in late-onset Alzheimer's disease (AD) and mild cognitive impairment using Alzheimer's Disease Neuroimaging Initiative (ADNI) and National Institute of Aging-Late Onset AD/National Cell Repository for AD (NIA-LOAD/NCRAD) Family Study participants, and identified a number of genes overlapped by CNV calls. To confirm the findings and identify other potential candidate regions, we analyzed array data from a unique cohort of 1617 Caucasian participants (1022 AD cases and 595 controls) who were clinically characterized and whose diagnosis was neuropathologically verified. All DNA samples were extracted from brain tissue. CNV calls were generated and subjected to quality control (QC). 728 cases and 438 controls who passed all QC measures were included in case/control association analyses including candidate gene and genome-wide approaches. Rates of deletions and duplications did not significantly differ between cases and controls. Case-control association identified a number of previously reported regions (*CHRFAM7A*, *RELN* and *DOPEY2*) as well as a new gene (*HLA-DRA*). Meta-analysis of *CHRFAM7A* indicated a significant association of the gene with AD and/or MCI risk ($P=0.006$, odds ratio = 3.986 (95% confidence interval 1.490–10.667)). A novel *APP* gene duplication was observed in one case sample. Further investigation of the identified genes in independent and larger samples is warranted.

Citation: Swaminathan S, Huentelman MJ, Corneveaux JJ, Myers AJ, Faber KM, et al. (2012) Analysis of Copy Number Variation in Alzheimer's Disease in a Cohort of Clinically Characterized and Neuropathologically Verified Individuals. PLoS ONE 7(12): e50640. doi:10.1371/journal.pone.0050640

Editor: Tricia A. Thornton-Wells, Vanderbilt University, United States of America

Received: July 25, 2012; **Accepted:** October 23, 2012; **Published:** December 5, 2012

Copyright: © 2012 Swaminathan et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: Genotyping of the TGen cohort was supported by Kronos Science; the National Institute of Neurological Disorders and Stroke (NINDS) (R01NS059873); the National Institute on Aging (NIA) (R01AG034504, R01AG031581, P30AG19610, Z01AG000950-06, P30AG10161, R01AG15819); the Banner Alzheimer's Foundation; the Johnnie B. Byrd Sr. Alzheimer's Disease Institute; the Medical Research Council; the Intramural Research Program of the National Institutes of Health (NIH); and the State of Arizona. Many data and biomaterials for the TGen cohort were collected from several NIA and National Alzheimer's Coordinating Center (NACC) (U01AG016976) funded sites. These include: John Hopkins Alzheimer's Disease Research Center (NIA AG05146); University of California, Los Angeles (NIA P50AG16570); The Kathleen Price Bryan Brain Bank, Duke University Medical Center (NIA AG05128, NINDS NS39764, National Institute of Mental Health MH60451 also funded by Glaxo Smith Kline); Massachusetts Alzheimer's Disease Research Center (P50AG005134); University of Michigan (NIH P50-AG08671); University of Kentucky (NIH AG05144); Washington University, St Louis Alzheimer's Disease Research Center (NIH P50AG05681); University of Washington, Seattle (NIH P50AG05136); Boston University Alzheimer's Disease Research Center (NIH P30-AG13846); Sun Health Research Institute Brain Donation Program of Sun City, Arizona (NIA P30AG19610); Arizona Alzheimer's Disease Core Center, Arizona Department of Health Services, contract 211002, Arizona Alzheimer's Research Center; Arizona Biomedical Research Commission, contracts 4001, 0011, 05-901 and 1001 to the Arizona Parkinson's Disease Consortium; Michael J. Fox Foundation for Parkinson's Research; Rush University Medical Center, Rush Alzheimer's Disease Center (NIH AG10161). Additional tissues include samples from the following sites: Newcastle Brain Tissue Resource (funding via the Medical Research Council (MRC), local National Health Service (NHS) trusts and Newcastle University); MRC London Brain Bank for Neurodegenerative Diseases (funding via the Medical Research Council); 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 the National Institute for Health Research (NIHR) Dementias and Neurodegenerative Diseases Research Network (DeNDRoN); The Netherlands Brain Bank (funding via numerous sources including Stichting MS Research, Brain Net Europe, Hersenstichting Nederland Breinbrekend Werk, Internationale Parkinson Fonds, Internationale Stichting Alzheimer Onderzoek). Data collection and sharing for the Alzheimer's Disease Neuroimaging Initiative (ADNI) project was funded by the ADNI (NIH U01AG024904; RC2AG036535). ADNI is funded by the NIA, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Abbott; Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Amorfix Life Sciences Ltd.; AstraZeneca; Bayer HealthCare; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals Inc.; Eli Lilly and Company; F. Hoffman-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics, N.V.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc; Meso Scale Diagnostics, LLC.; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Society at the University of California, San Diego. ADNI data are disseminated by the Laboratory of Neuro Imaging at the University of California, Los Angeles. This research was also supported by NIH grants P30AG010129, K01AG030514, the Dana Foundation, U01AG032984 Alzheimer's Disease Genetics Consortium grant, NIA R01AG19771, P30AG010133, the Indiana Economic Development Corporation (IEDC #87884), and the Foundation for the NIH for data analysis. Funding support for the National Institute of Aging-Late Onset AD/National Cell Repository for AD (NIA-LOAD/NCRAD) Family Study was provided through NIA grants U24AG026395 (NIA-LOAD Family Study), U24AG021886 (National Cell Repository for Alzheimer's Disease) and other NIA grants. Genotyping services were provided by the Center for Inherited Disease Research funded through a federal contract (HHSN268200782096C) from the NIH to the John Hopkins University. Samples from the NCRAD, which receives government support under a cooperative agreement (U24AG021886) awarded by the NIA were used in the ADNI and NIA-LOAD/NCRAD Family studies. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: Genotyping of the TGen cohort was supported by Kronos Science. This study was partly funded by Glaxo Smith Kline. Data collection and sharing for the Alzheimer's Disease Neuroimaging Initiative (ADNI) project was funded by the ADNI (NIH U01AG024904; RC2AG036535). ADNI is funded by the NIA, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Abbott; Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Amorfix Life Sciences Ltd.; AstraZeneca; Bayer HealthCare; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals Inc.; Eli Lilly and Company; F. Hoffman-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics, N.V.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc; Meso Scale Diagnostics, LLC.; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. There are no patents, products in development or marketed products to declare. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials, as detailed online in the guide for authors.

* E-mail: asaykin@iupui.edu

Introduction

Alzheimer's disease (AD) is the most common form of dementia characterized by loss of memory and other cognitive abilities, severe enough to disrupt daily life activities. An estimated 5.4 million Americans have AD, the sixth leading cause of death across all ages in the United States [1]. No treatments at present can slow or halt its progression. Amnesic mild cognitive impairment (MCI) is a clinical condition in which a person has memory problems not normal for his/her age, but not severe enough to interfere significantly with daily functioning. Approximately 14–18% of individuals aged 70 years and older have MCI, and every year 10–15% of these individuals will likely progress to dementia, particularly AD [2].

Genetic factors play a key role in AD development accounting for approximately 58–79% of the phenotypic variation [3]. Mutations in *APP*, *PSEN1* and *PSEN2* primarily cause early-onset AD (age at onset <60 or 65 years) [4]. The leading genetic risk factor for the more common late-onset AD (age at onset >60 or 65 years) is the *APOE* ϵ 4 allele. Large case-control genome-wide association studies (GWASs) have identified and replicated other AD risk loci including: *CLU*, *CRI1*, *PICALM*, *BLN1*, *EXOC3L2*, *MTHFD1L*, *MS4A4A/MS4A6E*, *CD2AP*, *CD33*, *ABCA7* and *CUGBP2* [5–12]. However it is estimated that the *APOE* ϵ 4 allele accounts for approximately 20% and the non-*APOE* loci cumulatively account for as much as 35% of the AD risk [10,13]. A recent study observed only a

marginal joint effect of known loci on memory independent from *APOE* [14]. The combined loci provided minimal improvement of prediction of AD beyond age, sex and *APOE*. Thus the loci do not explain all the genetic variation associated with AD, and other forms of genetic variation such as copy number variations (CNVs) may play a role.

