Molecular Psychiatry (2013), 1–6

www.nature.com/mp

ORIGINAL ARTICLE

Rare autosomal copy number variations in early-onset familial Alzheimer's disease

BV Hooli[1], ZM Kovacs-Vajna[2], K Mullin[1], MA Blumenthal[1], M Mattheisen[3], C Zhang[1], C Lange[4], G Mohapatra[5], L Bertram[6] and RE Tanzi[1]

Over 200 rare and fully penetrant pathogenic mutations in amyloid precursor protein (*APP*), presenilin 1 and 2 (*PSEN1* and *PSEN2*) cause a subset of early-onset familial Alzheimer's disease (EO-FAD). Of these, 21 cases of EO-FAD families carrying unique *APP* locus duplications remain the only pathogenic copy number variations (CNVs) identified to date in Alzheimer's disease (AD). Using high-density DNA microarrays, we performed a comprehensive genome-wide analysis for the presence of rare CNVs in 261 EO-FAD and early/mixed-onset pedigrees. Our analysis revealed 10 novel private CNVs in 10 EO-FAD families overlapping a set of genes that includes: *A2BP1*, *ABAT*, *CDH2*, *CRMP1*, *DMRT1*, *EPHA5*, *EPHA6*, *ERMP1*, *EVC*, *EVC2*, *FLJ35024* and *VLDLR*. In addition, CNVs encompassing two known frontotemporal dementia genes, *CHMP2B* and *MAPT* were found. To our knowledge, this is the first study reporting rare gene-rich CNVs in EO-FAD and early/mixed-onset AD that are likely to underlie pathogenicity in familial AD and perhaps related dementias.

Molecular Psychiatry advance online publication, 11 June 2013; doi:10.1038/mp.2013.77

Keywords: Alzheimer's; CNVs; *MAPT*; *CHMP2B*; *VLDLR*

INTRODUCTION

Alzheimer's disease (AD) is a genetically complex and heterogeneous disorder. Family history is the second biggest risk factor in AD following age.[1] The four established AD genes, amyloid precursor protein (*APP*), presenilin 1 and 2 (*PSEN1*, *PSEN2*) and apolipoprotein E (*APOE*) have been estimated to account for 30–50% of the genetic variance of AD.[2] Attempts to identify additional AD genes have been largely limited to the search for simple DNA sequence variants (mutations and single-nucleotide polymorphisms) that influence AD susceptibility or time-to-onset.[3,4] Genome-wide association studies have identified several common AD-associated polymorphisms with very small effect sizes.[5]

Recent studies estimate that structural variations in the genome, including copy number variations (CNVs)[6] make a significant contribution to genetic and phenotypic variation.[7,8] CNVs vary in size from a few kilobases (kb) to several megabases (Mb), but in most instances arbitrarily indicate DNA segments that are >1 kb in length. Several recent studies have successfully identified CNVs that underlie pathogenesis of complex diseases such as autism,[9–11] schizophrenia[12,13] and human immunodeficiency virus.[14,15] However, the contribution of CNVs to disease risk is highly complex. Information regarding CNV penetrance, frequency and functional implications is still rudimentary.[16,17,18] Nonetheless, the influence of large CNVs on phenotype is estimated to be clinically significant[16] and identification of large, rare CNVs predisposing to various diseases has proven to be successful.[19–21] Based on the above facts, we set out to identify rare pathogenic CNVs in our collection of early-onset familial AD (EO-FAD) cohorts. Specifically, we searched for large (>100 kb), novel (not previously reported in the database of genomic variants (DGV)) CNVs that co-segregate with disease status within pedigrees.

MATERIALS AND METHODS

Family cohorts

Two large family-based AD sample sets were used in this study: The National Institute of Mental Health Alzheimer's Disease Genetics Initiative Study (NIMH) consisting of 1439 individuals from 436 families of which 131 early/mixed-onset families were included here (that is, families with $\geqslant 1$ affected individual showing an age of onset <65) and the National Cell Repository for Alzheimer's Disease (NCRAD) consisting of 1108 samples from 331 pedigrees, of which 130 early/mixed-onset AD families were used here. Both AD family cohorts have been previously described.[22] Overall, a total of 1009 subjects from 261 early/mixed-onset families were used in the CNV analysis.

Genotyping and CNV analyses

DNA samples from 1009 subjects in the NIMH and NCRAD AD family cohorts were processed on Affymetrix Human Genome-Wide SNP 6.0 arrays (Affymetrix Inc., Santa Clara, CA, USA). A total of 43 samples that failed to pass quality control, including gender validation, array quality and large chromosomal abnormalities were excluded from the study, as described in detail elsewhere.[23,24] CNVs >100 kb were inferred from probe intensity data using PennCNV[25] (1 May 2010 version). Copy number polymorphisms (CNPs) identified in these 261 early/mixed-onset families based on $>70\%$ overlap with CNPs in DGV (last update November 2010) were excluded from further analysis.[26–29] The latest version of DGV lists close to 67K CNVs in about 16K genomic loci.

