Genome-Wide Survey of Large Rare Copy Number Variants in Alzheimer’s Disease Among Caribbean Hispanics

Mahdi Ghani,* 1 Dalila Pinto,† 1 Joseph H. Lee,‡ § ½, ¾, **†† Yakov Grinberg,* Christine Sato,* Danielle Moreno,* Stephen W. Scherer,† Richard Mayeux,* ½, ¾, **†† Peter St. George-Hyslop,* ½, ¾, **†† and Ekaterina Rogaeva*‡‡

*Tanz Centre for Research in Neurodegenerative Diseases, University of Toronto, Toronto, Ontario M5S 3H2, Canada, ††Department of Medicine, University of Toronto, Toronto, Ontario M5S 1A8, Canada, ¹Genetics and Genomic Biology, Hospital for Sick Children, Toronto, Ontario M5G 1L7, Canada, ‡Taub Institute for Research on Alzheimer’s Disease and the Aging Brain and ³Gertrude H. Sergievsky Center, Columbia University Medical Center, New York, New York 10032, ²Departments of Neurology, Psychiatry, and Medicine, College of Physicians and Surgeons, and ³Cambridge Institute for Medical Research and ⁴*Department of Clinical Neurosciences, University of Cambridge, Cambridge CB2 0XY, United Kingdom

ABSTRACT

Recently genome-wide association studies have identified significant association between Alzheimer’s disease (AD) and variations in CLU, PICALM, BIN1, CR1, MS4A4A/MS4A6E, CD2AP, CD33, EPHA1, and ABCA7. However, the pathogenic variants in these loci have not yet been found. We conducted a genome-wide scan for large copy number variation (CNV) in a dataset of Caribbean Hispanic origin (554 controls and 559 AD cases that were previously investigated in a SNP-based genome-wide association study using Illumina HumanHap 650Y platform). We ran four CNV calling algorithms to obtain high-confidence calls for large CNVs (>100 kb) that were detected by at least two algorithms. Global burden analyses did not reveal significant differences between cases and controls in CNV rate, distribution of deletions or duplications, total or average CNV size; or number of genes affected by CNVs. However, we observed a nominal association between AD and a ~470 kb duplication on chromosome 15q11.2 (P = 0.037). This duplication, encompassing up to five genes (TUBGCP5, CYFIP1, NIPA2, NIPA1, and WHAMML1) was present in 10 cases (2.6%) and 3 controls (0.8%). The dosage increase of CYFIP1 and NIPA1 genes was further confirmed by quantitative PCR. The current study did not detect CNVs that affect novel AD loci identified by recent genome-wide association studies. However, because the array technology used in our study has limitations in detecting small CNVs, future studies must carefully assess novel AD genes for the presence of disease-related CNVs.

KEYWORDS

gene deletion duplication Alzheimer’s Disease copy number variants

Alzheimer’s disease (AD) is the most common form of dementia, affecting ~30% of individuals over 80 years of age (Mayeux 2003). The hallmark of AD brain pathology is characterized by the accumulation of a neurotoxic proteolytic derivative of the amyloid precursor protein (APP; Aβ peptides) and the formation of intraneuronal tau-associated neurofibrillary tangles. The majority of AD cases are sporadic (~95%), with onset after 65 years of age. One half of the genetic variance is attributable to mutations in APP, PSEN1, PSEN2, and the APOE e4-allele, and they have been shown to cause the overproduction or reduced clearance of Aβ (Rogaeva et al. 2006). More recently, it was demonstrated that single nucleotide polymorphisms (SNP) in SORL1 are significantly associated with late-onset AD in several independent cohorts (Rogaeva et al. 2007; Reitz et al. 2011). In addition,
genome-wide association studies (GWAS) of large case-control datasets have identified significant association between late-onset AD and SNPs in CLU, PICALM, BIN1, CR1, MS4A4A/MS4A6E, CD2AP, CD33, EPHA1, and ABCA7 (Harold et al. 2009; Lambert et al. 2009; Carrasquillo et al. 2010; Hollingworth et al. 2011; Naj et al. 2011). Our recent GWAS of a Caribbean Hispanic cohort has supported the association with SNPs in the CLU, PICALM, and BIN1 genes (Lee JH et al. 2010). However, the pathogenic variants in these novel AD loci have not yet been found.