CNVs are deoxyribonucleic acid (DNA) regions (one kilobase (kb) to several megabases (Mb) in size) that have differences in copy number. These can result in the addition (copy number gains or duplications) or loss (copy number losses or deletions) of genetic material. CNVs often encompass a single gene or multiple genes and may affect their function [15]. The role of CNVs in late-onset AD has been investigated in prior studies [16–19]. Previously, we analyzed the role of CNVs in AD and MCI using data from participants in the Alzheimer's Disease Neuroimaging Initiative (ADNI) study [20] and the National Institute of Aging-Late Onset AD/National Cell Repository for AD (NIA-LOAD/NCRAD) Family Study [21]. For both studies, DNA extracted either from peripheral blood or brain tissue were used. Case/control association analyses including candidate gene and genome-wide approaches were performed to determine genes overlapped by CNVs only in cases (AD and/or MCI) but not controls. A number of genes were identified in the two studies including *ATXN1*, *CHRFAM7A*, *CSMD1*, *DOPEY2*, *ERBB4*, *GSTT1*, *HLA-DPB1*, *HNRNPCL1*, *IMMP2L*, *NRXN1*, *RELN* and *SLC35F2*.

The aim of the present report is to analyze the role of CNVs in AD using data from a unique cohort of clinically characterized and neuropathologically defined cases (AD) and controls (TGen cohort) [22]. All DNA samples were extracted from brain tissue. Case/control association analyses similar to the two previous studies were performed to determine the CNV burden in cases relative to controls and genes overlapped by CNVs detected in cases but not controls. Here we report analyses identifying a number of previously reported as well as new CNV regions.

Materials and Methods

Samples

The TGen cohort included samples extracted from brain tissue of 1617 Caucasian individuals (1022 AD cases and 595 controls). Recruitment information for the participants has been previously described [22]. Briefly, the United States cohort was obtained from 21 National Institute on Aging-supported Alzheimer's Disease Center brain banks and from the Miami Brain Bank [23,24]. Cohorts from other brain banks in the United States, United Kingdom, and the Netherlands were obtained similar to the original United States cohort. Genome-wide genotyping for all samples was performed using the Affymetrix Genome-Wide Human SNP 6.0 Array (Santa Clara, California, United States of America) as previously described [22]. *APOE* genotyping was done using Crook *et al.*'s method [25] or using a fluorescence-based allele-specific polymerase chain reaction (PCR), also called PCR Amplification of Specific Alleles, on array tape [26] by PreventionGenetics (Marshfield, Wisconsin, United States of America).

The ADNI data used in the preparation of the present report were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (<http://adni.loni.ucla.edu/>). ADNI's primary goal is to test whether imaging markers, genetic markers, other biological markers, and clinical and neuropsychological assessments can be combined to measure progression of MCI and early AD. More information on ADNI can be found on <http://www.adni-info.org/>. The Illumina Human610-Quad BeadChip (San Diego, California, United States of America) was used to perform genome-wide genotyping of the ADNI sample as previously described [20,27]. The *APOE* polymorphisms (rs429358 and rs7412) were genotyped separately.

The NIA-LOAD/NCRAD Family Study data used in the present report were obtained from the "NIA-Late Onset Alzheimer's Disease and National Cell Repository for Alzheimer's Disease Family Study: Genome-Wide Association Study for Susceptibility Loci" dataset (dbGaP Study Accession: phs000168.v1.p1, Project #2026) on the database of Genotypes and Phenotypes (http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000168.v1.p1) website. Recruitment information for NIA-LOAD Family Study and NCRAD participants has been previously described [12]. Genome-wide genotyping for all samples was performed using the Illumina Human610-Quad BeadChip at the Center for Inherited Disease Research (Baltimore, Maryland, United States of America). The *APOE* polymorphisms (rs429358 and rs7412) were genotyped at PreventionGenetics.

Ethics Statement

De-identification of samples in the TGen cohort was done before receipt, and the study met human studies institutional review board and the Health Insurance Portability and Accountability Act of 1996 regulations. The present work is declared not human-subjects research and is institutional review board exempt under regulation 45 CFR 46. The ADNI study was approved by

institutional review boards of all participating institutions and written informed consent was obtained from all participants or authorized representatives. All individuals in the NIA-LOAD/NCRAD Family Study were recruited after providing informed consent and with approval by the relevant institutional review boards. The study was conducted according to the principles in the Declaration of Helsinki.

Generation of CNV Calls and Quality Control

CNV calls were generated for the 1617 TGen samples using PennCNV (2011Jun16 version; <http://www.openbioinformatics.org/penncnv/>), a Hidden Markov model based program [28]. The PennCNV-Affy protocol (http://www.openbioinformatics.org/penncnv/penncnv_tutorial_affy_gw6.html) for the Affymetrix Genome-Wide Human SNP 6.0 Array was first performed to transform raw CEL files into a signal intensity file containing the Log R Ratio (LRR) and B Allele Frequency (BAF) values used by PennCNV to generate CNV calls. The Hidden Markov model "affygw6.hmm", population frequency of B allele "affygw6.hg18.pfb" and gcmodel "affygw6.hg18.gcmodel" files were used. Extensive quality control (QC) was performed on all samples. A genomic wave adjustment procedure using PennCNV's gcmodel file was applied as samples that have below optimal genomic wave QC values can be considered unreliable [29]. Frequency distribution plots of the number of CNV calls, LRR standard deviation (SD), BAF Drift and Waviness Factor (WF) were made. A sample was excluded if at least one of the above measures for the sample was greater than 90th percentile of the frequency distribution, i.e. the sample had >56 CNV calls, LRR SD>0.38, BAF Drift>0.01 or WF>0.02.

Due to complications of hemizygosity in males and X-chromosome inactivation in females, analyses were restricted to autosomes. To ensure we were including only high-confidence CNVs in the analysis, CNVs for which the difference of the log likelihood of the most likely copy number state and less likely copy number state was <10, CNVs called based on data <10 SNPs, and CNVs that had >50% overlap with centromeric, telomeric, and immunoglobulin regions as defined in Need *et al.* [30] were excluded. CNV calls were not filtered for size because both large and small variants could be of potential significance. A case sample observed to have a very large (~8.4 Mb) deletion on chromosome 19, and a control sample observed to have a very large (~22.4 Mb) duplication on chromosome 1, were excluded from the analyses as they may be possible outliers. The ~8.4 Mb deletion on chromosome 19 encompassed both sides of the centromere, but did not overlap any RefSeq or UCSC Genes according to the UCSC Genome Browser [31] (<http://genome.ucsc.edu/>). 1166 samples (728 cases, 438 controls) with 31045 CNV calls remained after all QC measures and were entered into case/control association analyses.

Case/control Association Analyses

Case/control analyses using permutation-based tests of association in the TGen study were performed similar to the ADNI [20] and NIA-LOAD/NCRAD Family [21] studies. PLINK v1.07 [32] (<http://pngu.mgh.harvard.edu/~purcell/plink/>) was used to investigate CNV call differences between cases (AD) and controls. Two approaches were used: a candidate gene approach including 317 AD genes identified from the AlzGene database (Updated 5 January 2011) (<http://www.alzgene.org/>) as having a positive association with AD in at least one study, and a genome-wide approach using 17938 genes from PLINK's gene list (hg18 coordinates). The AlzGene database is a publicly available online resource that provides a comprehensive and regularly updated