Genomic segment correlation with disease status

The CNV segments identified using PennCNV were analyzed for segregation with disease status in the affected subjects of our pedigree samples

[1]Department of Neurology, Genetics and Aging Research Unit, MassGeneral Institute for Neurodegenerative Disease, Massachusetts General Hospital, Charlestown, MA, USA; [2]Department of Information Engineering, University of Brescia, Brescia, Italy; [3]Channing Laboratory, Brigham and Women's Hospital, Boston MA, USA; [4]Department of Biostatistics, Harvard School of Public Health, Boston, MA, USA; [5]Molecular Pathology Unit, Massachusetts General Hospital, Boston, MA, USA and [6]Max-Planck Institute for Molecular Genetics, Neuropsychiatric Genetics Group, Berlin, Germany. Correspondence: Dr RE Tanzi, Department of Neurology, Genetics and Aging Research Unit, MassGeneral Institute for Neurodegenerative Disease, Massachusetts General Hospital, 114 16th Street, Charlestown, MA 02129, USA.
E-mail: tanzi@helix.mgh.harvard.edu
Received 24 January 2013; revised 19 March 2013; accepted 15 April 2013

using two methods: (a) utilizing a novel algorithm that we developed for this study to perform CNV segregation analysis (genomic segment correlator, GSC) and (b) visual analysis for segregation using the UCSC genome browser. The physical start and end positions of the CNVs, phenotype and pedigree information of each sample were supplied as input to GSC (Supplementary Figure 3). CNVs with >50% overlap in genomic region were treated as a single event because of the ambiguity of CNV end points in calls inferred from micro-array intensity data. Each family was separately analyzed for CNVs that co-segregated with the disease status. Genetic heterogeneity was accounted for by excluding late-onset AD patients (onset >65 years of age) from segregation analysis within pedigrees. Only those CNV segments that were present in affected members of the test family, and absent in 'controls' were subjected to further analysis. The 'control' samples are unaffected subjects from the entire set of early/mixed-onset families, as well as, the remaining late-onset AD families—a total of 1222 unaffected subjects. Following visual confirmation on UCSC Genome Browser using custom tracks in BED format,[30] CNVs showing inheritance patterns consistent with disease causation were prioritized for CNV confirmation. See supplementary for more details on GSC.

For visual analysis in the UCSC Genome Browser method, custom tracks were created comprising all the PennCNV segments, and were integrated into the browser. A window spanning a 2-Mb genomic interval was visually analyzed for CNVs segregating with disease—and in a sliding window manner we then covered all autosomes. Rare CNVs showing pathogenic inheritance in the early/mixed-onset AD families were compared against CNVs detected in the rest of the NIMH and NCRAD late-onset families ($n = 506$) consisting of 1547 subjects. Only rare CNVs that co-segregated with AD exclusively in affected subjects in the early/mixed-onset families were included in further analysis.

CNV validation

CNVs that co-segregated with AD were validated using Fluidigm Digital 48.776 arrays (Fluidigm Corporation, San Francisco, CA, USA) and TaqMan copy number probes.[31] Depending on the availability of tissue samples, several CNVs were also confirmed using fluorescence *in-situ* hybridization as described previously[32] using Epstein–Barr virus-transformed lymphoblast cell lines derived from the subjects.

RESULTS

PennCNV was utilized to make CNV calls in all 261 early/mixed-onset AD families. The total number of CNVs detected, CNV burden (average CNVs per individual), average CNV segment size, copy number (CN) gains and CN losses detected per sample are listed in Table 1. The overall frequency of CNVs identified in nearly 1000 subjects from 261 early/mixed-onset AD families (Table 1) are similar to that described in previous studies utilizing the Affymetrix SNP 6.0 microarray.[23,33] CNV segments that showed <70% overlap with CNPs reported on DGV were analyzed for segregation with disease. These analyses confirmed *APP* locus duplication previously reported in two EO-FAD families.[34] In addition to these, we observed 10 (NIMH sample) and 7 (NCRAD families) novel CNVs, which co-segregated with disease exclusively in EO-FAD families. These 17 CNVs were absent from all other NIMH and NCRAD AD families (total of 2796 samples from 825 families), were not found in any unaffected subjects, and did not overlap with published CNVs or CNPs in DGV—suggesting that they are AD specific. Subsequent CNV confirmation assays validated 10 of the 17 CNVs (Table 2); the majority of false-positive CNVs (five out of seven) involved genomic duplications. All the 10 validated CNVs were novel and private in the early/mixed-onset AD families. Five of the 10 novel CNVs identified (Table 2) were heterozygous deletions for which the average familial onset age was 60.5 years, and the average CNV size was 224 kb. These CNVs overlapped with the genes: *CHMP2B*, *POU1F1*, *KANK1*, *DMRT1*, *DMRT3*, *FLJ35024*, *VLDLR* and *A2BP1*. Five of the 10 novel CNVs identified were CN gains (Table 2) for which the average familial onset age was 58.75 years, and the average CNV size was 501 kb. These CNVs overlapped with the genes: *MAPT*, *CDH2*, *ERMP1*, *EVC*, *EVC2*, *CRMP1* and *EPHA6*. For 6 of the 10 CNVs, the *APOE-ε4* allele also co-segregated with AD, suggesting that the genes affected by the CNVs and the *APOE-ε4* allele combine as genetic modifiers for AD risk. Note that two CNVs were found to encompass known frontotemporal lobar dementia genes, that is, *CHMP2B* and *MAPT* (Table 2).

Finally, we set out to confirm that the 10 EO-FAD families co-segregating the novel CNVs with AD do not contain mutations in genes that have been previously implicated in EO-FAD and frontotemporal lobar dementia.[35] To this end, we re-sequenced *APP*, *PSEN1*, *PSEN2*, *MAPT* and *GRN* for pathogenic mutations in probands from the entire collection of NIMH and NCRAD families, including the 261 early/mixed-onset families (Supplementary Table 1). No pathogenic sequence variants were detected in the 10 EO-FAD families that co-segregated with novel private CNVs with AD. However, 11 pathogenic sequence variants were identified in other families, along with a novel mutation (M84V) in *PSEN1*. Similar to a previous finding,[36] the known pathogenic *PSEN1* EO-FAD mutation, A79V, was found to co-segregate with AD in a late-onset AD family (average onset age = 74.5 years).