Copy number variants (CNV) have been associated with several neuropsychiatric disorders, such as autism, schizophrenia, and bipolar disorder (Cook and Scherer 2008; Lee and Scherer 2010). Furthermore, rare duplications of the APP locus are associated with dominant early-onset AD, which support the possibility of the existence of disease-related CNVs in other AD genes (McNaughton et al. 2010). In fact, recently Brouwers et al. proposed that the association between CR1 and AD might be explained by intragenic CNVs that translate into two major CR1 isoforms (Brouwers et al. 2011). To date, there are only two published genome-wide case-control studies that assess the contribution of CNVs to AD in North American populations (Heinzen et al. 2010; Swaminathan et al. 2011). However, in both studies, CNV calls were detected using a single method (PennCNV), and no overall case-control differences were observed in the CNV rate, size, presence of rare genic CNVs, or number of genes disrupted by CNVs. The only borderline association was reported by Heinzen et al. for a ~500 kb duplication at 15q13.3 affecting the CHRNA7 gene that encodes the neuronal nicotinic cholinergic receptor (P = 0.053, uncorrected for multiple testing) (Heinzen et al. 2010).

To evaluate the contribution of rare genomic variants to risk of late-onset AD, we analyzed a Caribbean Hispanic dataset that was previously assessed in a SNP-based GWAS (Lee JH et al. 2010). We focused our investigation on large rare CNVs that might contribute significantly to disease risk, as was previously demonstrated in other neuropsychiatric disorders (Kirov et al. 2009; Zhang et al. 2009; Glessner et al. 2010). To maximize CNV discovery, we used multiple CNV detection methods.

### METHODS

#### Sample collection and genotyping

The study was approved by the Institutional Review Boards of Columbia University and the University of Toronto. The Caribbean Hispanic case-control dataset, consisting of participants predominantly originating from the Dominican Republic and Puerto Rico, was described previously (Lee JH et al. 2010). Briefly, the dataset included 559 unrelated cases with late-onset AD and 554 unrelated controls similar in age and sex distribution. The mean (SD) age at onset of AD was 80.0 (8.0) years, and the mean (SD) age at last examination of the controls was 78.9 (6.4) years. In both the control and AD groups, 70% of the participants were women. The diagnosis of AD was based on the National Institute of Neurological Disorders and Stroke-Alzheimer’s Disease and Related Disorders Association criteria (McKhann et al. 1984).

All DNA samples were isolated from whole blood and were randomly distributed in genotyping plates. All samples were genotyped on Illumina HumanHap 650Y arrays at the same laboratory. The dataset consisted only of samples that previously passed SNP-based quality control procedures (e.g. gender miscalls and relatedness checks) (Lee JH et al. 2010). Our preliminary analysis was done as a blind study, and the affection status of the samples was only disclosed after the CNV detection procedures were completed.

#### Quality control and CNV detection

Raw intensity array data were normalized within and across samples using Illumina’s BeadStudio software v3.3.7. To maximize CNV discovery, we ran four different CNV calling algorithms, QuantiSNP (Colella et al. 2007), iPattern (Pinto et al. 2011), PennCNV (Wang et al. 2007), and CNVpartition (implemented in BeadStudio). To obtain high-confidence CNV calls, a stringent CNV dataset was generated by taking the CNV calls by iPattern that were also found by at least one additional algorithm (either PennCNV or QuantiSNP). Specifically, each CNV detected by two methods was merged using the outside probe boundaries (i.e. union of the CNVs) as described previously (Pinto et al. 2010), and it needed to overlap in at least 50% of its length. Previously (using Illumina 1M arrays) we showed that stringent CNVs >30 kb detected by both iPattern and QuantiSNP were confirmed by quantitative PCR (qPCR) to be true events at 95% confidence (Pinto et al. 2010). Here, given the lower resolution of the current 650K array, we assumed that a comparable sensitivity would be able to detect large CNVs (>100 kb). To minimize overestimation of reported boundaries, the third algorithm was only used for support. The fourth algorithm, CNV partition, was used to visualize large CNVs.

Poor quality samples were excluded from the study if they met the following criteria: chip call rate < 97%; log R ratio standard deviation > 0.27; B allele frequency standard deviation > 0.17; and PennCNV wave factor > 0.04 or ≤ 0.04 (Diskin et al. 2008). We excluded CNV calls when they failed stringent quality control (QC) criteria: <5 probes, <100 kb size, or low confidence QuantiSNP score (log Bayes factor < 15), as these CNVs were likely to be unreliable at the current array resolution. We also excluded CNV calls within hypervariable centromere proximal bands and those overlapping immunoglobulin regions, as both are known to be prone to artificial CNV calling and thus false discoveries.