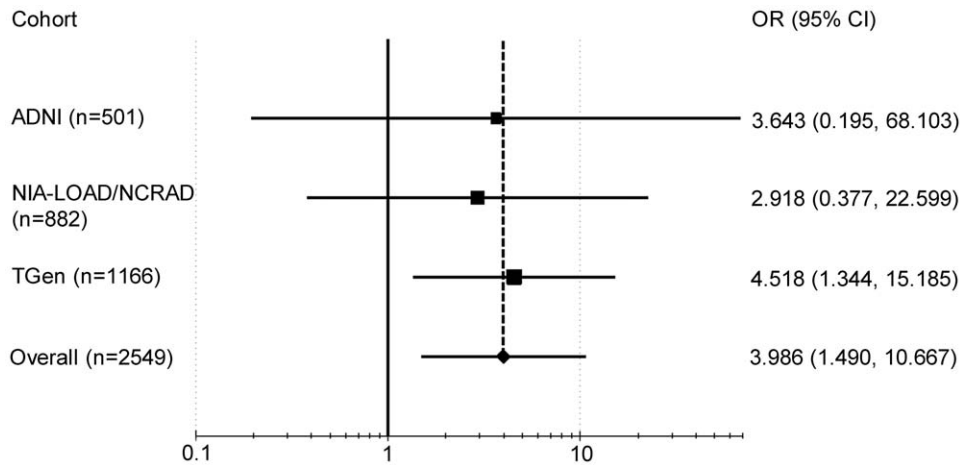


Figure 1. Forest plot of the *CHRFAM7A* gene. The plot represents the meta-analysis of the *CHRFAM7A* gene using results from the Alzheimer's Disease Neuroimaging Initiative (ADNI) study, the National Institute of Aging-Late Onset AD/National Cell Repository for AD (NIA-LOAD/NCRAD) Family Study and the TGen study. The odds ratio (OR) and 95% confidence interval (CI) for the odds ratio for each study are represented by black squares and horizontal lines. The summary odds ratio is depicted as a black diamond. doi:10.1371/journal.pone.0050640.g001

catalog of genetic case/control and family association studies in AD [33]. In both approaches, CNV segments either partially or completely overlapping genes were analyzed. The analyses included both deletions and duplications. 50000 null permutations were performed to generate one-sided empirical P values testing genes overlapped by CNV calls in more cases than controls. The analyses focused on genes overlapped by CNV calls in cases, but not in controls, to identify genes that may play a role in AD susceptibility. 317 genes were considered in the candidate gene approach and 17938 genes were considered in the genome-wide approach. Genes that achieved $P < 0.05$ (one-sided) were considered significant.

Meta-analysis

We performed a meta-analysis for the *CHRFAM7A* gene using results from the ADNI, NIA-LOAD/NCRAD Family and TGen studies to determine differences in frequency of CNV calls overlapping the gene between cases (AD and/or MCI) and controls. A fixed-effects model was run and a summary odds ratio (OR) was calculated using the Mantel Haenszel method. MetaAnalyst Beta 3.13 [34] (http://tuftscaes.org/meta_analyst/index.html) and Comprehensive Meta-Analysis Version 2 [35] were used for the meta-analysis and generation of the forest plot (Figure 1). The UCSC Genome Browser [31] (<http://genome.ucsc.edu/>) (March 2006 (NCBI36/hg18) assembly) was used to create representative plots of the CNV calls (Figures 2 and 3). The Genome Browser track for the Affymetrix Genomewide 6.0 array was obtained from the PennCNV website (http://www.openbioinformatics.org/penncnv/penncnv_download.html).

Results

Sample Demographics and CNV Call Characteristics

The sample demographics and CNV call characteristics of the 728 cases and 438 controls who passed all QC measures are shown in Tables 1 and 2. Significant ($P < 0.05$; two-sided) differences in gender, absence or presence of the *APOE* $\epsilon 4$ allele, age at death, Braak stage and the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) score were observed between cases and controls. 31045 CNV calls (24188 deletions and 6857

duplications) were observed with an average of 45 SNPs per CNV call and an average CNV call length of 64.76 kb. A higher CNV call rate and a lower average CNV call size were observed in deletions compared to duplications. Rates of deletions and duplications did not significantly differ between cases and controls. There were no significant differences in the rates of deletions and rates of duplications when males and females were analyzed separately (data not shown). A large proportion of deletions and duplications were found in the 0.1–0.5 Mb size range (Table 3).

Case/control Association Analyses

The candidate gene approach identified 32 of the 317 genes tested (10.09%) and the genome-wide approach identified 939 of the 17938 genes tested (5.23%) to be overlapped by CNV calls only in cases (AD) but not controls in the TGen study. A significant ($P = 0.0003$; Fisher's exact test; two-sided) enrichment of the candidate genes relative to the genome was observed.

Candidate gene approach. We identified 32 candidate genes in the TGen study overlapped by CNV calls from at least one case (AD) but no controls (Table 4). Representative plots of two genes (*APP* and *DOPEY2*) are shown in Figure 2. The *HLA-DRA* gene was overlapped by deletions in nine cases (uncorrected $P = 0.0140$; one-sided). This gene was also found to be overlapped by deletions in two controls in the ADNI study. Two genes (*RELN* overlapped by deletions in two cases and *DOPEY2* overlapped by duplications in four cases) identified in this study were also reported from only cases (AD and/or MCI) in the ADNI and NIA-LOAD/NCRAD Family studies. One AD sample (*APOE* $\epsilon 2/\epsilon 3$ genotype, age at death = 67) had a novel *APP* gene duplication supported by 443 sequential SNP and CNV probes. The *CHRFAM7A* gene reported in the ADNI and NIA-LOAD/NCRAD Family studies was overlapped by deletions in 10 cases and two controls, and duplications in 12 cases and one control (corrected $P = 0.0198$; one-sided) in this study (Figure 3).

Genome-wide approach. We also identified 939 genes across the genome overlapped by CNV calls only in cases (AD) but not controls in the TGen study. Genes overlapped by CNV calls from at least four cases but not controls in the TGen study are shown in Table 5. The *HLA-DRA* gene overlapped by deletions in nine cases in the TGen study and identified in the candidate gene

A



B

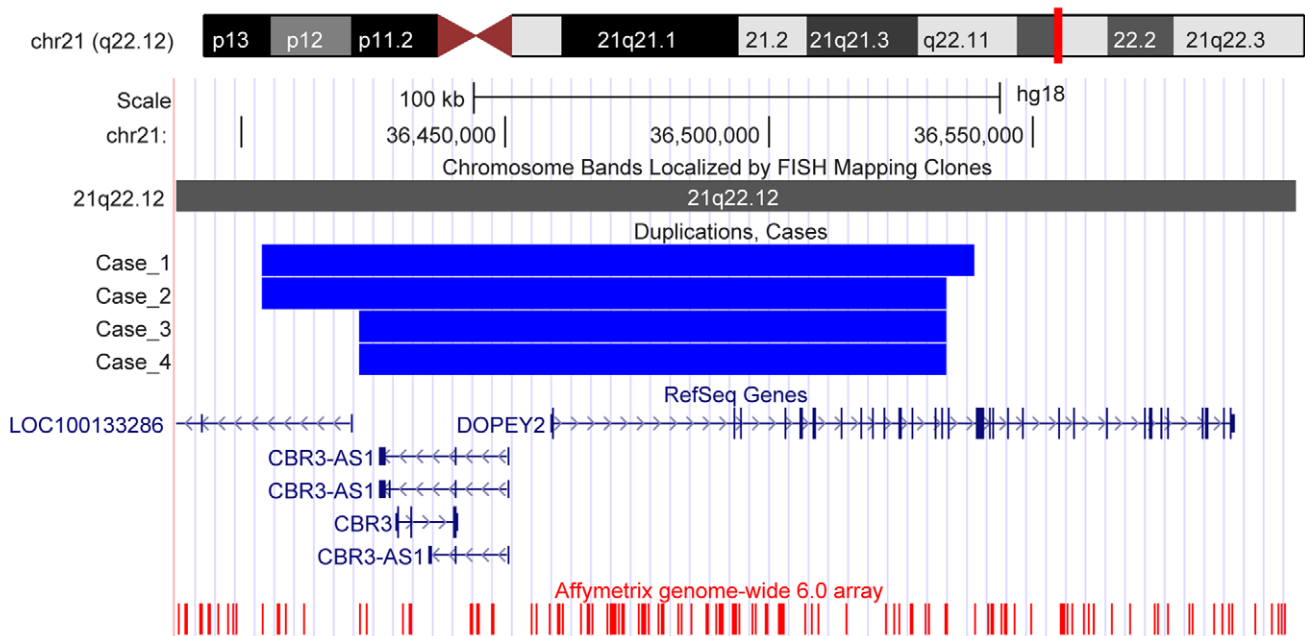


Figure 2. Duplications overlapping the candidate genes *APP* and *DOPEY2* in samples of the TGen cohort. Representative UCSC Genome Browser (March 2006 (NCBI36/hg18) assembly) plots of duplications overlapping the candidate genes: (A) *APP* and (B) *DOPEY2*, in samples of the TGen cohort. The chromosomal location of the gene and probes on the Affymetrix Genomewide 6.0 array are shown. The region with the duplication for each sample relative to the gene is represented by a blue rectangle.
doi:10.1371/journal.pone.0050640.g002

approach was also found from this approach (uncorrected $P=0.0144$; one-sided). The *CHRFAM7A* gene reported in the ADNI and NIA-LOAD/NCRAD Family Studies and mentioned in the candidate gene approach had an uncorrected $P=0.0046$;

one-sided. Deletions in 10 cases and two controls, and duplications in 12 cases and one control, overlapped the *CHRFAM7A* gene in the TGen study (Figure 3). Other genes reported in the ADNI and NIA-LOAD/NCRAD Family studies were overlapped by CNV