DISCUSSION

The primary goal of this study was to identify novel CNVs that co-segregate with EO-FAD in families for which mutations in *APP, PSEN1* and *PSEN2* have been ruled out. To date, the only established CNV in AD is the duplication of the *APP* gene.[34] Considering the exceedingly rare frequency of the *APP* CNV in EO-FAD, we searched for other rare and novel CNVs co-segregating with AD in EO-FAD kindreds. For this purpose, we developed and used a novel algorithm GSC (Supplementary Figure 3). Using the GSC, we identified 10 novel and private CNVs in 10 EO-FAD families. The potential pathogenicity of these 10 novel CNVs is supported by the nature of these mutations—all deletions or duplications within various genes, observed specifically in EO-FAD patients with age of onset <65 years. The potential causality of these novel and rare CNVs for EO-FAD is further supported by their

Table 1. Overview of CNV burden in the two collections of Alzheimer's disease family cohorts

Cohort	*Families*	*Subjects*		*Total (AO)*	*APOE-ε4 positive (%)*	*CNVs detected (CNV burden)*	*Avg. size (bp)*	*No. of CN loss (burden)*	*No. of CN gain (burden)*
NCRAD	130	498	Affected	332 (63.31)	250 (75.3)	4581 (13.8)	351 343	2503 (7.54)	2078 (6.26)
			Unaffected	161 (63.83)	91 (56)	2093 (13)	251 427	1041 (6.47)	1052 (6.53)
NIMH	131	511	Affected	317 (64.51)	236 (74.4)	5284 (16.7)	241 910	2446 (7.72)	2837 (8.95)
			Unaffected	194 (70.3)	100 (51.5)	3403 (17.54)	240 417	1706 (8.8)	1697 (8.75)

Abbreviations: AO, average onset age; CN, copy number; CNV, copy number variation; CNV burden, average CNVs per individual; EO-FAD, early-onset familial Alzheimer's disease; NCRAD, National Cell Repository for Alzheimer's Disease; NIMH, National Institute of Mental Health Alzheimer's Disease Genetics Initiative Study. Overall summary of CNVs identified from microarray probe intensity data in EO-FAD and early/mixed-onset Alzheimer's family samples. The average values per subject are indicated in the parenthesis. The number and fraction of subjects positive for the ε4 risk allele are listed in the column APOE- ε4 positive.

Table 2. List of rare CNVs identified in NIMH and NCRAD EO-FAD and early/mixed-onset AD pedigrees

CNV ID	*Overlapping gene(s)*	*Subject IDs*	*Dx*	*Onset*	*Array CN*	*CN state*	*CNV region (hg18)*	*Size (kB)*
FAD-CNV1	*ERMP1*	*ND1*						
		ND1-II.25	AC	40	1	Loss	Chr9: 5 744 105–5 867 748	124
		ND1-II.06	PS	42	1	Loss	Chr9: 5 744 105–5 867 748	124
FAD-CNV2	*EVC2, EVC, CRMP1*	*ND2*						
		ND2-II.94	AC	52	3	Gain	Chr4: 5 602 184–5 837 823	236
		ND2-II.75	AC	56	3	Gain	Chr4: 5 602 184–5 845 805	244
FAD-CNV3	*A2BP1, ABAT*	*ND3*						
		ND3-II.26	PR	56	1	Loss	Chr16: 7 991 014–8 100 555	110
		ND3-II.25	PS	68	1	Loss	Chr16: 7 994 156–8 100 555	106
FAD-CNV4	*EPHA5*	*ND4*						
		ND4-II.24	PR	37	3	Gain	Chr4: 63 268 479–63 813 833	545
		ND4-II.70	PS	63	3	Gain	Chr4: 63 268 479–63 809 059	541
FAD-CNV5	*CDH2*	*ND5*						
		ND5-II.84	AC	64	3	Gain	Chr18: 23 693 824–24 181 680	488
		ND5-II.72	PS	65	3	Gain	Chr18: 23 693 824–24 180 173	486
		ND5-II.73	PS	--	2	Diploid	--	--
FAD-CNV6	*EPHA6*	*ND6*						
		ND6-II.50	PR	64	3	Gain	Chr3: 96 949 558–97 684 405	735
		ND6-II.81	AC	69	3	Gain	Chr3: 96 937 158–97 678 067	741
FAD-CNV7	*KANK1, DMRT1*	*NH1*						
		NH1-II.32	AC	64	1	Loss	Chr9: 587 476–992 280	405
		NH1-II.85	ND	--	2	Diploid	--	--
		NH1-II.86	PR	65	1	Loss	Chr9: 666 266–992 280	326
		NH1-II.93	ND	--	2	Diploid	--	--
		NH1-II.94	ND	--	1	Loss	Chr9: 589 612–992 280	403
FAD-CNV8	*CHMP2B, POU1F1*	*NH2*						
		NH2-II.06	AC	61	1	Loss	Chr3: 87 319 231–87 650 334	331
		NH2-II.45	PR	72	1	Loss	Chr3: 87 319 617–87 650 334	331
		NH2-II.70	ND	--	2	Diploid	--	--
		NH2-II.90	ND	--	2	Diploid	--	--
FAD-CNV9	*FLJ35024, VLDLR*	*NH3*						
		NH3-II.56	PR	62	1	Loss	Chr9: 2 414 322–2 565 408	151
		NH3-II.57	AC	75	1	Loss	Chr9: 2 414 322–2 565 408	151
FAD-CNV10	*MAPT*	NH4						
		NH4-II.92	PR	59	3	Gain	Chr17: 41 292 942–41 466 517	173
		NH4-II.93	PR	60	3	Gain	Chr17: 41 292 942–41 467 674	174
		NH4-II.64	PR	61	3	Gain	Chr17: 41 300 173–41 467 674	167
		NH4-II.94	ND	--	2	Diploid	--	
		NH4-II.95	ND	--	2	Diploid	--	
		NH4-II.80	PR	71	2	Diploid	--	
ND-APP[a]	*APP*	*ND7*						
		ND7-II.36	AC	49	3	Gain	Chr21: 26 125 668–26 505 191	380
		ND7-II.67	AC	52	3	Gain	Chr21: 26 125 668–26 523 359	398
		ND7-II.03	PS	70	2	Diploid	--	--
		ND7-II.98	ND	69	2	Diploid	--	--
		ND7-II.44	ND	74	2	Diploid	--	--
NH-APP[a]	*APP*	*NH5*						
		NH5-II.70	AC	43	3	Gain	Chr21: 23 984 747–27 466 529	3482
		NH5-II.38	AC	48	3	Gain	Chr21: 23 987 177–27 466 529	3479
		NH5-II.49	ND	--	3	Gain	Chr21: 23 987 177–27 466 008	3479
		NH5-II.50	AC	50	3	Gain	Chr21: 23 993 039–27 458 052	3465
		NH5-II.51	ND	--	2	Diploid	--	--

