Title

Genome sequencing for early-onset or atypical dementia: high diagnostic yield and frequent observation of multiple contributory alleles

Journal

Molecular Case Studies, 5(6)

ISSN

2373-2865

Authors

Cochran, J Nicholas
McKinley, Emily C
Cochran, Meagan
et al.

Publication Date

2019-12-01

DOI

10.1101/mcs.a003491

Peer reviewed

COLD SPRING HARBOR
Molecular Case Studies

RESEARCH ARTICLE

Genome sequencing for early-onset or atypical dementia: high diagnostic yield and frequent observation of multiple contributory alleles

J. Nicholas Cochran,[1] Emily C. McKinley,[2] Meagan Cochran,[1] Michelle D. Amaral,[1] Bryan A. Moyers,[1] Brittany N. Lasseigne,[1,5] David E. Gray,[1] James M.J. Lawlor,[1] Jeremy W. Prokop,[1,3] Ethan G. Geier,[4] James M. Holt,[1] Michelle L. Thompson,[1] J. Scott Newberry,[1] Jennifer S. Yokoyama,[4] Elizabeth A. Worthey,[1] David S. Geldmacher,[2] Marissa Natelson Love,[2] Gregory M. Cooper,[1] Richard M. Myers,[1,6] and Erik D. Roberson[2,6]

[1]HudsonAlpha Institute for Biotechnology, Huntsville, Alabama 35806, USA; [2]Alzheimer's Disease Center, Department of Neurology, University of Alabama at Birmingham, Birmingham, Alabama 35294, USA; [3]Department of Pediatrics and Human Development, Michigan State University, East Lansing, Michigan 48824, USA; [4]Department of Neurology, Memory and Aging Center, University of California San Francisco, San Francisco, California 94158, USA

Abstract We assessed the results of genome sequencing for early-onset dementia. Participants were selected from a memory disorders clinic. Genome sequencing was performed along with *C9orf72* repeat expansion testing. All returned sequencing results were Sanger-validated. Prior clinical diagnoses included Alzheimer's disease, frontotemporal dementia, and unspecified dementia. The mean age of onset was 54 (41–76). Fifty percent of patients had a strong family history, 37.5% had some, and 12.5% had no known family history. Nine of 32 patients (28%) had a variant defined as pathogenic or likely pathogenic (P/LP) by American College of Medical Genetics and Genomics standards, including variants in *APP*, *C9orf72*, *CSF1R*, and *MAPT*. Nine patients (including three with P/LP variants) harbored established risk alleles with moderate penetrance (odds ratios of ~2–5) in *ABCA7*, *AKAP9*, *GBA*, *PLD3*, *SORL1*, and *TREM2*. All six patients harboring these moderate penetrance variants but not P/LP variants also had one or two *APOE* ε4 alleles. One patient had two *APOE* ε4 alleles with no other established contributors. In total, 16 patients (50%) harbored one or more genetic variants likely to explain symptoms. We identified variants of uncertain significance (VUSs) in *ABI3*, *ADAM10*, *ARSA*, *GRID2IP*, *MME*, *NOTCH3*, *PLCD1*, *PSEN1*, *TM2D3*, *TNK1*, *TTC3*, and *VPS13C*, also often along with other variants. In summary, genome sequencing for early-onset dementia frequently identified multiple established or possible contributory alleles. These observations add support for an oligogenic model for early-onset dementia.

[Supplemental material is available for this article.]

Corresponding author:
ncochran@hudsonalpha.org

Ontology terms: Alzheimer disease; frontotemporal dementia

Published by Cold Spring Harbor Laboratory Press

doi:10.1101/mcs.a003491

[5]Current address: Department of Cell, Developmental and Integrative Biology, University of Alabama at Birmingham, Birmingham, Alabama 35294, USA
[6]These authors contributed equally to this work.

INTRODUCTION

Genomic technologies are increasingly being used in clinical settings, but clinical large-scale sequencing for adult-onset neurological conditions has not been heavily applied. Possible reasons include the use of disease-specific gene panels and uncertain genetic yield, despite promising signals for yield using comprehensive approaches (Blauwendraat et al. 2018). We sought to assess the diagnostic yield and burden of variants implicated in neurodegenerative disease with genome sequencing and *C9orf72* expansion testing in cases of early-onset dementia.