Finally, we removed samples that had an excessive number of CNVs detected by each algorithm (i.e. samples with a number of CNV calls exceeding the third quartile plus three times the interquartile range). The resulting cutoff for the number of CNVs per sample was 67 CNV calls for PennCNV, 35 calls for QuantiSNP, and 35 calls for iPattern. Chromosome X and all CNVs > 1 Mb detected by any algorithm were inspected manually. Samples with excessive aggregate length of CNVs, as well as samples with CNVs > 7.5 Mb (likely karyotyping abnormalities) were visually inspected by plotting their intensities and allelic ratios, and removed from burden analyses (supporting information, Table S1).

For the purpose of burden analysis, CNVs with more than 50% of their length overlapping segmental duplications were discarded; CNVs found in >1% of cases and controls were not considered further. A total of 392 cases (106 males, 286 females) and 357 controls (104 males, 253 females) passed all QC steps and were used in subsequent analyses. The female/male ratio and age at onset in the dataset that passed all QC steps remained similar to original dataset: ~70% females, the mean age at onset (SD) of AD cases was 77.1 (8.5) years, and the mean age at last examination (SD) of the controls was 79.5 (6.1) years.

#### CNV burden analyses

To determine whether cases show a greater genome-wide burden of rare CNVs compared with controls, CNV burden analyses were conducted using PLINK v1.07 and a permutation procedure (one-sided, 100,000 permutations) (Purcell et al. 2007). P values were estimated for the number of CNVs per individual (CNV rate), for CNV sample proportion (fraction of samples with one or more CNVs), and...
Table 1 Global rare CNV burden analyses with respect to CNV size and CNV rate

| CNV Rate | CNV Sample Proportion | Total CNV Size (kb) | Average CNV Size (kb) |
|-----------|-----------------------|--------------------|----------------------|
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctrl | Baseline              | Case/ctrl ratio    | P                    |
|          |                       |                     |                      |
| Case/ctr...
Global burden analyses were further extended by stratifying rare CNVs according to size (e.g. >500 kb or >1 Mb) and CNVs with genetic content. None of these strategies revealed significant differences between cases and controls (Tables 1 and 2). For instance, the case/control ratio for all genic CNVs was 0.94 (1.1 for deletions and 0.9 for duplications), and no considerable enrichment was found for CNV size in any range or excess of gene-disrupting CNVs (Tables 1 and 2).

**Candidate novel CNVs**

We observed 12 CNVs >1 Mb that were detected in eight AD cases and four controls. Six CNVs were found only in AD cases and were observed neither in Hispanic controls nor in the DGV (Table S3), suggesting that they might be novel structural abnormalities with potential functional significance for AD. For example, in case NY1811 (age at onset 73), we observed a 1.9 Mb duplication on chr 2p16.3 (Figure S2A) encompassing the entire neurexin1 gene (NRXN1) that encodes a neuronal cell surface protein involved in cell recognition and cell adhesion. Genome-wide CNV studies previously implicated NRXN1 deletions in autism and schizophrenia (Ching et al. 2010; Magri et al. 2010). Our study is the first report describing a duplication of NRXN1 in an AD case.

In AD case NY1261 (age at onset 89), we detected a 1.4 Mb deletion on chromosome 17p13.1-2 and a 563 kb deletion on 3p21.31 (Figure S2B). Together both deletions affect 90 genes, including several genes implicated in synaptic function (DLG4, NLGN2, CHRNBI, GABARAP, and PITPNM3) (Table S2). In AD case RX1107 (age at onset 88), we detected a 2.9 Mb deletion on chromosome 7q35-q36.1 (Figure S2C) that disrupts the CNTNAP2 gene encoding the contactin-associated protein-like 2 protein, a member of the neurexin family that mediates interactions between neurons and glial cells. SNPs in the CNTNAP2 were reported to be significantly associated with schizophrenia and bipolar disorder in GWAS studies (Wang et al. 2010; O’Dushlahane et al. 2011). Intriguingly, variants in CNTNAP2 were also implicated in pseudoexfoliation syndrome (Krumbeigel et al. 2011) among patients who show a selective downregulation of clustatin (CLU) expression in their eyes (Zenk el et al. 2006). Notably, the association between CLU SNPs and AD was confirmed in several studies at a genome-wide significance level (Harold et al. 2009; Lambert et al. 2009; Carrasquillo et al. 2010).