Abbreviations: AD, Alzheimer's disease; CN, copy number; CNV, copy number variation; EO-FAD, early-onset familial Alzheimer's disease; NCRAD, National Cell Repository for Alzheimer's Disease; NIMH, National Institute of Mental Health Alzheimer's Disease Genetics Initiative Study.
CNV-type and phenotype details of the test families carrying the novel CNVs that show putative pathogenic inheritance. Physical location corresponds to hg18 version of the genomic assembly and the genes listed are encompassed by the CNVs within 1MB proximal genomic region. The diagnosis (Dx) of the families are coded as AC for autopsy confirmed, PR for probable AD, PS for possible AD and ND for no dementia. Array CN indicates the change in genomic copy number corresponding to the gain or loss in CN state. [a]Includes two previously reported *APP* duplication families.[35]

tight co-segregation with AD in these 10 families. All but one (FAD-CNV7) of the 10 CNVs are exclusively observed in affected subjects in these families, whereas absent in the rest of the affected and unaffected individuals in our AD families ($N = 561$ families; $N = 1750$ subjects). In addition, all 10 CNVs are absent in the publicly available CNV database on the DGV.

Evidence for possible genetic heterogeneity was observed in three of the families. In one family, NH1, FAD-CNV7 was observed in a single unaffected subject (NH1-II.94) for whom the last diagnosis was made in 2006 with no information on current age or latest disease status. In another family, NH4, one subject (NH4-II.80) was listed as 'probable dementia' but did not carry FAD-CNV10, a duplication of *MAPT*. It should be noted, however, that family NH4 likely involves genetic heterogeneity (similar to that observed in the control family ND7—APP duplication). FAD-CNV10 co-segregates with AD in all three early-onset patients (onset <62 years), but not in the sole late-onset family member (onset $=71$ years). Finally, family ND5 may also represent a case of genetic heterogeneity since one subject (ND5-II.73) did not carry FAD-CNV5 while two EO-FAD patients with onset <65 years did carry. The non-carrier is listed as 'possible dementia' with no onset age available; unfortunately, no other information is available on this subject. Collectively, our findings support the probable pathogenicity of these CNVs as novel forms of genomic mutation leading to EO-FAD.

The 10 CNVs identified in this study affect genes implicated in a variety of neuronal function pathways and some have been previously been implicated in other neurodegenerative disorders. FAD-CNV1 was detected in two affected siblings (onset ages: 40 and 42 years, both *APOE-ε4* negative), with no unaffected siblings, from family ND1. This 124-kb CN loss segment, overlaps *CIP150* (*KIAA1432)* and *ERMP1*. *CIP150* has been reported to be essential for phosphorylation and localization of Cx43,[37] one of the most ubiquitous gap junction protein. Gap junction proteins are specialized in intercellular connection between various cell types, notably at the synapses and is associated with both neuroprotective[38] and neurodestructive[39] functions. ERMP1 has not been characterized in great detail, but encodes a zinc-dependent endoplasmic reticulum metallopeptidase, with decreased expression in a transgenic AD mouse model;[40] higher *ERMP1* expression has been reported in epilepsy mouse models.[41]

Two affected siblings from family ND2 (onset ages: 52 and 56, *APOE-ε4* negative) with no unaffected siblings, carry FAD-CNV2 (240-kb gain), which overlaps *EVC*, *EVC2* and *CRMP1*. Mutations in *EVC* and *EVC2* cause the bone growth disorder (dwarfism), Ellis–Van Creveld syndrome (OMIM: 225500) and the skeletal disorder, Weyers acrofacial dysostosis (OMIM: 193530), however, gene function is unknown.[42] *CRMP1* belongs to cytosolic phosphoprotein family and is highly expressed in the brain.[43] *CRMP1* has a role in *SEMA3A*- (a gene associated with AD[44]) mediated regulation of neural growth and axonal guidance, and microtubule assembly.[43,45–47] Considering the role of CRMPs in axonal guidance and regeneration in the brain, gain in *CRMP1* could compromise neuroplasticity with aging and thus lead to AD. Interestingly, CRMP2 has been reported to be hyper-phosphorylated as an early event in AD pathogenesis.[48]

FAD-CNV3 (110 kb, CN loss) was observed in two affected siblings (onset age: 56 and 68, *APOE-ε4* negative) with no unaffected siblings. The closest genes in the proximity are *A2BP1* located 0.3-Mb telomeric, and *ABAT*, which is 0.57-Mb centromeric to the CNV breakpoint. Studies suggest CNVs can alter expression levels of genes more than to 2 Mb away;[16,49] thus, both genes could possibly undergo CNV-mediated dysregulation. *A2BP1* controls splice variants in many genes—mainly neurodevelopmental, and loss in gene function is associated with various conditions, including: mental retardation, epilepsy, autism spectrum disorder and obesity.[9,50–52] *ABAT* is involved in catalyzing neurotransmitter GABA (gamma-aminobutyric acid), and associated with various neuronal disorders.[53,54]