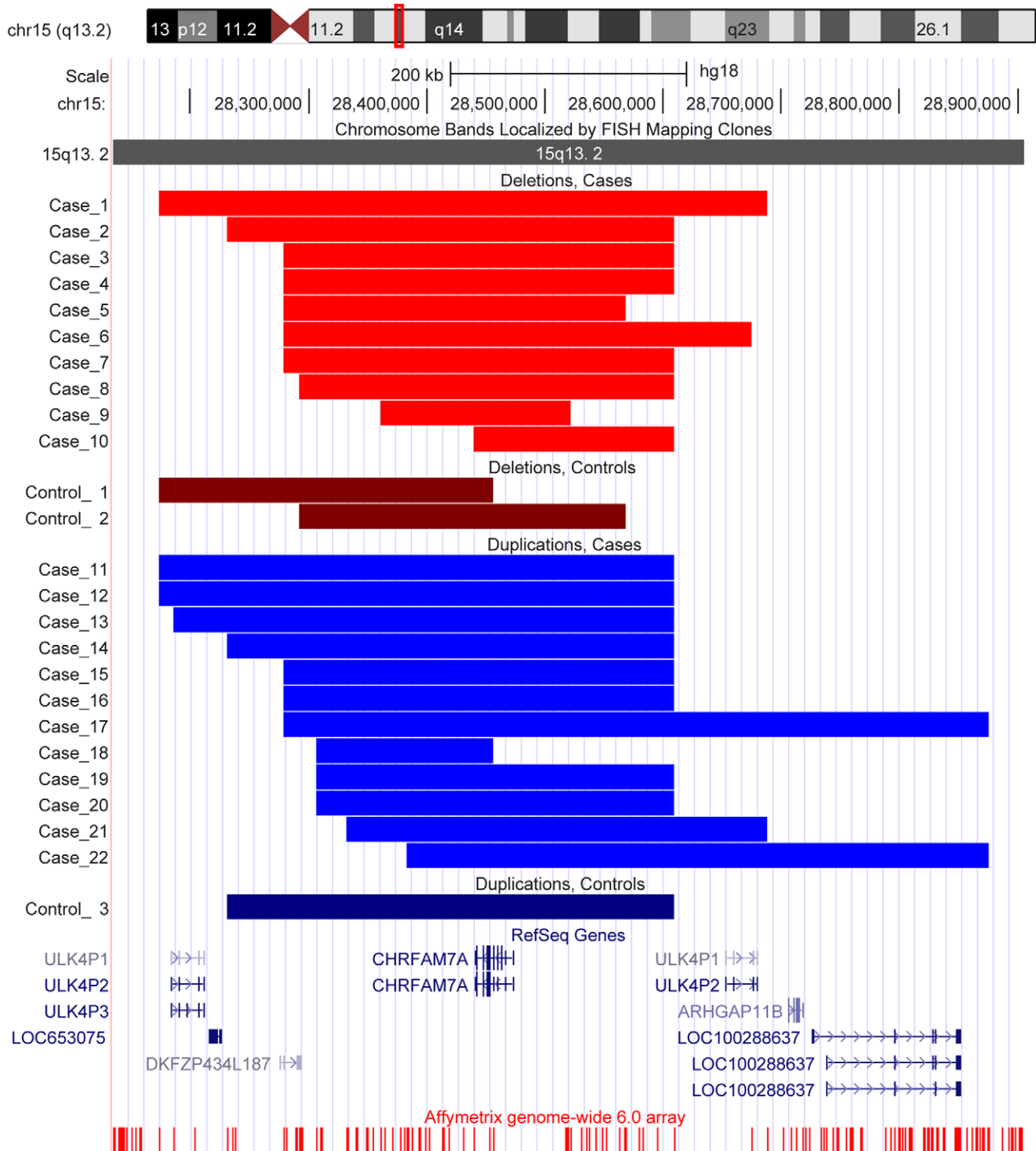


Figure 3. Deletions and duplications overlapping the *CHRFAM7A* gene in samples of the TGen cohort. Representative UCSC Genome Browser (March 2006 (NCBI36/hg18) assembly) plots of deletions and duplications overlapping the *CHRFAM7A* gene in samples of the TGen cohort. The chromosomal location of the gene and probes on the Affymetrix Genomewide 6.0 array are shown. The region with the deletion for each sample relative to the gene is highlighted by a red rectangle and the region with the duplication for each sample relative to the gene is highlighted by a blue rectangle.

doi:10.1371/journal.pone.0050640.g003

calls in the TGen study although they did not reach significance at uncorrected $P < 0.05$. These include *CSMD1* (deletions in 65 cases and 32 controls, duplications in one case and one control), *ERBB4*

(deletions in 71 cases and 35 controls, duplication in one control), *HNRNPCL1* (deletions in 19 cases and eight controls, duplications in one case and two controls), *IMMP2L* (deletions in six cases and

Table 1. Sample demographics in the TGen cohort.

| | Cases | Controls | P (two-sided) |
|--|--------------------------|--------------------------|---------------|
| Number of samples | 728 | 438 | – |
| Gender (Males/Females) | 180/548 | 274/164 | <0.001 |
| APOE group ($\epsilon 4$ negative/ $\epsilon 4$ positive) | 93/173 | 138/45 | <0.001 |
| Age at onset | 72.84 \pm 6.14 (n=60) | – | – |
| Age at death | 82.46 \pm 7.58 (n=266) | 80.80 \pm 9.17 (n=183) | 0.037 |
| Braak stage | 5.09 \pm 0.83 (n=168) | 1.62 \pm 0.76 (n=96) | <0.001 |
| CERAD ^a score | 2.31 \pm 0.84 (n=26) | 0.83 \pm 0.64 (n=47) | <0.001 |

^aCERAD = The Consortium to Establish a Registry for Alzheimer's Disease.
doi:10.1371/journal.pone.0050640.t001

five controls, duplication in one control), *NRXN1* (deletions in two cases and three controls), and *SLC35F2* (duplications in two cases and one control). Deletions in three cases and three controls, and duplications in two cases and three controls in the TGen study, overlapped the *CHRNA7* gene identified in a previous study [16]. No CNV calls were found in the TGen study overlapping other previously reported regions including *CRI* [17], 14q11.2 [18] and 15q11.2 [19].

Meta-analysis

A meta-analysis for the *CHRFAM7A* gene was performed using findings from the ADNI, NIA-LOAD/NCRAD Family and TGen studies (Table 6 and Figure 1). Overall, the gene was overlapped by CNV calls in 38 of 1797 (2.115%) cases (AD and/or MCI) and four of 752 (0.532%) controls. A significant association was observed for the gene with AD and/or MCI risk ($P=0.006$, summary OR = 3.986 (95% confidence interval 1.490–10.667)).

Discussion

The present report represents an initial CNV analysis in the TGen cohort, a unique cohort of clinically characterized and neuropathologically verified individuals. After extensive QC, case (AD)/control association analyses using candidate gene and genome-wide approaches were performed. Genes enriched in cases relative to controls were determined, suggesting possible involvement of these genes in AD susceptibility.

Rates of deletions and duplications did not significantly differ between cases and controls. This is different from findings in the ADNI and NIA-LOAD/NCRAD Family studies and could be due to different participant selection criteria, random sampling variation, different QC criteria, and that the NIA-LOAD/NCRAD Family Study and TGen study analyses included only AD and control samples, whereas the ADNI study analyses included MCI samples in addition to AD and control samples.