In addition, we generated a list of 29 genes that were affected by CNVs in two or more AD patients, that were not seen in our Caribbean Hispanic controls, and that were absent or rare in the DGV and 5000 Caucasian controls (Table 3). Some of these genes have potential functional connections to neurological disorders. For instance, in two AD patients, we detected deletions affecting the protein-tyrosine phosphatase receptor-type delta gene (PTPRD), which has been associated with restless legs syndrome (Morris et al. 2010; Yang et al. 2011). One of the deletions (128 kb) removes exon 9 of PTPRD, and the other one (135 kb) removes exon 4. Also, two patients (RM4073 and RM4285) had a 622 kb duplication on chr 5q12.1 affecting five genes, including the NDUF42 gene that encodes a chaperone for mitochondrial complex I assembly and that was found to be implicated in attention-deficit/hyperactivity disorder (Lesch et al. 2011). Two other duplications were detected at 3p26.3, disrupting the contactin 6 gene (CNTN6). Structural and sequence variations in several members of the contactin gene family were associated with neuropsychiatric disorders (e.g. schizophrenia and autism) (Fernandez et al. 2008; Burbach and van der Zwaag 2009; Cottrell et al. 2011).

**DISCUSSION**

We conducted a genome-wide scan for large CNVs (≥100 kb) in a case-control dataset of Caribbean Hispanic origin that was previously investigated in a SNP-based GWAS (Lee JH et al. 2010). To generate results with high confidence, we focused on CNVs that were identified by at least two algorithms. We detected 1774 stringent CNVs (Table S4). First, we tested the hypothesis that rare CNVs (≥1%) with a potentially strong impact on AD risk in individual patients might contribute to the overall disease risk, as was previously observed in other common neuropsychiatric disorders (Kirow et al. 2009; Zhang et al. 2009; Glessner et al. 2010). However, the burden analyses of rare CNVs did not find significant differences between cases and controls in CNV rate, total or average CNV size, or the number of genes affected by CNVs.