Patients were selected from the Memory Disorders Clinic at the University of Alabama at Birmingham (UAB). Inclusion criteria were clinician-diagnosed early-onset dementia. When possible, unaffected parents were included as participants to allow filtering for de novo variants in patients without a family history (a fruitful approach in pediatric genetic disorders [Vissers et al. 2010; Bowling et al. 2017] and amyotrophic lateral sclerosis [ALS] [Chesi et al. 2013; Steinberg et al. 2015]). In addition, unaffected siblings past the age of onset of the patient were enrolled as participants when possible for variant filtering and segregation.

Before starting the analysis, we set criteria for return of results to patients. First, we used the American College of Medical Genetics and Genomics (ACMG) criteria for pathogenicity (Richards et al. 2015) to identify highly penetrant causal variation. For moderately penetrant variants, we set criteria to return (i) *APOE* ε4 status for early-onset Alzheimer's disease (EOAD), (ii) any variant with a disease-associated odds ratio >2 in multiple reports as an "established risk variant," or (iii) one strong report with a disease-associated odds ratio >2 with replication included in the study design as a "likely risk variant."

RESULTS

Clinical Presentation and Family History

Prior clinical diagnoses for patients included EOAD, frontotemporal dementia (FTD), and other unspecified dementias. Twenty-one patients were female and 11 were male. Twenty-eight self-reported Caucasian, four self-reported African–American, and all reported non-Hispanic ethnicity. The mean age of onset was 54 (range 41–76). Ten patients had ages of onset in their 40s, 17 in their 50s, four in their 60s, and one in his 70s. Two of the patients with onset in their 60s had an age of onset below age 65 (a typically used threshold for early-onset dementia [Lambert et al. 2014]). The three patients with onset after 65 were included based on clinician discretion. The patient with onset in the 70s had uncertainty in the reported onset date and an unusual leukoenceophalopathy presentation (in which *NOTCH3* variants of unknown significance [VUSs] were identified), and all three patients with ages of onset after 65 had strong family history (modified Goldman score of 1 or 1.5 as defined below).

In addition to enrolling patients, we also enrolled reportedly unaffected family members for variant filtering and segregation analyses. Thirty-one unaffected relatives were enrolled, 29 of whom had genome sequencing (two were only checked for variants by Sanger). Only two families had complete trios (mother, father, and proband) to allow for searching for de novo variants, of which none of interest was identified. In total, 20 unaffected siblings, nine unaffected parents, and two unaffected cousins were enrolled.

A strong family history of dementia was reported for 50% of patients (16/32), whereas 37.5% (12/32) had some family history, and 12.5% (4/32) had no reported family history. Our definition of family history is based on a modification of a four-point scoring system first put forward by Goldman et al. (2005) in which we modified the score as follows: (1) At least

three people in two generations affected with EOAD, FTD, or ALS, with one person being a first-degree relative of the other two; (1.5) same as (1) but with LOAD (late-onset Alzheimer's disease) instead of EOAD; (2) at least three relatives with dementia or ALS but in whom criteria for autosomal dominant inheritance were not met; (3) a single affected first- or second-degree family member with early-onset dementia or ALS; (3.5) a single affected first- or second-degree family member with late-onset dementia or ALS; or (4) no contributory family history or unknown family history. We considered a score of 1 or 1.5 as strong family history, a score of 2, 3, or 3.5 as some family history, and a score of 4 as no reported family history. All family history information is listed alongside phenotype and variant information in Supplemental Table 1.

To protect patient information, more detailed diagnoses and phenotype information beyond that provided here and listed in Supplemental Table 1 are only provided in the controlled access data set, NIAGADS project NG00082, to qualified researchers approved for access.