The candidate gene approach revealed a number of interesting genes (Table 4 and Figure 2). The *HLA-DRA* (major histocompatibility complex, class II, DR alpha) gene on chromosome 6 is a human leukocyte antigen (HLA) class II alpha chain paralogue which plays an important role in the immune system by presenting peptides derived from extracellular proteins. Variants in *HLA-DRA* have been associated with Parkinson disease [36,37] and multiple sclerosis [38,39], but not with AD. Other *HLA* alleles however have been investigated for a possible role in AD [40–43].

The *CHRFAM7A* (*CHRNA7* (cholinergic receptor, nicotinic, alpha 7, exons 5–10) and *FAM7A* (family with sequence similarity 7A, exons A–E) fusion) gene (Figure 3) is located on chromosome 15. It is formed as a hybrid of a partially duplicated *CHRNA7* gene and the *FAM7A* gene [44,45]. It is highly polymorphic and individuals with and without this gene have been identified. A 2-bp deletion polymorphism at position 497–498 in exon 6 of this gene has been observed to be significantly over-represented in participants with AD, dementia with Lewy bodies and Pick's disease compared to controls [46].

Although *CHRFAM7A* is transcribed, its translation and possible function of the resulting protein is uncertain. The gene is expressed in the hippocampus, a brain region known to be first affected in AD. Recently, it has been suggested to possibly modulate $\alpha 7$ subunit receptor-mediated synaptic transmission and cholinergic anti-inflammatory response [47]. It may also be a dominant negative modulator of *CHRNA7* function and important for receptor regulation in humans [48]. Duplications overlapping the *CHRNA7* gene and an approximately 300 kb region upstream of the gene were identified in a previous study investigating the role of CNVs in AD [16]. A meta-analysis using the *CHRFAM7A* gene results from the ADNI, NIA-LOAD/NCRAD Family and TGen studies indicated a significant association of the gene with AD and/or MCI risk at an OR of 3.986. This association is striking relative to known AD markers and is similar to the OR for one copy of the *APOE* $\epsilon 4$ allele. However, unlike the *APOE* $\epsilon 4$ allele which is a change in a single base within the gene, CNV calls overlapping the *CHRFAM7A* gene involves more than one base and often a large region with less than perfect definition of boundaries given available methodology. In addition, in a CNV analysis, a gene may be overlapped by deletions or duplications

Table 2. Characteristics of copy number variation calls from samples in the TGen cohort.

| | Cases (n = 728) | Controls (n = 438) |
|---------------------------------------|-----------------|--------------------|
| Deletions | | |
| Number of copy number variation calls | 15177 | 9011 |
| Rate per sample | 20.85 | 20.57 |
| Average size (kilobase) | 45.34 | 46.75 |
| Duplications | | |
| Number of copy number variation calls | 4334 | 2523 |
| Rate per sample | 5.95 | 5.76 |
| Average size (kilobase) | 140.8 | 131.1 |

doi:10.1371/journal.pone.0050640.t002

Table 3. TGen samples grouped by copy number variation call size.

| Call size (Megabase) | Cases (n = 728) | | Controls (n = 438) | |
|----------------------|-----------------|--------------------|--------------------|--------------------|
| | Deletions n (%) | Duplications n (%) | Deletions n (%) | Duplications n (%) |
| 0.1–0.5 | 647 (88.87) | 625 (85.85) | 387 (88.36) | 368 (84.02) |
| 0.5–1.0 | 95 (13.05) | 164 (22.53) | 61 (13.93) | 101 (23.06) |
| 1.0–1.5 | 10 (1.37) | 25 (3.43) | 7 (1.60) | 12 (2.74) |
| 1.5–2.0 | 4 (0.55) | 7 (0.96) | 0 (0.00) | 2 (0.46) |
| >2.0 | 2 (0.27) | 11 (1.51) | 5 (1.14) | 1 (0.23) |

doi:10.1371/journal.pone.0050640.t003

Table 4. Genes overlapped by copy number variation calls from at least one case and no control samples in the TGen cohort using the candidate gene approach.

| Chromosome | Region | Start (base pair) | End (base pair) | Number of cases |
|------------|-----------------|-------------------|-----------------|--------------------|
| 1 | <i>FAM63A</i> | 149234172 | 149245957 | 1 |
| 1 | <i>SOAT1</i> | 177529639 | 177591076 | 1 |
| 1 | <i>AGT</i> | 228904891 | 228916959 | 1 ^a |
| 2 | <i>LHCGR</i> | 48767416 | 48836384 | 4 ^a |
| 6 | <i>HLA-G</i> | 29902734 | 29906878 | 1 |
| 6 | <i>HLA-A</i> | 30018309 | 30021633 | 2 |
| 6 | <i>HLA-E</i> | 30565249 | 30569072 | 3 ^b |
| 6 | <i>BAT1</i> | 31605974 | 31618204 | 1 ^a |
| 6 | <i>HLA-DRA</i> | 32515624 | 32520802 | g ^{b,c,d} |
| 6 | <i>HLA-DQB1</i> | 32735634 | 32742444 | 1 ^c |
| 7 | <i>MAGI2</i> | 77484309 | 78920826 | 1 |
| 7 | <i>CD36</i> | 80069439 | 80146529 | 1 |
| 7 | <i>RELN</i> | 102899472 | 103417198 | 2 |
| 8 | <i>NAT2</i> | 18293034 | 18303003 | 1 |
| 10 | <i>ALDH18A1</i> | 97355675 | 97406557 | 1 |
| 10 | <i>EBF3</i> | 131523536 | 131652081 | 1 |
| 11 | <i>PICALM</i> | 85346132 | 85457756 | 1 |
| 12 | <i>C12orf41</i> | 47333261 | 47362302 | 3 |
| 12 | <i>ALDH2</i> | 110688728 | 110732167 | 2 |
| 14 | <i>PSEN1</i> | 72672931 | 72756862 | 1 |
| 14 | <i>NGB</i> | 76801586 | 76807408 | 1 ^d |
| 17 | <i>SERPINF2</i> | 1592879 | 1605309 | 1 |
| 17 | <i>MYH13</i> | 10144907 | 10217047 | 2 ^e |
| 17 | <i>MYH8</i> | 10234366 | 10265992 | 1 ^e |
| 17 | <i>MAPT</i> | 41327543 | 41461546 | 1 |
| 19 | <i>GALP</i> | 61379200 | 61388956 | 1 |
| 21 | <i>APP</i> | 26174731 | 26465003 | 1 |
| 21 | <i>DOPEY2</i> | 36458708 | 36588442 | 4 |
| 21 | <i>CBS</i> | 43346369 | 43369493 | 1 |
| 21 | <i>S100B</i> | 46842958 | 46849463 | 1 |
| 22 | <i>COMT</i> | 18309308 | 18336530 | 1 |
| 22 | <i>BCR</i> | 21852551 | 21990224 | 1 |

^aA case sample had copy number variation calls overlapping the *AGT*, *LHCGR* and *BAT1* genes.^{b,c,d}Three different case samples had copy number variation calls overlapping the *HLA-DRA* and *HLA-E*, *HLA-DRA* and *HLA-DQB1*, and *HLA-DRA* and *NGB* genes respectively.^eA different case sample had copy number variation calls overlapping the *MYH13* and *MYH8* genes.

doi:10.1371/journal.pone.0050640.t004

Table 5. Genes overlapped by copy number variation calls in at least four case samples and no control samples in the TGen cohort using the genome-wide approach.

| Chromosome | Region | Start (base pair) | End (base pair) | Number of cases |
|------------|-------------------|-------------------|-----------------|-----------------|
| 1 | <i>LCE1D</i> | 151035850 | 151037281 | 6 |
| 2 | <i>GTF2A1L</i> | 48698451 | 48813790 | 4 ^a |
| 2 | <i>LHCGR</i> | 48767416 | 48836384 | 4 ^a |
| 4 | <i>TMPRSS11E2</i> | 69465107 | 69515259 | 6 |
| 6 | <i>HLA-DRA</i> | 32515624 | 32520802 | 9 ^b |
| 7 | <i>RP9</i> | 33100934 | 33115527 | 4 |
| 11 | <i>OR52B4</i> | 4345156 | 4346101 | 4 ^c |
| 11 | <i>GRM5</i> | 87880625 | 88420838 | 4 |
| 11 | <i>ST3GAL4</i> | 125731305 | 125789743 | 4 ^c |
| 21 | <i>CBR3</i> | 36429132 | 36440730 | 4 ^d |
| 21 | <i>DOPEY2</i> | 36458708 | 36588442 | 4 ^d |
| 22 | <i>DGCR2</i> | 17403794 | 17489967 | 5 ^b |

^aFour case samples had copy number variation calls overlapping the *GTF2A1L* and *LHCGR* genes.