In addition, we conducted a case-control analysis of large genic CNVs, including common variants, using PLINK regional analysis. The only nominally significant result that survived qPCR confirmation was detected for a duplication on chromosome 15q11.2 affecting up to five genes, including NIPA1 and CYFIP1 (P = 0.037). Duplications affecting the NIPA1 and CYFIP1 in control populations are cataloged at the DGV based on four studies (Pinto et al. 2007; Zogopoulos et al. 2007; Itsara et al. 2009; Shaikh et al. 2009) with similar frequencies to our controls (0.5%); this duplication was reported in 24 out of 5056 individuals.
| Sample   | Chr:start-end                  | Size    | CNV type | RefSeq genes   | Genes overlapped by CNVs in ≥ 2 AD-cases and absent in Caribbean Hispanic controls without-AD<sup>a</sup> | DGV<sup>b</sup> (diverse populations) | 5000 additional Caucasian-controls<sup>b</sup> | #RefSeq genes | Disrupted genes | Genes with exons_in_cnv | Mouse_MGI neuronal-phenotypes | Synaptic genes | Flanked_by_segdup<sup>a</sup> |
|----------|-------------------------------|---------|----------|----------------|-----------------------------------------------------------------------------------------------|----------------------------------------|---------------------------------------------|---------------|------------------|------------------------|---------------------------------|----------------|------------------------|
| RM3803   | 7:16882053-17397887           | 515,835 | Gain     | AHR,AGR3      | AHR,AGR3                                                                                       | Rare                                   | Absent           | 2             | AGR3            | AHR,AGR3                                  | —                  | —                      | —                      |
| RX0319   | 7:16887616-17381309           | 493,694 | Gain     | AHR,AGR3      | AHR,AGR3                                                                                       | Rare                                   | Absent           | 2             | FBLN1          | ATXN10, FBLN1                            | FBLN1              | —                      | —                      |
| RM4226   | 22:44368703-44637724          | 269,022 | Gain     | ATXN10, FBLN1 | ATXN10, FBLN1                                                                                   | Absent                                 | Absent           | 2             | FBLN1          | ATXN10, FBLN1                            | —                  | —                      | —                      |
| NY2608   | 22:44315725-44638874          | 323,150 | Gain     | ATXN10, FBLN1 | ATXN10, FBLN1                                                                                   | Absent                                 | Absent           | 2             | C1orf101, ADSS | C1orf101, ADSS                           | —                  | —                      | —                      |
| RM0949   | 1:242659903-242869328         | 209,426 | Gain     | C1orf101, ADSS| C1orf101, ADSS                                                                                   | Absent                                 | Absent           | 2             | C1orf101, ADSS | C1orf101, ADSS                           | —                  | —                      | —                      |
| RX1192   | 1:242659903-242869328         | 209,426 | Gain     | C1orf101, ADSS| C1orf101, ADSS                                                                                   | Absent                                 | Absent           | 2             | C1orf101, ADSS | C1orf101, ADSS                           | —                  | —                      | —                      |
| RM6264   | 3:643003-1118424              | 475,422 | Gain     | CNTN6         | CNTN6                                                                                           | Rare                                   | Rare             | 1             | CNTN6          | CNTN6                                      | —                  | —                      | —                      |
| NY0553   | 3:1118424-1756739             | 638,316 | Gain     | CNTN6         | CNTN6                                                                                           | Rare                                   | Rare             | 1             | CNTN6          | CNTN6                                      | —                  | —                      | —                      |
| RM5037   | 5:60056277-1118424            | 374,987 | Gain     | EEF2K, CDR2, POLR3E, C16orf52, PDZD9, WWA3A | EEF2K, CDR2, POLR3E, C16orf52, PDZD9, WWA3A                                                      | Rare                                   | Rare             | 6             | CDR2, C16orf65 | EEF2K, C16orf65, CDR2, POLR3E, C16orf52, WWA3A | —                  | —                      | Yes                    |
| RX0490   | 16:21907345-22282331          | 374,987 | Gain     | EEF2K, CDR2, POLR3E, C16orf52, PDZD9, WWA3A | EEF2K, CDR2, POLR3E, C16orf52, PDZD9, WWA3A                                                      | Rare                                   | Rare             | 6             | CDR2, C16orf65 | EEF2K, C16orf65, CDR2, POLR3E, C16orf52, WWA3A | —                  | —                      | Yes                    |
| RM4285   | 5:60056277-60678352           | 622,076 | Gain     | ERC28, ELOVL7, ZSWIM6, C5orf43, NDUFAF2 | ERC28, ELOVL7, ZSWIM6, C5orf43, NDUFAF2                                                        | Absent                                 | Absent           | 5             | ZSWIM6         | ERCC8, ELOVL7, ZSWIM6, C5orf43, NDUFAF2   | —                  | —                      | —                      |
| RM4073   | 5:60056277-60678352           | 622,076 | Gain     | ERC28, ELOVL7, ZSWIM6, C5orf43, NDUFAF2 | ERC28, ELOVL7, ZSWIM6, C5orf43, NDUFAF2                                                        | Absent                                 | Absent           | 5             | ZSWIM6         | ERCC8, ELOVL7, ZSWIM6, C5orf43, NDUFAF2   | —                  | —                      | —                      |