Genomic Analyses

Nine of 32 (28%) patients had a highly penetrant variant relevant to their clinical diagnosis (ACMG P/LP [Richards et al. 2015]), whereas seven (22%) had multiple moderately penetrant risk alleles (Fig. 1). Individual cases are discussed next, with variants identified summarized by Table 1 and listed alongside phenotype information in Supplemental Table 1. Note that, in general, the variants identified as either high confidence or possible contributors to disease are rare (median allele frequency in Table 1 is 2.4 per 100,000) and are predicted damaging by Combined Annotation Dependent Depletion (CADD) score (Kircher et al. 2014) (median in Table 1 is a Phred-scaled CADD score of 29.5 or approximately the top 0.1% most predicted damaging variants in the genome).

Pathogenic or Likely Pathogenic Diagnoses

Variants were first evaluated using ACMG criteria for pathogenicity, and all P/LP variants were returned to patients (Richards et al. 2015). We provide a summary below, with detail on the ACMG evidence codes for variants provided in the Supplemental ACMG Pathogenicity Evidence Details.

An APP Pathogenic Variant in Two Siblings. Two siblings with ages of onset in the mid-to-late 40s and a family history of EOAD suggestive of dominant inheritance harbored a pathogenic variant in *APP* (NM_000484.3:c.2149G > T, p.(Val717Phe)), a well-established

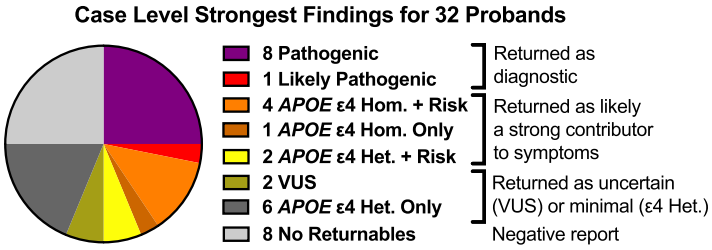

Figure 1. Summary of genomic analysis results for 32 patients with early-onset or familial dementia. Pathogenic variants were observed in *APP* (x2), *C9orf72* (x3), and *MAPT* (x3). A likely pathogenic variant was observed in *CSF1R*. Five patients were *APOE* ε4 homozygous, with four of these patients also harboring additional risk variants in *AKAP9*, *GBA*, *PLD3*, and *TREM2*. Two patients were *APOE* ε4 heterozygous and had additional risk variants in *SORL1* and *TREM2*. Two patients had variants of uncertain significance (VUS) in *MAPT* and *NOTCH3*. For six patients, the only returnable finding was *APOE* ε4 heterozygosity. Eight patients had no returnable findings.