Two of three affected siblings (onset ages 64 and 65, *APOE-ε4* positive) carry FAD-CNV5, while the third affected sibling (unknown onset age) did not. The 488-kb gain in CN encompasses *CDH2*, encoding *N*-cadherin, which is expressed in brain, skeletal and cardiac muscles and has critical roles in synaptic adhesion, dendritic morphology and neuritic growth.[55–59] *CDH2* has been extensively studied in AD. For example, inhibition of *N*-cadherin function has been reported to accelerate Aβ-triggered synapse damage.[60] Moreover, dissociation of *N*-cadherin-mediated synaptic contact by Aβ has been proposed to cause neuronal cell death, synaptic loss and tau phosphorylation in AD brain.[61]

Two affected siblings in family ND6 show the presence of FAD-CNV6, a 0.73-Mb intergenic gain in CN. *EPHA6* is located 0.33-Mb telomeric and the only described gene in close proximity. *EPHA6* is highly expressed in brain and involved in forming neural networks,[62] and has been shown to cause learning and memory impairment in knock-out mice,[63] suggesting that *EPHA6* dysregulation as a possible pathogenic pathway leading to AD. The 400-kb deletion (FAD-CNV7) shows partial overlap (59%) with CNPs in DGV, and overlaps *KANK1*, *DMRT1* and *DMRT3*. Locus 9p24.3 has been reported in numerous disorders including, chromosome 9p deletion syndrome[64] (OMIM: 158170), cerebral palsy (OMIM: 607704), obsessive compulsive disorder[65] suggesting a role for these loci with neuronal function and various disorders.

Two affected siblings in family NH2, onset age 45 and 61, contain a 331-kb deletion (FAD-CNV8) overlapping *CHMP2B* and *POU1F1*, while two unaffected siblings were negative for the CNV. Both affected individuals are homozygous for the *APOE-ε4* allele, while the two unaffected siblings were heterozygous for *APOE-ε4* allele with current ages of 75 and 83. *CHMP2B* has been previously shown to harbor rare pathogenic mutations leading to frontotemporal dementia (OMIM: 609512) and amyotrophic lateral sclerosis (ALS, OMIM: 600795). *CHMP2B* is expressed in all neuronal populations and colocalizes with granulovacuolar degeneration—one of the pathological hallmarks in AD.[66,67] The partially deleted adjacent gene, *POU1F1* (OMIM: 173110), has been reported to cause pituitary hormone deficiency and is associated with mental retardation.[68,69] *CHMP2B* has also been implicated in endosomal trafficking, the disruption of which could lead to neurodegeneration.[70]

FAD-CNV9, a 151-kb CN loss, was detected in two affected subjects (AAO: 62, 75 and *APOE-ε4* positive) in family NH3 overlapping *FLJ35024* and *VLDLR*. *VLDLR* has been previously studied as a candidate gene in AD, albeit without evidence for strong association (www.alzgene.org).[71] *VLDLR* has been reported to be involved in multiple AD-related pathways, including acting as a receptor for APOE as well as roles in Tau phosphorylation and synaptic functioning via interaction with *RELN*.[72,73] Numerous pathogenic mutations in *VLDLR* have also been reported to cause cerebral ataxia and mental retardation (OMIM: 224050).

Finally, we also observed three affected siblings in family NH4 to carry a 170-kb duplication spanning *MAPT* (FAD-CNV10). *MAPT* encodes the microtubule-associated Tau protein, which in its hyper-phosphorylated and paired helical filament state creates neurofibrillary tangles in AD and tauopathies.[1] Tau is ubiquitous in neuronal axons and crucial to microtubule assembly. Pathogenic mutations in *MAPT*, including CNVs, have been previously implicated in frontotemporal dementia and 17q21.31 microduplication with variable phenotype.[74] Rovelet-Lecrux *et al.*[75] previously described a 439-kb microduplication at the 17q21.31 locus encompassing the *MAPT*, *IMP5*, *CRHR1* and *STH* genes in the proband of a family in which three patients displayed an frontotemporal lobar dementia phenotype. In our study, only *MAPT* was duplicated in all three affected siblings. Thus, this is the first evidence for duplication of solely *MAPT* leading to a dementia phenotype. It remains to be determined (pending autopsy confirmation) whether family NH4 is affected with AD, frontotemporal lobar dementia or a disorder that encompasses both forms of these dementias.

In summary, we have systematically analyzed EO-FAD and early/mixed-onset AD families for CNVs putatively conferring pathogenicity. We have identified a dozen rare, novel and large CNV regions that co-segregate with disease status within families. The genes implicated by these CNVs have roles in wide range of

neuronal pathways critical to normal functioning of the brain. Further studies will be required to elucidate the precise pathogenic mechanisms underlying the co-segregation of these CNVs with AD in our family data sets.

Supplemental data description

Supplementary Table 1 shows the results of our screening for known mutations in *APP*, *PSEN1*, *PSEN2*, *MAPT* and *GRN*. The table lists the pathogenic mutations identified in early-onset AD pedigrees in the three known genes in AD and frontotemporal dementia.

Supplementary Figure 1 shows the fluorescent *in-situ* hybridization (FISH) images of lymphoblastoid cell lines derived from probands and the corresponding ΔCt values from Fluidigm CNV assay. Supplementary Figure 2 shows the pedigree structure of the families found to carry CNVs that show pathogenic form of inheritance.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This study was funded by grants from the NIMH and the Cure Alzheimer's Fund.