(continued)
### Table 3 Continued

| Sample  | Chr:start-end     | Size   | CNVtype | RefSeq genes          | Genes overlapped by CNVs in ≥ 2 AD-cases and absent in Caribbean Hispanic controls without-AD<sup>a</sup> | DGV<sup>b</sup> (diverse populations) | 5000 additional Caucasian-controls<sup>b</sup> #RefSeq genes | Disrupted_ genes | Genes with exons_in_cnv | Mouse_MGI_neuronal-phenotypes | Synaptic_genes | Flanked_by_segdups |
|---------|-------------------|--------|---------|------------------------|------------------------------------------------------------------------------------------------|--------------------------------------|----------------------------------------------------------|----------------|------------------------|---------------------------------|----------------|---------------------|
| RM5824  | 3:99076070-99196494 | 120,425 | Gain    | MINA, GABRR3, CRYBG3 | MINA, GABRR3, CRYBG3 | Absent | Absent | 3 | GABRR3 | MINA, GABRR3, CRYBG3 | — | — |
| RX1129  | 3:99085427-99196494 | 111,068 | Gain    | MINA, GABRR3, CRYBG3 | MINA, GABRR3, CRYBG3 | Absent | Absent | 2 | MSH3, RASGRF2 | MSH3, RASGRF2 | — | — |
| NY0709  | 5:80056924-80293554 | 236,631 | Gain    | MSH3, RASGRF2        | MSH3, RASGRF2        | Absent | Absent | 3 | GABRR3 | MINA, GABRR3, CRYBG3 | — | — |
| NY0725  | 5:80056924-80293554 | 236,631 | Gain    | MSH3, RASGRF2        | MSH3, RASGRF2        | Absent | Absent | 2 | MSH3, RASGRF2 | MSH3, RASGRF2 | — | — |
| RM4073  | 9:8893857-9022105  | 128,249 | Loss    | 9p24.1, 9p23         | PTPRD                | Absent | Absent | 1 | PTPRD | PTPRD       | PTPRD | PTPRD |
| NY2092  | 9:9850735-9985938  | 135,204 | Loss    | PTPRD                | PTPRD                | Absent | Absent | 1 | PTPRD | PTPRD       | PTPRD | PTPRD |
| NY1942  | 3:12610706-12792622 | 181,917 | Gain    | RAF1, TMEM40         | RAF1, TMEM40         | Rare   | Rare   | 2 | RAF1   | RAF1, TMEM40 | — | — |
| RM5553  | 3:1261545-12781123 | 165,379 | Gain    | RAF1, TMEM40         | RAF1, TMEM40         | Rare   | Rare   | 2 | RAF1   | RAF1, TMEM40 | — | — |
| RX1208  | 13:19109434-19362188 | 252,755 | Gain    | MPHOSPH8, PSPC1, ZMYM5 | MPHOSPH8, PSPC1, ZMYM5 | Rare   | Rare   | 3 | MPHOSPH8 | MPHOSPH8, PSPC1, ZMYM5 | — | — |
| NY0295  | 13:19109434-19362188 | 252,755 | Gain    | MPHOSPH8, PSPC1, ZMYM5 | MPHOSPH8, PSPC1, ZMYM5 | Rare   | Rare   | 3 | MPHOSPH8 | MPHOSPH8, PSPC1, ZMYM5 | — | — |
| NY1684  | 13:19109434-19362188 | 252,755 | Gain    | MPHOSPH8, PSPC1, ZMYM5 | MPHOSPH8, PSPC1, ZMYM5 | Rare   | Rare   | 3 | MPHOSPH8 | MPHOSPH8, PSPC1, ZMYM5 | — | — |

<sup>a</sup> Controls from the Caribbean Hispanic dataset.<br>
<sup>b</sup> Controls from databases that were not specifically screened for AD.
NIPA1 encodes a magnesium transporter associated with early endosomes in neuronal and epithelial cells (Rainier et al. 2003; van der Zwaag et al. 2010). CYFIP1 forms a complex at synapses with the fragile X mental retardation protein (FMRP) and elf4E (FMRP-CYFIP1-elf4E complex). FMRP acts as an APP translation repressor (Lee EK et al. 2010), releasing CYFIP1 from the FMRP-CYFIP1-elf4E complex in response to synaptic stimulation (Napoli et al. 2008). Therefore, unbalanced dosage of CYFIP1 might result in altered APP turnover in AD patients. Of note, this region belongs to a larger region at chromosome 15q11-q13 that has been introduced as one of the most reliable "cytogenetic regions of interest" for genomic aberrations in autism spectrum disorders (Vorstman et al. 2006). It is important that the association between AD and the 15q11.2 duplication be validated in follow-up studies using large case-control datasets.

Our study does not support the previously reported marginal association between AD and the ~500 kb duplication on chromosome 15q13.3 affecting the CHRNA7 locus in a genome-wide scan of a North American dataset (Heinzen et al. 2010), which is 9.5 Mb away from the duplication on 15q11.2 discussed above. We observed an equal number of cases (n = 2; 0.5%) and controls (n = 2; 0.6%) with duplications affecting CHRNA7, whereas Heinzen et al. detected this CNV in six cases (2%) and one control (0.3%) (Heinzen et al. 2010).

In addition, a higher copy number of a complex multiallelic segment (DGV variation_0316) containing the olfactory receptor genes on chromosome 14q11.2 was reported to be associated with a decrease in chromosome 14q11.2 was reported to be associated with a decrease in negative/positive rate. In general, a case-control setting can only test common CNVs from SNP-intensity data could lead to a high false positive rate. In summary, in a stringent genome-wide investigation for the identification of cases (n = 2; 0.5%) and controls (n = 2; 0.6%) with duplications affecting CHRNA7, whereas Heinzen et al. detected this CNV in six cases (2%) and one control (0.3%) (Heinzen et al. 2010).