Table 1. Variant table

Gene	Chr.	HGVS DNA	HGVS protein	Variant type	Predicted effect	dbSNP ID	gnomAD alleles per 100,000	CADD score
APP	21	NM_000484.3:c.2149G > T	p.(Val717Phe)	SNV	Missense	rs63750264	0	35
C9orf72	9	NM_001256054.1:c.-45 + 163_-45 + 168GGGGCC [(24_?)]	NA	Insertion	Repeat expansion	rs143561967	0	NA
ABCA7	19	NM_019112.3:c.5035G > T	p.(Glu1679*)	SNV	Stop gained	rs770510230	1.2	37
APOE	19	NM_000041.3:c.388T > C	p.(Cys130Arg)	SNV	Missense	rs429358	14254	0.007
APOE	19	NM_000041.3:c.526C > T	p.(Arg176Cys)	SNV	Missense	rs7412	6538	30
PSEN1	14	NM_000021.3:c.103C > T	p.(Arg35Trp)	SNV	Missense	rs746691776	5.7	26
ABCA7	19	NM_019112.3: c.2126_2132delAGCAGGG	p.(Glu709fs*86)	Deletion	Frameshift	rs547447016	148	35
ARSA	22	NM_000487.5:c.256C > T	p.(Arg86Trp)	SNV	Missense	rs199476352	0	34
ARSA	22	NM_000487.5:c.585G > T	p.(Trp195Cys)	SNV	Missense	rs6151415	5323	27
MAPT	17	NM_005910.5:c.1216C > T	p.(Arg406Trp)	SNV	Missense	rs63750424	1.6	35
APP	21	NM_000484.3:c.1090C > T	p.(Leu364Phe)	SNV	Missense	rs749453173	1.4	21.7
GRID2IP	7	NM_001145118.1: c.429 + 2T > G	NA	SNV	Splice	rs1413118387	0	22.4
CSF1R	5	NM_005211.3:c.2699G > A	p.(Arg900Lys)	SNV	Missense	NA (private)	0	33
PLD3	19	NM_012268.3:c.694G > A	p.(Val232Met)	SNV	Missense	rs145999145	325	29.9
APP	21	NM_000484.3:c.742G > A	p.(Asp248Asn)	SNV	Missense	rs200103591	15	24
ABI3	17	NM_016428.2:c.290T > A	p.(Val97Glu)	SNV	Missense	NA (private)	0	33
SORL1	11	NM_003105.5:c.314T > C	p.(Met105Thr)	SNV	Missense	rs982581946	3.2	23.8
TREM2	6	NM_018965.3:c.140G > A	p.(Arg47His)	SNV	Missense	rs75932628	248	33
TREM2	6	NM_018965.3:c.259G > A	p.(Asp87Asn)	SNV	Missense	rs142232675	115	22.8
AKAP9	7	NM_005751.4:c.7638A > G	p.(Ile2546Met)	SNV	Missense	rs144662445	82	0.009
GBA	1	NM_000157.3:c.1448T > C	p.(Leu483Pro)	SNV	Missense	rs421016	131	24.8
VPS13C	15	NM_020821.2:c.10954C > T	p.(Arg3652*)	SNV	Stop gained	rs138846118	18	49
VPS13C	15	NM_020821.2:c.1988delC	p.(Thr663fs*2)	Deletion	Frameshift	rs1019238429	0	35
PLCD1	3	NM_006225.3:c.631C > T	p.(Arg211Trp)	SNV	Missense	rs752156828	0.4	30
NOTCH3	19	NM_000435.2:c.133G > C	p.(Asp45His)	SNV	Missense	rs142031490	0.8	27.6
NOTCH3	19	NM_000435.2:c.154G > A	p.(Gly52Arg)	SNV	Missense	rs148166997	1.2	29.5
MAPT	17	NM_005910.5:c.1174A > G	p.(Ile392Val)	SNV	Missense	rs991713081	0	24.6
ADAM10	15	NM_001110.3:c.359T > C	p.(Ile120Thr)	SNV	Missense	rs144890810	4.3	14.35
TTC3	21	NM_001320703.1:c.5677G > A	p.(Val1893Met)	SNV	Missense	NA (private)	0	14.62
SORL1	11	NM_003105.5:c.1247G > A	p.(Arg416Gln)	SNV	Missense	rs377550239	3.9	34
MME	3	NM_007289.2:c.1241A > G	p.(Tyr414Cys)	SNV	Missense	rs202095767	0.4	27.7
TM2D3	15	NM_078474.2:c.206C > T	p.(Pro69Leu)	SNV	Missense	rs140152371	3.7	33
TNK1	17	NM_001251902.1:c.393C > G	p.(His131Gln)	SNV	Missense	rs767381816	0	24.5
KCNQ1	11	NM_000218.2:c.1552C > T	p.(Arg518*)	SNV	Stop Gained	rs17215500	11	39

Note that many individuals had multiple candidate contributory variants, which is not captured when considering variants individually. For an expanded table that indicates multiple candidate variants, see Supplemental Table 1. All variants were observed in the heterozygous state except for *APOE* (NM_000041.3:c.388T > C, p.(Cys130Arg)), which was observed in both the heterozygous and homozygous state. *APOE* (NM_000041.3:c.526C > T, p.(Arg176Cys)) is also noted here, because confirmation of its absence along with *APOE* (NM_000041.3:c.388T > C, p.(Cys130Arg)) indicates the *APOE* ε4 allele.
(HGVS) Human Genome Variation Society, (dbSNP) Single Nucleotide Polymorphism Database, (gnomAD) Genome Aggregation Database, (CADD) Combined Annotation Dependent Depletion, (SNV) single-nucleotide variant.

pathogenic variant (see Supplemental ACMG Pathogenicity Evidence Details). This variant is an example of one that would have been identified on commonly used panels for genetic testing for EOAD.