^bA different case sample had copy number variation calls overlapping the *HLA-DRA* and *DGCR2* genes.

^cA different case sample had copy number variation calls overlapping the *OR52B4* and *ST3GAL4* genes.

^dFour different case samples had copy number variation calls overlapping the *CBR3* and *DOPEY2* genes.

doi:10.1371/journal.pone.0050640.t005

which may have different effects on the protein. Nonetheless, the strength of CNV signals overlapping the *CHRFAM7A* gene in the present three samples warrants further investigation and replication in additional independent samples.

Two genes (*RELN* and *DOPEY2*) in the three studies were found to be overlapped by CNV calls from cases (AD and/or MCI) but not controls. The *RELN* (reelin) gene on chromosome 7 encodes the glycoprotein reelin, which activates a signaling pathway required for proper positioning of neurons within laminated nervous system parenchyma. Gene variants have been associated with AD [49] and the protein has been observed to have increased expression in pyramidal neurons of the hippocampus in AD individuals and in cognitively intact controls with AD-associated pathology [50]. The *DOPEY2* (dopey family member 2, also known as *C21orf5*) gene located on chromosome 21 in the Down syndrome critical region is a potential Down syndrome candidate

gene [51,52]. Overexpression of the gene may be associated with the neurological phenotypes and mental retardation observed in Down syndrome patients. Although the *DOPEY2* gene was identified in the candidate gene approach in the TGen study, CNV calls overlapping this gene also completely overlapped the *CBR3* (carbonyl reductase 3) and *CBR3-AS1* (*CBR3* antisense RNA 1). The *CBR3* protein, which catalyzes reduction of carbonyl compounds to their corresponding alcohols, has been found to be differentially expressed in the cerebellum of patients with atypical cases of frontotemporal lobar degeneration with fused in sarcoma-positive inclusions compared to controls [53]. Although *DOPEY2* is a known candidate, it is possible that *CBR3* and *CBR3-AS1* could be pathologic given their involvement in other neurodegenerative diseases. One AD sample (*APOE* $\epsilon 2/\epsilon 3$ genotype, age at death = 67) was identified with an *APP* duplication. The patient started suffering from forgetfulness at approximately 59 years and

Table 6. Meta-analysis of the *CHRFAM7A* gene.

| Cohort | ADNI (n = 501) | NIA-LOAD/NCRAD (n = 882) | TGen (n = 1166) | Overall (n = 2549) |
|--|----------------|--------------------------|-----------------|--------------------|
| Cases | | | | |
| Number of samples | 358 | 711 | 728 | 1797 |
| CNV ^a overlapping gene (n) | 4 | 12 | 22 | 38 |
| No CNV ^a overlapping gene (n) | 354 | 699 | 706 | 1759 |
| Controls | | | | |
| Number of samples | 143 | 171 | 438 | 752 |
| CNV ^a overlapping gene (n) | 0 | 1 | 3 | 4 |
| No CNV ^a overlapping gene (n) | 143 | 170 | 435 | 748 |
| Odds ratio | 3.643 | 2.918 | 4.518 | 3.986 |
| 95% confidence interval | 0.195–68.103 | 0.377–22.599 | 1.344–15.185 | 1.490–10.667 |
| P | 0.387 | 0.305 | 0.015 | 0.006 |

^aCNV = copy number variation.

doi:10.1371/journal.pone.0050640.t006

was admitted to a nursing home at approximately age 63. The patient had no reported family history of dementia. *APP* duplications have been associated with early-onset [54–57] but not late-onset AD, and it is possible that this patient may have had an early-onset sporadic form of AD (age at onset <60 or 65 years). Replication in independent samples and laboratory validation can help confirm the role of these genes in AD susceptibility.

The genome-wide approach revealed the candidate genes: *HLA-DRA* and *CHRFAM7A* as well as identified genes reported in the two previous studies. The *CSMD1*, *HNRNPCL1*, *IMMP2L* and *SLC35F2* genes have not been previously associated with AD. The *NRXN1* gene has been associated with autism [58], schizophrenia [59], and has been shown to have reduced expression with increasing AD severity [60]. The *ERBB4* gene may play a possible role in the progression of AD pathology [61–63].

It is important to note the limitations of the present report. Although the same software (PennCNV) was used in the three studies, different QC criteria were used for the selection of samples. The ADNI and NIA-LOAD/NCRAD Family study samples were genotyped on the Illumina Human610-Quad BeadChip and used similar QC criteria. The TGen study samples were genotyped on the Affymetrix Genome-Wide Human SNP 6.0 Array and thus we used a slightly different QC criteria. To our knowledge, there does not appear to be consensus on a well defined set of QC criteria for inclusion of the most appropriate samples in CNV analyses. The QC criterion applied in the present study may have been too stringent, leading to samples possibly having informative CNV data being excluded. A direct comparison of CNV calls from the three studies would be difficult as two different genotyping platforms were used. Probes from the two platforms may not correspond with each other with respect to their location. Replication in additional independent data sets and future molecular studies will help confirm the findings.

In sum, we have conducted an initial CNV analysis in samples from a cohort of clinically characterized and neuropathologically verified individuals. Rates of deletions and duplications did not significantly differ between cases and controls. Gene-based association analysis identified a number of genes including those reported in the ADNI and NIA-LOAD/NCRAD Family studies (*CHRFAM7A*, *RELN* and *DOPEY2*) as well as a new gene (*HLA-DRA*). Meta-analysis from the three studies revealed a significant association for *CHRFAM7A* with AD and/or MCI risk. Replication in independent samples will be necessary to confirm these findings. Targeted analyses of the identified regions will help determine the biological role of these variants. Overall, there appears to be some consistency of CNVs across AD cohorts and this variation holds promise for revealing novel risk factors and disease mechanisms.

Acknowledgments

Many data and biomaterials for the TGen cohort were collected from several National Institute on Aging (NIA) and National Alzheimer's Coordinating Center (NACC) funded sites. We thank Amanda J. Myers (University of Miami, Department of Psychiatry) and John A. Hardy (Reta Lila Weston Institute, University College London) for collecting and preparing the series. Marcelle Morrison-Bogorad, Tony Phelps, and Walter Kukull 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: Juan C. Troncoso, Olga Pletnikova; University of California, Los Angeles: Harry Vinters, Justine Pomakian; The Kathleen Price Bryan

Brain Bank, Duke University Medical Center: Christine Hulette, John F. Ervin; Stanford University: Dikran Horoupian, Ahmad Salehi; New York Brain Bank, Taub Institute, Columbia University (NYBB): Jean Paul Vonsattel, Katerina Mancevska; Massachusetts Alzheimer's Disease Research Center: E. Tessa Hedley-Whyte, M.P. Frosch, Karlotta Fitch; University of Michigan: Roger Albin, Lisa Bain, Eszter Gombosi; University of Kentucky: William Markesbery, Sonya Anderson; Mayo Clinic, Jacksonville: Dennis W. Dickson, Natalie Thomas; University Southern California: Caroll A. Miller, Jenny Tang, Dimitri Diaz; Washington University, St Louis Alzheimer's Disease Research Center: Dan McKeel, John C. Morris, Eugene Johnson, Jr., Virginia Buckles, Deborah Carter; University of Washington, Seattle: Thomas Montine, Aimee Schantz; University of Pennsylvania School of Medicine, Alzheimer's Disease Research Center: John Q. Trojanowski, Virginia M Lee, Viviana Van Deerlin, Terry Schuck; Boston University Alzheimer's Disease Research Center: Ann C. McKee, Carol Kubilus; Sun Health Research Institute Brain Donation Program of Sun City, Arizona: Joseph Rogers, Thomas G. Beach, Lucia I. Sue; Emory University: Bruce H. Wainer, Marla Gearing; University of Texas, Southwestern Medical School: Charles L. White, III, Roger Rosenberg, Marilyn Howell, Joan Reisch; University of California, Davis: William Ellis, Mary Ann Jarvis; Rush University Medical Center, Rush Alzheimer's Disease Center: David A. Bennett, Julie A. Schneider, Karen Skish, Wayne T. Longman; University of Miami/NPF Brain Endowment Bank: Deborah C. Mash, Margaret J. Basile, Mitsuko Tanaka; Oregon Health & Science University: Randy Wotjler. Additional tissues include samples from the following sites: Newcastle Brain Tissue Resource: C.M. Morris, Ian G. McKeith, Robert H. Perry; MRC London Brain Bank for Neurodegenerative Diseases: Simon Lovestone, Safa Al-Sarraj, Claire Troakes; South West Dementia Brain Bank: Seth Love, Patrick Kehoe, Laura Palmer; The Netherlands Brain Bank: Inge Huitinga, Marleen Rademaker, Michiel Kooreman; Institut de Neuropatologia, Servei Anatomia Patologica, Universitat de Barcelona: Isidre Ferrer Abizanda, Susana Casas Boluda.

Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.ucla.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.ucla.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf. We also thank the following people: genotyping at The Translational Genomics Research Institute (TGen): Matthew Huentelman and David Craig, and sample processing, storage and distribution at the National Cell Repository for Alzheimer's Disease (NCRAD): Kelley Faber and Colleen Mitchell.

For the National Institute of Aging-Late Onset AD/National Cell Repository for AD (NIA-LOAD/NCRAD) Family Study, we thank the NIH Genome-Wide Association Studies Data Repository for providing the "Genetic Consortium for Late Onset Alzheimer's Disease" dataset (dbGaP Study Accession: phs000168.v1.p1, Project #2026). Investigators who contributed to the NIA-LOAD/NCRAD Family Study can be found on the dbGaP website (http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000168.v1.p1). We would like to thank Susan LaRusse Eckert, Stephan Dean and Jennifer Williamson-Catania at Columbia University, Michele Goodman at Indiana University, Creighton H. Phelps, Marcelle Morrison-Bogorod, Marilyn Miller, and Stephen Snyder, at the NIA, for their guidance and help in coordinating the project.

The authors would like to thank contributors who collected samples as well as patients and their families, whose help and participation made this work possible. We would also like to thank John West for his help.

Author Contributions

Conceived and designed the experiments: SS MJH LS AS. Performed the experiments: SS MJH JJC AJM KMF TF RM LS MT JH EMR AS. Analyzed the data: SS MJH LS SK AS. Contributed reagents/materials/analysis tools: MJH JJC AJM KMF AS. Wrote the paper: SS MJH JJC AJM KMF TF RM LS SK MT JH EMR AS.