REFERENCES

1 Tanzi RE, Bertram L. Twenty years of the Alzheimer's disease amyloid hypothesis: A genetic perspective. *Cell* 2005; **120**: 545–555.
2 Gatz M, Reynolds CA, Fratiglioni L, Johansson B, Mortimer JA, Berg S *et al.* Role of genes and environments for explaining Alzheimer disease. *Arch Gen Psychiatry* 2006; **63**: 168–174.
3 Hooli BV, Tanzi RE. A current view of Alzheimer's disease. *F1000 Biol Rep* 2009 1.
4 Bertram L, Lill CM, Tanzi RE. The genetics of Alzheimer disease: back to the future. *Neuron* 2010; **68**: 270–281.
5 Bertram L, Tanzi RE. The genetics of Alzheimer's disease. *Prog Mol Biol Transl Sci* 2012; **107**: 79–100.
6 Feuk L, Carson AR, Scherer SW. Structural variation in the human genome. *Nat Rev Genet* 2006; **7**: 85–97.
7 Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, Thorne N *et al.* Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science* 2007; **315**: 848–853.
8 Varki A, Geschwind DH, Eichler EE. Explaining human uniqueness: genome interactions with environment, behaviour and culture. *Nat Rev Genet* 2008; **9**: 749–763.
9 Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T *et al.* Strong association of de novo copy number mutations with autism. *Science* 2007; **316**: 445–449.
10 Levy D, Ronemus M, Yamrom B, Lee YH, Leotta A, Kendall J *et al.* Rare *de novo* and transmitted copy-number variation in autistic spectrum disorders. *Neuron* 2011; **70**: 886–897.
11 Sanders SJ, Ercan-Sencicek AG, Hus V, Luo R, Murtha MT, Moreno-De-Luca D *et al.* Multiple recurrent *de novo* CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron* 2011; **70**: 863–885.
12 Vacic V, McCarthy S, Malhotra D, Murray F, Chou HH, Peoples A *et al.* Duplications of the neuropeptide receptor gene VIPR2 confer significant risk for schizophrenia. *Nature* 2011; **471**: 499–503.
13 Bassett AS, Scherer SW, Brzustowicz LM. Copy number variations in schizophrenia: Critical review and new perspectives on concepts of genetics and disease. *Am J Psychiatry* 2010; **167**: 899–914.
14 Gonzalez E, Kulkarni H, Bolivar H, Mangano A, Sanchez R, Catano G *et al.* The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science* 2005; **307**: 1434–1440.
15 Townson JR, Barcellos LF, Nibbs RJ. Gene copy number regulates the production of the human chemokine CCL3-L1. *Eur J Immunol* 2002; **32**: 3016–3026.
16 Lee C, Scherer SW. The clinical context of copy number variation in the human genome. *Expert Rev Mol Med* 2010; **12**: e8.
17 Girirajan S, Eichler EE. Phenotypic variability and genetic susceptibility to genomic disorders. *Hum Mol Genet* 2010; **19**: R176–R187.
18 Ionita-Laza I, Rogers AJ, Lange C, Raby BA, Lee C. Genetic association analysis of copy-number variation (CNV) in human disease pathogenesis. *Genomics* 2009; **93**: 22–26.
19 Pagnamenta AT, Holt R, Yusuf M, Pinto D, Wing K, Betancur C *et al.* A family with autism and rare copy number variants disrupting the Duchenne/Becker muscular dystrophy gene DMD and TRPM3. *J Neurodev Disord* 2011; **3**: 124–131.
20 Chartier-Harlin MC, Kachergus J, Roumier C, Mouroux V, Douay X, Lincoln S *et al.* Alpha-synuclein locus duplication as a cause of familial Parkinson's disease. *Lancet* 2004; **364**: 1167–1169.
21 Rovelet-Lecrux A, Hannequin D, Raux G, Le Meur N, Laquerriere A, Vital A *et al.* APP locus duplication causes autosomal dominant early-onset Alzheimer disease with cerebral amyloid angiopathy. *Nat Genet* 2006; **38**: 24–26.
22 Bertram L, Schjeide BM, Hooli B, Mullin K, Hiltunen M, Soininen H *et al.* No association between CALHM1 and Alzheimer's disease risk. *Cell* 2008; **135**: 993–994, author reply 4-6.
23 Ku CS, Pawitan Y, Sim X, Ong RT, Seielstad M, Lee EJ *et al.* Genomic copy number variations in three Southeast Asian populations. *Hum Mutat* 2010; **31**: 851–857.
24 de Andrade M, Atkinson EJ, Bamlet WR, Matsumoto ME, Maharjan S, Slager SL *et al.* Evaluating the influence of quality control decisions and software algorithms on SNP calling for the affymetrix 6.0 SNP array platform. *Hum Hered* 2011; **71**: 221–233.