C9orf72 Expansion Carriers. Testing for a pathogenic G_4C_2 hexanucleotide expansion at the *C9orf72* locus associated with ALS and FTD was ordered for 30 of 32 patients (with two excluded for technical reasons, see Methods). GeneDx conducted a repeat-primed PCR test with 95% sensitivity and 98% specificity (Akimoto et al. 2014) to detect *C9orf72* expansions. As a technical aside, *C9orf72* expansions were not detectable using ExpansionHunter (Dolzhenko et al. 2017) or STRetch (Dashnow et al. 2018) in genome sequencing libraries prepared with polymerase chain reaction (PCR) amplification assessed here. ExpansionHunter detects *C9orf72* expansions in PCR-free genome preparations (Dolzhenko et al. 2017), so PCR-free genome preparations or secondary testing (such as testing conducted by GeneDx here) is necessary for detection of *C9orf72* expansions (and would also be necessary for other repeat expansions). Three patients with FTD (one of whom also had ALS) with ages of onset in the 40s and 50s harbored a pathogenic expansion in *C9orf72* (see Supplemental ACMG Pathogenicity Evidence Details).

Some studies have suggested that additional contributing alleles could lower age of onset and/or alter clinical presentation for *C9orf72* expansion carriers (van Blitterswijk et al. 2012, 2014; Pottier et al. 2015; Giannoccaro et al. 2017; Farhan et al. 2018). Consistent with this, all three *C9orf72* expansion carriers harbored other possibly contributory variants.

One carrier had three additional variants that may be contributory: an "established risk" stop gained variant in *ABCA7* (NM_019112.3:c.5035G > T, p.(Glu1679*)), one *APOE* ε4 allele, and a VUS in *PSEN1* (NM_000021.3:c.103C > T, p.(Arg35Trp)) (see Supplemental ACMG Pathogenicity Evidence Details). These variants may have contributed to the patient's family history of multiple neurodegenerative diseases including ALS and EOAD.

Another carrier had a different "established risk" variant in *ABCA7* (NM_019112.3:c.2126_2132delAGCAGGG, p.(Glu709fs*86)) (see Supplemental ACMG Pathogenicity Evidence Details), along with memory symptoms and a family history of AD, consistent with a possible contributory role of *ABCA7*.

The third carrier had two VUSs in *ARSA*, associated with recessive metachromatic leukodystrophy (discussed further in Supplemental ACMG Pathogenicity Evidence Details).

A MAPT Pathogenic Variant in Three Alzheimer's Disease Patients. Three patients with EOAD (one patient also exhibited FTD signs) with ages of onset in the mid-50s to early 60s harbored a pathogenic variant in *MAPT* (NM_005910.5:c.1216C > T, p.(Arg406Trp)). Although *MAPT* pathogenic variants are typically associated with FTD (Cruts et al. 2012), this variant has been reported in patients with clinically diagnosed Alzheimer's disease (AD) in multiple studies (see Supplemental ACMG Pathogenicity Evidence Details). This variant would not have been detected on many AD-specific panels, which often test for only *APP*, *PSEN1*, and *PSEN2*.

All three of these patients exhibited a possible contribution from another allele, just as in *C9orf72* expansion carriers. One patient had a loss-of-function "established risk" variant in *ABCA7* (NM_019112.3:c.2126_2132delAGCAGGG, p.(Glu709fs*86)). Another patient had a VUS in *APP* (NM_000484.3:c.1090C > T, p.(Leu364Phe)). The third patient had a loss-of-function splice variant in *GRID2IP* (NM_001145118.1:c.429 + 2T > G), which, although not yet firmly associated with EOAD and thus not yet returnable, was implicated in a recent large sequencing study (Raghavan et al. 2018).

The presence of this rare variant in three individuals enrolled at the same clinic suggests they may share a common ancestor. However, none of these individuals is aware of any extended family members participating in the study. Furthermore, the patients are not detectably related by the software used for routine checks of close familial relationships (KING).