We did not detect CNVs (including common variants) that affect the well-confirmed AD loci reported by large GWAS (CLU, PICALM, BIN1, CR1, MS4A4A/MS4A6E, CD2AP, CD33, EPHA1, and ABCA7) (Harold et al. 2009; Lambert et al. 2009; Carraquillo et al. 2010; Hollingworth et al. 2011; Naj et al. 2011). However, as the array technology used in the current study has limitations in detecting small CNVs, future studies must carefully assess the new AD loci using a qPCR approach to detect small CNVs. For instance, by using a higher copy number of a complex multiallelic segment (DGV variation_0316) containing the olfactory receptor genes on chromosome 14q11.2 was reported to be associated with a decrease in age at onset of AD using genotypes obtained from Affymetrix SNP 6.0 arrays (controls were not evaluated) (Shaw et al. 2011). Although this region is ~200 kb in size, it is poorly covered with SNPs in the 650Y array used for our study (three SNPs). Therefore, we were unable to assess the contribution of this region to AD in our dataset.

ACKNOWLEDGMENTS

We thank Boobra Thiruvanichandrapuram for technical assistance and the Centre for Applied Genomics at the Hospital for Sick Children for data sharing and database support. This work was supported by National Institutes of Health and National Institute on Aging grants R37-Ag15473 (R.M.) and P01-Ag07232; by the Blanchett Hooker Rockefeller Foundation; by the Charles S. Robertson Gift from the Banbury Fund (R.M.); by the W. Garfield Weston Foundation (E.R.); by the Canadian Institutes of Health Research, Ontario Research Fund (E.R. and P.S.H.); and by the Howard Hughes Medical Institute, the Wellcome Trust, the Alzheimer Society of Ontario, the Canada Foundation for Innovation, the Ontario Mental Health Foundation, Genome Canada, and the Alzheimer Society of Canada (P.S.H.). D.P. is supported by Canadian Institutes of Health Research fellowship 213997. S.W.S. holds the GlaxoSmithKline-CIHR Endowed Chair in Genetics and Genomics at the Hospital for Sick Children and the University of Toronto.
itsaras, A., G. M. Cooper, C. Baker, S. Girirajan, J. Li, Hollingworth, P., D. Harold, R. Sims, A. Gerrish, J. C. Lambert, Pinto, D., C. Marshall, L. Feuk, and S. W. Scherer, 2007. Copy number analysis of large copy number variants and hotspots of human genetic disease. Am. J. Hum. Genet. 84(2): 148–161.

Kirov, G., D. Grozeva, N. Norton, D. Ivanov, K. K. Mantrippragada et al., 2009. Support for the involvement of large copy number variants in the pathogenesis of schizophrenia. Hum. Mol. Genet. 18(8): 1497–1503.

Krumbiegel, M., F. Pasutto, U. Schlotzer-Schrehardt, S. Uebe, M. Zénkel et al., 2011. Genome-wide association study with DNA pooling identifies variants at CNTNAP2 associated with pseudoxenfoliation syndrome. Eur. J. Hum. Genet. 19(2): 186–193.

Lambert, J. C., S. Heath, G. Even, D. Campion, K. Sleegers et al., 2010. The clinical context of copy number pathogenesis of schizophrenia. Hum. Mol. Genet. 19(10): 2059–2068.

Lee, J. H., R. Cheng, S. Barral, C. Reitz, M. Medrano et al., 2010. Duplication of amyloid precursor protein (APP), but not prion deficiency in eyes with pseudoexfoliative material. Invest. Ophthalmol. Vis. Sci. 47(5): 1982–1989.

McKhann, G., D. Drachman, M. Folstein, R. Katzman, D. Price et al., 1984. Clinical diagnosis of Alzheimer’s disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer’s Disease. Neurology 34(7): 939–944.

McNaughton, D., W. Knight, R. Guerreiro, N. Ryan, J. Lowe et al., 2010. 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.e13–426.e21.

Morris, L. G., S. Veeriah, and T. A. Chan, 2010. Genetic determinants at the interface of cancer and neurodegenerative disease. Oncogene 29(24): 3453–3464.

Naj, A. C., G. Jun, G. W. Beecham, L. S. Wang, B. N.vardarajan et al., 2011. Common variants at 5q44A4/MSA4A6E, CD2AP, CD33 and EPFA1 are associated with late-onset Alzheimer’s disease. Nat. Genet. 43(5): 436–441.