References

- Alzheimer's Association, Thies W, Bleiler L (2011) 2011 Alzheimer's disease facts and figures. *Alzheimers Dement* 7: 208–244.
- Petersen RC, Roberts RO, Knopman DS, Boeve BF, Geda YE, et al. (2009) Mild cognitive impairment: ten years later. *Arch Neurol* 66: 1447–1455.
- Gatz M, Reynolds CA, Fratiglioni L, Johansson B, Mortimer JA, et al. (2006) Role of genes and environments for explaining Alzheimer disease. *Arch Gen Psychiatry* 63: 168–174.
- Bekris LM, Yu CE, Bird TD, Tsuang DW (2010) Genetics of Alzheimer disease. *J Geriatr Psychiatry Neurol* 23: 213–227.
- Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, et al. (2009) Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat Genet* 41: 1088–1093.
- Lambert JC, Heath S, Even G, Campion D, Sleegers K, et al. (2009) Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. *Nat Genet* 41: 1094–1099.
- Jun G, Naj AC, Beecham GW, Wang LS, Buros J, et al. (2010) Meta-analysis confirms CR1, CLU, and PICALM as Alzheimer disease risk loci and reveals interactions with APOE genotypes. *Arch Neurol* 67: 1473–1484.
- Seshadri S, Fitzpatrick AL, Ikram MA, DeStefano AL, Gudnason V, et al. (2010) Genome-wide analysis of genetic loci associated with Alzheimer disease. *JAMA* 303: 1832–1840.
- Naj AC, Beecham GW, Martin ER, Gallins PJ, Powell EH, et al. (2010) Dementia revealed: novel chromosome 6 locus for late-onset Alzheimer disease provides genetic evidence for folate-pathway abnormalities. *PLoS Genet* 6: e1001130.
- Naj AC, Jun G, Beecham GW, Wang LS, Vardarajan BN, et al. (2011) Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. *Nat Genet* 43: 436–441.
- Hollingworth P, Harold D, Sims R, Gerrish A, Lambert JC, et al. (2011) Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. *Nat Genet* 43: 429–435.
- Wijsman EM, Pankratz ND, Choi Y, Rothstein JH, Faber KM, et al. (2011) Genome-wide association of familial late-onset Alzheimer's disease replicates BIN1 and CLU and nominates CUGBP2 in interaction with APOE. *PLoS Genet* 7: e1001308.
- Slooter AJ, Cruts M, Kalmijn S, Hofman A, Breteler MM, et al. (1998) Risk estimates of dementia by apolipoprotein E genotypes from a population-based incidence study: the Rotterdam Study. *Arch Neurol* 55: 964–968.
- Verhaaren BF, Vernooij MW, Koudstaal PJ, Uitterlinden AG, Duijn CM, et al. (2012) Alzheimer's Disease Genes and Cognition in the Nondemented General Population. *Biol Psychiatry*.
- Cook EH, Jr., Scherer SW (2008) Copy-number variations associated with neuropsychiatric conditions. *Nature* 455: 919–923.
- Heinzen EL, Need AC, Hayden KM, Chiba-Falek O, Roses AD, et al. (2010) Genome-wide scan of copy number variation in late-onset Alzheimer's disease. *J Alzheimers Dis* 19: 69–77.
- Brouwers N, Cauwenbergh CV, Engelborghs S, Lambert JC, Bettens K, et al. (2012) Alzheimer risk associated with a copy number variation in the complement receptor 1 increasing C3b/C4b binding sites. *Mol Psychiatry* 17: 223–233.
- Shaw CA, Li Y, Wiszniewska J, Chasse S, Zaidi SN, et al. (2011) Olfactory copy number association with age at onset of Alzheimer disease. *Neurology* 76: 1302–1309.
- Ghani M, Pinto D, Lee JH, Grinberg Y, Sato C, et al. (2012) Genome-wide survey of large rare copy number variants in Alzheimer's disease among Caribbean hispanics. *G3 (Bethesda)* 2: 71–78.
- Swaminathan S, Kim S, Shen L, Risacher SL, Foroud T, et al. (2011) Genomic Copy Number Analysis in Alzheimer's Disease and Mild Cognitive Impairment: An ADNI Study. *Int J Alzheimers Dis* 2011: 729478.
- Swaminathan S, Shen L, Kim S, Inlow M, West JD, et al. (2012) Analysis of Copy Number Variation in Alzheimer's Disease: the NIA-LOAD/NCRAD Family Study. *Curr Alzheimer Res* 9: 801–814.
- Corneveaux JJ, Myers AJ, Allen AN, Pruzin JJ, Ramirez M, et al. (2010) Association of CR1, CLU and PICALM with Alzheimer's disease in a cohort of clinically characterized and neuropathologically verified individuals. *Hum Mol Genet* 19: 3295–3301.
- Webster JA, Gibbs JR, Clarke J, Ray M, Zhang W, et al. (2009) Genetic control of human brain transcript expression in Alzheimer disease. *Am J Hum Genet* 84: 445–458.
- Myers AJ, Gibbs JR, Webster JA, Rohrer K, Zhao A, et al. (2007) A survey of genetic human cortical gene expression. *Nat Genet* 39: 1494–1499.
- Crook R, Hardy J, Duff K (1994) Single-day apolipoprotein E genotyping. *J Neurosci Methods* 53: 125–127.
- Hawkins JR, Khripin Y, Valdes AM, Weaver TA (2002) Miniaturized sealed-tube allele-specific PCR. *Hum Mutat* 19: 543–553.
- Saykin AJ, Shen L, Foroud TM, Potkin SG, Swaminathan S, et al. (2010) Alzheimer's Disease Neuroimaging Initiative biomarkers as quantitative phenotypes: Genetics core aims, progress, and plans. *Alzheimers Dement* 6: 265–273.
- Wang K, Li M, Hadley D, Liu R, Glessner J, et al. (2007) PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res* 17: 1665–1674.
- Diskin SJ, Li M, Hou C, Yang S, Glessner J, et al. (2008) Adjustment of genomic waves in signal intensities from whole-genome SNP genotyping platforms. *Nucleic Acids Res* 36: e126.
- Need AC, Ge D, Weale ME, Maia J, Feng S, et al. (2009) A genome-wide investigation of SNPs and CNVs in schizophrenia. *PLoS Genet* 5: e1000373.
- Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, et al. (2002) The human genome browser at UCSC. *Genome Res* 12: 996–1006.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81: 559–575.
- Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE (2007) Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nat Genet* 39: 17–23.
- Wallace BC, Schmid CH, Lau J, Trikalinos TA (2009) Meta-Analyst: software for meta-analysis of binary, continuous and diagnostic data. *BMC Med Res Methodol* 9: 80.
- Borenstein M, Hedges L, Higgins J, Rothstein H (2005) *Comprehensive Meta-analysis Version 2*, Biostat, Englewood NJ.
- Hamza TH, Zabetian CP, Tenesa A, Laederach A, Montimurro J, et al. (2010) Common genetic variation in the HLA region is associated with late-onset sporadic Parkinson's disease. *Nat Genet* 42: 781–785.
- Guo Y, Deng X, Zheng W, Xu H, Song Z, et al. (2011) HLA rs3129882 variant in Chinese Han patients with late-onset sporadic Parkinson disease. *Neurosci Lett* 501: 185–187.
- International Multiple Sclerosis Genetics Consortium, Hafler DA, Compston A, Sawcer S, Lander ES, et al. (2007) Risk alleles for multiple sclerosis identified by a genomewide study. *N Engl J Med* 357: 851–862.
- Hoppensbrouwers IA, Aulchenko YS, Janssens AC, Ramagopalan SV, Broer L, et al. (2009) Replication of CD58 and CLEC16A as genome-wide significant risk genes for multiple sclerosis. *J Hum Genet* 54: 676–680.
- Ma SL, Tang NL, Tam CW, Lui VW, Suen EW, et al. (2008) Association between HLA-A alleles and Alzheimer's disease in a southern Chinese community. *Dement Geriatr Cogn Disord* 26: 391–397.
- Lehmann DJ, Barnardo MC, Fuggle S, Quiroga I, Sutherland A, et al. (2006) Replication of the association of HLA-B7 with Alzheimer's disease: a role for homozygosity? *J Neuroinflammation* 3: 33.
- Guerini FR, Tinelli C, Calabrese E, Agliardi C, Zanzottera M, et al. (2009) HLA-A*01 is associated with late onset of Alzheimer's disease in Italian patients. *Int J Immunopathol Pharmacol* 22: 991–999.
- Listi F, Candore G, Balistreri CR, Grimaldi MP, Orlando V, et al. (2006) Association between the HLA-A2 allele and Alzheimer disease. *Rejuvenation Res* 9: 99–101.
- Riley B, Williamson M, Collier D, Wilkie H, Makoff A (2002) A 3-Mb map of a large Segmental duplication overlapping the alpha7-nicotinic acetylcholine receptor gene (CHRNA7) at human 15q13-q14. *Genomics* 79: 197–209.
- Gault J, Robinson M, Berger R, Drebing C, Logel J, et al. (1998) Genomic organization and partial duplication of the human alpha7 neuronal nicotinic acetylcholine receptor gene (CHRNA7). *Genomics* 52: 173–183.
- Feher A, Juhasz A, Rimanoczy A, Csibri E, Kalman J, et al. (2009) Association between a genetic variant of the alpha-7 nicotinic acetylcholine receptor subunit and four types of dementia. *Dement Geriatr Cogn Disord* 28: 56–62.
- de Lucas-Cerrillo AM, Maldifassi MC, Arnalich F, Renart J, Atienza G, et al. (2011) Function of partially duplicated human alpha77 nicotinic receptor subunit CHRFAM7A gene: potential implications for the cholinergic anti-inflammatory response. *J Biol Chem* 286: 594–606.
- Araud T, Graw S, Berger R, Lee M, Neveu E, et al. (2011) The chimeric gene CHRFAM7A, a partial duplication of the CHRNA7 gene, is a dominant negative regulator of alpha7*nAChR function. *Biochem Pharmacol* 82: 904–914.
- Seripa D, Matera MG, Franceschi M, Daniele A, Bizzarro A, et al. (2008) The RELN locus in Alzheimer's disease. *J Alzheimers Dis* 14: 335–344.
- Kramer PL, Xu H, Woltjer RL, Westaway SK, Clark D, et al. (2011) Alzheimer disease pathology in cognitively healthy elderly: a genome-wide study. *Neurobiol Aging* 32: 2113–2122.
- Rachidi M, Delezoide AL, Delabar JM, Lopes C (2009) A quantitative assessment of gene expression (QAGE) reveals differential overexpression of DOPEY2, a candidate gene for mental retardation, in Down syndrome brain regions. *Int J Dev Neurosci* 27: 393–398.
- Rachidi M, Lopes C, Delezoide AL, Delabar JM (2006) C21orf5, a human candidate gene for brain abnormalities and mental retardation in Down syndrome. *Cytogenet Genome Res* 112: 16–22.
- Martins-de-Souza D, Guest PC, Mann DM, Roebor S, Rahmoune H, et al. (2012) Proteomic analysis identifies dysfunction in cellular transport, energy, and protein metabolism in different brain regions of atypical frontotemporal lobar degeneration. *J Proteome Res* 11: 2533–2543.
- Rovelet-Lecrux A, Hannequin D, Raux G, Le Meur N, Laquerriere A, et al. (2006) APP locus duplication causes autosomal dominant early-onset Alzheimer disease with cerebral amyloid angiopathy. *Nat Genet* 38: 24–26.

55. Sleegers K, Brouwers N, Gijselinck I, Theuns J, Goossens D, et al. (2006) APP duplication is sufficient to cause early onset Alzheimer's dementia with cerebral amyloid angiopathy. *Brain* 129: 2977–2983.
56. McNaughton D, Knight W, Guerreiro R, Ryan N, Lowe J, et al. (2012) Duplication of amyloid precursor protein (APP), but not prion protein (PRNP) gene is a significant cause of early onset dementia in a large UK series. *Neurobiol Aging* 33: 426 e413–421.
57. Antonell A, Gelpi E, Sanchez-Valle R, Martinez R, Molinuevo JL, et al. (2012) Breakpoint sequence analysis of an AbetaPP locus duplication associated with autosomal dominant Alzheimer's disease and severe cerebral amyloid angiopathy. *J Alzheimers Dis* 28: 303–308.
58. Autism Genome Project Consortium, Szatmari P, Paterson AD, Zwaigenbaum L, Roberts W, et al. (2007) Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nat Genet* 39: 319–328.
59. Rujescu D, Ingason A, Cichon S, Pietilainen OP, Barnes MR, et al. (2009) Disruption of the neurexin 1 gene is associated with schizophrenia. *Hum Mol Genet* 18: 988–996.
60. Gomez Ravetti M, Rosso OA, Berretta R, Moscato P (2010) Uncovering molecular biomarkers that correlate cognitive decline with the changes of hippocampus' gene expression profiles in Alzheimer's disease. *PLoS One* 5: e10153.
61. Chaudhury AR, Gerecke KM, Wyss JM, Morgan DG, Gordon MN, et al. (2003) Neuregulin-1 and erbB4 immunoreactivity is associated with neuritic plaques in Alzheimer disease brain and in a transgenic model of Alzheimer disease. *J Neuropathol Exp Neurol* 62: 42–54.
62. Woo RS, Lee JH, Yu HN, Song DY, Baik TK (2010) Expression of ErbB4 in the apoptotic neurons of Alzheimer's disease brain. *Anat Cell Biol* 43: 332–339.
63. Woo RS, Lee JH, Yu HN, Song DY, Baik TK (2011) Expression of ErbB4 in the neurons of Alzheimer's disease brain and APP/PS1 mice, a model of Alzheimer's disease. *Anat Cell Biol* 44: 116–127.