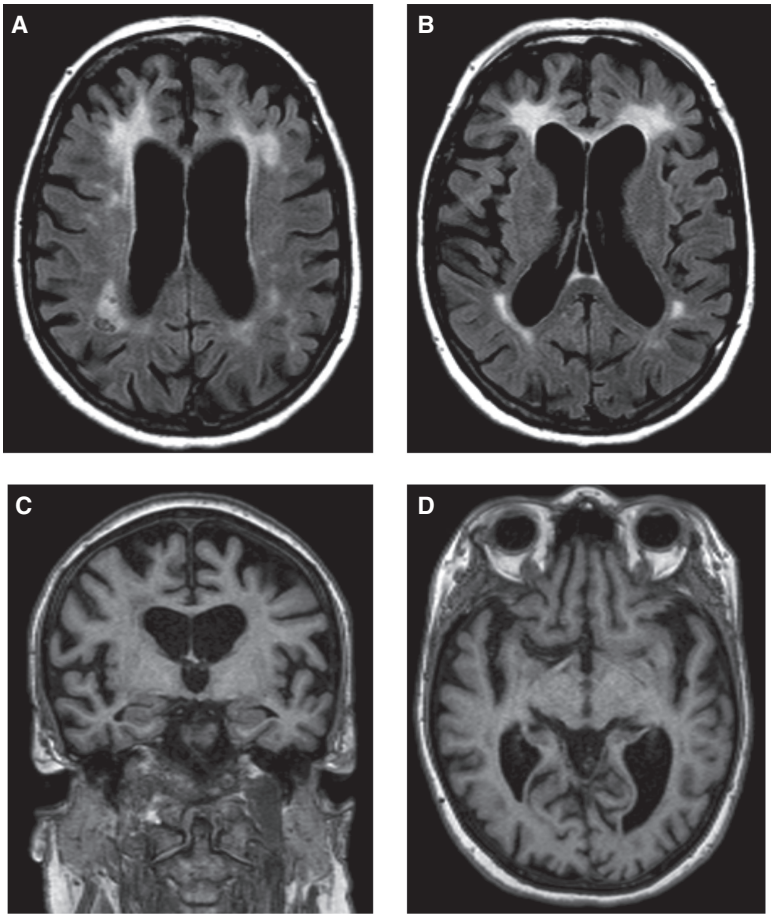

Figure 2. Neuroimaging findings in a *CSF1R* variant carrier. (*A*,*B*) Frontal-predominant, mildly asymmetric (R > L) white matter hyperintensities on axial FLAIR images. (*C*,*D*) Global cerebral atrophy on coronal and axial MPRAGE images. Radiological orientation with patient's R side displayed on L.

A CSF1R Variant in an FTD Patient. A patient presenting with behavioral variant FTD (bvFTD) harbored a likely pathogenic variant in *CSF1R* (NM_005211.3:c.2699G > A, p.(Arg900Lys)) (see Supplemental ACMG Pathogenicity Evidence Details). Patients with variants in *CSF1R* can present with bvFTD, but the underlying pathology of pathogenic *CSF1R* variants is leukoencephalopathy (Rademakers et al. 2011; Stabile et al. 2016). Consistent with this, this patient had white matter abnormalities on MRI, with frontal-predominant confluent white matter hyperintensity (Fig. 2A) and global atrophy (Fig. 2B–D). This variant would not have been detected on typical panels testing for FTD.

High-Impact Risk Alleles

One unique aspect of this study is that we returned to patients moderately penetrant risk variants that meet criteria we have described. Intriguingly, rare variants meeting these criteria were observed only in patients who also carried one or two *APOE* ε4 alleles, the most common moderately penetrant risk allele for AD (see Supplemental ACMG Pathogenicity Evidence Details). In all cases, *APOE* ε4 alleles were returned as "established risk variants." The presence of one *APOE* ε4 allele was returned as likely only a small contributor to symptoms, whereas the presence of two *APOE* ε4 alleles or one or two *APOE* ε4 alleles in

combination with a rare moderately penetrant risk variant was returned with language indicating that such a combination of variants is likely to explain a large portion of the genetic contribution to symptoms (but with the caveat that family members should not be presymptomatically tested given incomplete penetrance). We continue with detail on some cases falling into this category.