Napolì, L., V. Mercaldo, P. P. Boyl, B. Eleuteri, F. Zalfa et al., 2008. The fragile X syndrome protein represses activity-dependent translation through CYFIP1, a new 4E-BP. Cell 134(6): 1042–1054.

O’Dushlaine, C., E. Kenny, E. Heron, G. Donohoe, M. Gill et al., 2011. Molecular pathways involved in neuronal cell adhesion and membrane scaffolding contribute to schizophrenia and bipolar disorder susceptibility. Mol. Psychiatry 16(3): 286–292.

Pinto, D., C. Marshall, L. Feuk, and S. W. Scherer, 2007. Copy number variation in control population cohorts. Hum. Mol. Genet. 16 Spec No. 2: R168–73.

Pinto, D., A. T. Pagnamenta, L. Klei, R. Anney, D. Merico et al., 2010. Functional impact of global rare copy number variation in autism spectrum disorders. Nature 466(7304): 368–372.

Pinto, D., K. Darvishi, X. Shi, D. Rajan, D. Rigler et al., 2011. Comprehensive assessment of array-based platforms and calling algorithms for detection of copy number variants. Nat. Biotechnol. 29: 512–520.

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. Ferreira et al., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81(3): 559–575.

Rainier, S., J. H. Chai, D. Tokarz, R. D. Nicholls, and J. K. Fink, 2003. NIPA1 gene mutations cause autosomal dominant hereditary spastic paraplegia (SPG6). Am. J. Hum. Genet. 73(4): 967–971.

Reitz, C., R. Cheng, E. Rogaeva, J. H. Lee, S. Tokuhiro et al., 2011. Meta-analysis of the Association Between Variants in SORL1 and Alzheimer Disease. Arch. Neurol. 68(1): 99–106.

Rogaeva, E., T. Kawarai, and P. S. George-Hyslop, 2006. Genetic complexity of Alzheimer’s disease: successes and challenges. J. Alzheimers Dis. 9(3, Suppl): 381–387.

Rogaeva, E., Y. Meng, J. H. Lee, Y. Gu, T. Kawarai et al., 2007. The neuronal soritin-related receptor SORL1 is genetically associated with Alzheimer disease. Nat. Genet. 39(2): 168–177.

Shaikh, T. H., X. Gai, J. C. Perin, J. T. Glessner, H. Xie et al., 2009. High-resolution mapping and analysis of copy number variations in the human genome: a data resource for clinical and research applications. Genome Res. 19(9): 1682–1690.

Shaw, C. A., Y. Li, J. Wisniewska, S. Chasse, S. N. Zaidi et al., 2011. Olfactory copy number association with age at onset of Alzheimer disease. Neurology 76(15): 1302–1309.

Swaminathan, S., S. Kin, L. Shen, S. L. Rüscher, T. Foroud et al., 2011. Genetic copy number analysis in Alzheimer’s Disease and mild cognitive impairment: an ADNI study. Int. J. Alzheimers Dis. 2011: 729478.

van der Zwaag, B., W. G. Staal, R. Hochstenbach, M. Poot, H. A. Spierenburg et al., 2011. Geometric copy number analysis in Alzheimer’s Disease and mild cognitive impairment: an ADNI study. Int. J. Alzheimers Dis. 2011: 729478.

Wang, K., M. Li, D. Hadley, R. Liu, J. Glessner 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(11): 1665–1674.

Wang, K. S., X. F. Liu, and N. Aragam, 2010. A genome-wide meta-analysis identifies novel loci associated with schizophrenia and bipolar disorder. Schizophr. Res. 124(1–3): 192–199.

Yang, Q., L. Li, H. Yang, G. Q. Shen, Q. Chen et al., 2011. Family-based and population-based association studies validate PTPRD as a risk factor for restless legs syndrome. Mov. Disord. 26(3): 516–519.

Zenkel, M., F. E. Kruse, A. G. Junemann, G. O. Naumann, and U. Schlotzer-Schrehardt, 2006. Clusterin deficiency in eyes with pseudoxenfoliation syndrome may be implicated in the aggregation and deposition of pseudoxenfoliative material. Invest. Ophthalmol. Vis. Sci. 47(5): 1982–1990.

Zhang, D., L. Cheng, Y. Qian, N. Allely-Rodriguez, J. R. Kelsoe et al., 2009. Singleton deletions throughout the genome increase risk of bipolar disorder. Mol. Psychiatry 14(4): 376–380.