25 Wang K, Li M, Hadley D, Liu R, Glessner J, Grant SF *et al.* PennCNV: An integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res* 2007; **17**: 1665–1674.
26 Database of Genomic Variants [database on the Internet] 2011, Available from http://projects.tcag.ca/variation/project.html.
27 Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y *et al.* Detection of large-scale variation in the human genome. *Nat Genet* 2004; **36**: 949–951.
28 Zhang J, Feuk L, Duggan GE, Khaja R, Scherer SW. Development of bioinformatics resources for display and analysis of copy number and other structural variants in the human genome. *Cytogenet Genome Res* 2006; **115**: 205–214.
29 Maiti S, Kumar KH, Castellani CA, O'Reilly R, Singh SM. Ontogenetic de novo copy number variations (CNVs) as a source of genetic individuality: Studies on two families with MZD twins for schizophrenia. *PLoS One* 2011; **6**: e17125.
30 UCSC Genome Bioinformatics: FAQ [database on the Internet] 2012, Available from http://genome.ucsc.edu/FAQ/FAQformat.html#format1.
31 Qin J, Jones RC, Ramakrishnan R. Studying copy number variations using a nanofluidic platform. *Nucleic Acids Res* 2008; **36**: e116.
32 Mohapatra G, Moore DH, Kim DH, Grewal L, Hyun WC, Waldman FM *et al.* Analyses of brain tumor cell lines confirm a simple model of relationships among fluorescence in situ hybridization, DNA index, and comparative genomic hybridization. *Genes Chromosomes Cancer* 1997; **20**: 311–319.
33 Pang AW, MacDonald JR, Pinto D, Wei J, Rafiq MA, Conrad DF *et al.* Towards a comprehensive structural variation map of an individual human genome. *Genome Biol* 2010; **11**: R52.
34 Hooli BV, Mohapatra G, Mattheisen M, Parrado AR, Roehr JT, Shen Y *et al.* Role of common and rare APP DNA sequence variants in Alzheimer disease. *Neurology* 2012; **78**: 1250–1257.
35 Alzheimer Disease & Frontotemporal Dementia Mutation Database [database on the Internet] 1998, cited 21 March 2011.
36 Kauwe JS, Jacquart S, Chakraverty S, Wang J, Mayo K, Fagan AM *et al.* Extreme cerebrospinal fluid amyloid beta levels identify family with late-onset Alzheimer's disease presenilin 1 mutation. *Ann Neurol* 2007; **61**: 446–453.
37 Akiyama M, Ishida N, Ogawa T, Yogo K, Takeya T. Molecular cloning and functional analysis of a novel Cx43 partner protein CIP150. *Biochem Biophys Res Commun* 2005; **335**: 1264–1271.
38 Farahani R, Pina-Benabou MH, Kyrozis A, Siddiq A, Barradas PC, Chiu FC *et al.* Alterations in metabolism and gap junction expression may determine the role of astrocytes as "good samaritans" or executioners. *Glia* 2005; **50**: 351–361.
39 Perez Velazquez JL, Frantseva MV, Naus CC. Gap junctions and neuronal injury: Protectants or executioners? *Neuroscientist* 2003; **9**: 5–9.
40 Martin B, Brenneman R, Becker KG, Gucek M, Cole RN, Maudsley S. iTRAQ analysis of complex proteome alterations in 3xTgAD Alzheimer's mice: understanding the interface between physiology and disease. *PLoS One* 2008; **3**: e2750.
41 Bergren SK, Rutter ED, Kearney JA. Fine mapping of an epilepsy modifier gene on mouse Chromosome 19. *Mamm Genome* 2009; **20**: 359–366.
42 Sund KL, Roelker S, Ramachandran V, Durbin L, Benson DW. Analysis of Ellis van Creveld syndrome gene products: implications for cardiovascular development and disease. *Hum Mol Genet* 2009; **18**: 1813–1824.
43 Schmidt EF, Strittmatter SM. The CRMP family of proteins and their role in Sema3A signaling. *Adv Exp Med Biol* 2007; **600**: 1–11.
44 Good PF, Alapat D, Hsu A, Chu C, Perl D, Wen X *et al.* A role for semaphorin 3A signaling in the degeneration of hippocampal neurons during Alzheimer's disease. *J Neurochem* 2004; **91**: 716–736.
45 Yamashita N, Uchida Y, Ohshima T, Hirai S, Nakamura F, Taniguchi M *et al.* Collapsin response mediator protein 1 mediates reelin signaling in cortical neuronal migration. *J Neurosci* 2006; **26**: 13357–13362.
46 Mukherjee J, DeSouza LV, Micallef J, Karim Z, Croul S, Siu KW *et al.* Loss of collapsin response mediator Protein1, as detected by iTRAQ analysis, promotes

invasion of human gliomas expressing mutant EGFRvIII. *Cancer Res* 2009; **69**: 8545–8554.
47 Kurnellas MP, Li H, Jain MR, Giraud SN, Nicot AB, Ratnayake A *et al.* Reduced expression of plasma membrane calcium ATPase 2 and collapsin response mediator protein 1 promotes death of spinal cord neurons. *Cell Death Differ* 2010; **17**: 1501–1510.
48 Cole AR, Noble W, van Aalten L, Plattner F, Meimaridou R, Hogan D *et al.* Collapsin response mediator protein-2 hyperphosphorylation is an early event in Alzheimer's disease progression. *J Neurochem* 2007; **103**: 1132–1144.
49 Weterman MA, van Ruissen F, de Wissel M, Bordewijk L, Samijn JP, van der Pol WL *et al.* Copy number variation upstream of PMP22 in Charcot-Marie-Tooth disease. *Eur J Hum Genet* 2010; **18**: 421–428.
50 Hammock EA, Levitt P. Developmental expression mapping of a gene implicated in multiple neurodevelopmental disorders, a2bp1 (fox1). *Dev Neurosci* 2011; **33**: 64–74.
51 Bhalla K, Phillips HA, Crawford J, McKenzie OL, Mulley JC, Eyre H *et al.* The de novo chromosome 16 translocations of two patients with abnormal phenotypes (mental retardation and epilepsy) disrupt the A2BP1 gene. *J Hum Genet* 2004; **49**: 308–311.
52 Martin CL, Duvall JA, Ilkin Y, Simon JS, Arreaza MG, Wilkes K *et al.* Cytogenetic and molecular characterization of A2BP1/FOX1 as a candidate gene for autism. *Am J Med Genet B Neuropsychiatr Genet* 2007; **144B**: 869–876.
53 Barnby G, Abbott A, Sykes N, Morris A, Weeks DE, Mott R *et al.* Candidate-gene screening and association analysis at the autism-susceptibility locus on chromosome 16p: evidence of association at GRIN2A and ABAT. *Am J Hum Genet* 2005; **76**: 950–966.
54 Wegerer M, Adena S, Pfennig A, Czamara D, Sailer U, Bettecken T *et al.* Variants within the GABA transaminase (ABAT) gene region are associated with somatosensory evoked EEG potentials in families at high risk for affective disorders. *Psychol Med* 2013; **43**: 1207–1217.
55 Aiga M, Levinson JN, Bamji SX. N-cadherin and neuroligins cooperate to regulate synapse formation in hippocampal cultures. *J Biol Chem* 2011; **286**: 851–858.
56 Lefort CT, Wojciechowski K, Hocking DC. N-cadherin cell-cell adhesion complexes are regulated by fibronectin matrix assembly. *J Biol Chem* 2011; **286**: 3149–3160.
57 Malinverno M, Carta M, Epis R, Marcello E, Verpelli C, Cattabeni F *et al.* Synaptic localization and activity of ADAM10 regulate excitatory synapses through N-cadherin cleavage. *J Neurosci* 2010; **30**: 16343–16355.
58 Rieger S, Senghaas N, Walch A, Koster RW. Cadherin-2 controls directional chain migration of cerebellar granule neurons. *PLoS Biol* 2009; **7**: e1000240.
59 Tan ZJ, Peng Y, Song HL, Zheng JJ, Yu X. N-cadherin-dependent neuron-neuron interaction is required for the maintenance of activity-induced dendrite growth. *Proc Natl Acad Sci USA* 2010; **107**: 9873–9878.
60 Andreyeva A, Nieweg K, Horstmann K, Klapper S, Muller-Schiffmann A, Korth C *et al.* C-terminal fragment of N-cadherin accelerates synapse destabilization by amyloid-beta. *Brain* 2012; **135**(Pt 7): 2140–2154.
61 Ando K, Uemura K, Kuzuya A, Maesako M, Asada-Utsugi M, Kubota M *et al.* N-cadherin regulates p38 MAPK signaling via association with JNK-associated leucine zipper protein: implications for neurodegeneration in Alzheimer disease. *J Biol Chem* 2011; **286**: 7619–7628.
62 Orioli D, Klein R. The Eph receptor family: axonal guidance by contact repulsion. *Trends Genet* 1997; **13**: 354–359.
63 Savelieva KV, Rajan I, Baker KB, Vogel P, Jarman W, Allen M *et al.* Learning and memory impairment in Eph receptor A6 knockout mice. *Neurosci Lett* 2008; **438**: 205–209.
64 Hayashi S, Imoto I, Aizu Y, Okamoto N, Mizuno S, Kurosawa K *et al.* Clinical application of array-based comparative genomic hybridization by two-stage screening for 536 patients with mental retardation and multiple congenital anomalies. *J Hum Genet* 2011; **56**: 110–124.
65 Willour VL, Yao Shugart Y, Samuels J, Grados M, Cullen B, Bienvenu 3rd OJ *et al.* Replication study supports evidence for linkage to 9p24 in obsessive-compulsive disorder. *Am J Hum Genet* 2004; **75**: 508–513.
66 Funk KE, Mrak RE, Kuret J. Granulovacuolar degeneration (GVD) bodies of Alzheimer's disease (AD) resemble late-stage autophagic organelles. *Neuropathol Appl Neurobiol* 2011; **37**: 295–306.
67 Yamazaki Y, Takahashi T, Hiji M, Kurashige T, Izumi Y, Yamawaki T *et al.* Immunopositivity for ESCRT-III subunit CHMP2B in granulovacuolar degeneration of neurons in the Alzheimer's disease hippocampus. *Neurosci Lett* 2010; **477**: 86–90.
68 Sun Y, Zhang F, Gao J, Gao X, Guo T, Zhang K *et al.* Positive association between POU1F1 and mental retardation in young females in the Chinese Han population. *Hum Mol Genet* 2006; **15**: 1237–1243.
69 Turton JP, Reynaud R, Mehta A, Torpiano J, Saveanu A, Woods KS *et al.* Novel mutations within the POU1F1 gene associated with variable combined pituitary hormone deficiency. *J Clin Endocrinol Metab* 2005; **90**: 4762–4770.
70 Urwin H, Authier A, Nielsen JE, Metcalf D, Powell C, Froud K *et al.* Disruption of endocytic trafficking in frontotemporal dementia with CHMP2B mutations. *Hum Mol Genet* 2010; **19**: 2228–2238.
71 Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE. Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nat Genet* 2007; **39**: 17–23.
72 Carter CJ. Convergence of genes implicated in Alzheimer's disease on the cerebral cholesterol shuttle: APP, cholesterol, lipoproteins, and atherosclerosis. *Neurochem Int* 2007; **50**: 12–38.
73 Forster E, Bock HH, Herz J, Chai X, Frotscher M, Zhao S. Emerging topics in Reelin function. *Eur J Neurosci* 2010; **31**: 1511–1518.
74 Rovelet-Lecrux A, Campion D. Copy number variations involving the microtubule-associated protein tau in human diseases. *Biochem Soc Trans* 2012; **40**: 672–676.
75 Rovelet-Lecrux A, Hannequin D, Guillin O, Legallic S, Jurici S, Wallon D *et al.* Frontotemporal dementia phenotype associated with MAPT gene duplication. *J Alzheimers Dis* 2010; **21**: 897–902.

Supplementary Information accompanies the paper on the Molecular Psychiatry website (http://www.nature.com/mp)