A Case with APOE *ε4 Homozygosity along with Variants in* PLD3, APP, *and* ABI3. In a patient with EOAD whose symptoms began in the late 40s with enrolled unaffected parents, we observed an example of how EOAD may occur from a combination of inherited alleles from each parent, consistent with previous observations that EOAD can appear recessive in nature (Wingo et al. 2012). The patient had two *APOE* ε4 alleles (returned as "established risk,") a *PLD3* variant (NM_012268.3:c.694G > A, p.(Val232Met)) (returned as "likely risk,"), an *APP* variant (NM_000484.3:c.742G > A, p.(Asp248Asn)) (returned as a VUS), and a private variant in *ABI3* (NM_016428.2:c.290T > A, p.(Val97Glu)) (not returned but predicted damaging by PolyPhen-2 [Adzhubei et al. 2010] and SIFT [Ng and Henikoff 2003], with a CADD score of 33) (see Supplemental ACMG Pathogenicity Evidence Details). The *ABI3* variant was not returned to the patient because of insufficient evidence to consider the variant as a returnable VUS or risk variant, but is highlighted because a different coding variant in *ABI3* (NM_012268.3:c.1124T > C, p.(Ser209Phe)) (Sims et al. 2017) was associated with AD in a rigorous case-control study with an odds ratio of 1.4, yet is not predicted to be as damaging (CADD = 13.5) and is relatively common in population databases (allele frequency of 0.6%). Therefore, we speculate that perhaps the variant we observed could have an effect of similar or greater magnitude given its higher predicted deleteriousness and absence from population databases. One of the *APOE* ε4 alleles with the variants in *PLD3* and *APP* was inherited from a parent with neurologic symptoms but not EOAD. The other parent harbored an *APOE* ε4 allele and the *ABI3* variant and did not have neurologic symptoms. This case serves as an example of how EOAD may arise with either no family history or limited family history of late-onset disease.

A Case with APOE *ε4 Heterozygosity and a* SORL1 *Variant.* An individual with EOAD with onset in the mid 50s and a strong family history of AD had one *APOE* ε4 allele and a variant in *SORL1* (NM_003105.5:c.314T > C, p.(Met105Thr)). Although *SORL1* variants are not completely penetrant, loss-of-function variants in *SORL1* confer one of the highest levels of risk for AD outside of dominant pathogenic variants and *APOE*. Loss-of-function *SORL1* variant carriers in cases from a recent study (Raghavan et al. 2018) are present at an odds ratio of ~4 compared to population databases, a likely underestimate given that some individuals in population databases may develop AD. Indeed, a recent meta-analysis suggests the odds ratio for loss-of-function *SORL1* variants could be as high as 12.3 for all AD and 27.5 for EOAD (Campion et al. 2019).

For the *SORL1* variant identified here, we checked independent data sets for replication, and observed one M105T carrier in one study (Sassi et al. 2016), three M105T carriers in Alzheimer's Disease Sequencing Project (ADSP) exomes (Bis et al. 2018), and two M105T carriers in ADSP genomes (one in an AD case and in one a mild cognitive impairment case) with no controls harboring the variant in any of these data sets. No other carriers were identified in cases or controls in four other studies (see Supplemental ACMG Pathogenicity Evidence Details). In addition to these four studies, there is one record in ClinVar from GeneDx (RCV000489328.1), but it lacked a denominator of the number of cases tested and thus was not considered in calculating the replication statistic. Taken together, *SORL1* M105T is observed six times out of 13,390 AD cases compared to 11 of 189,196 individuals at a population level for a replication-only odds ratio of 7.7 ($P = 0.0005$ by Fisher's exact test). This variant did not completely segregate with disease in four family members of our patient. However, the age-of-onset range for similar variants in *SORL1* can be up to 24 yr

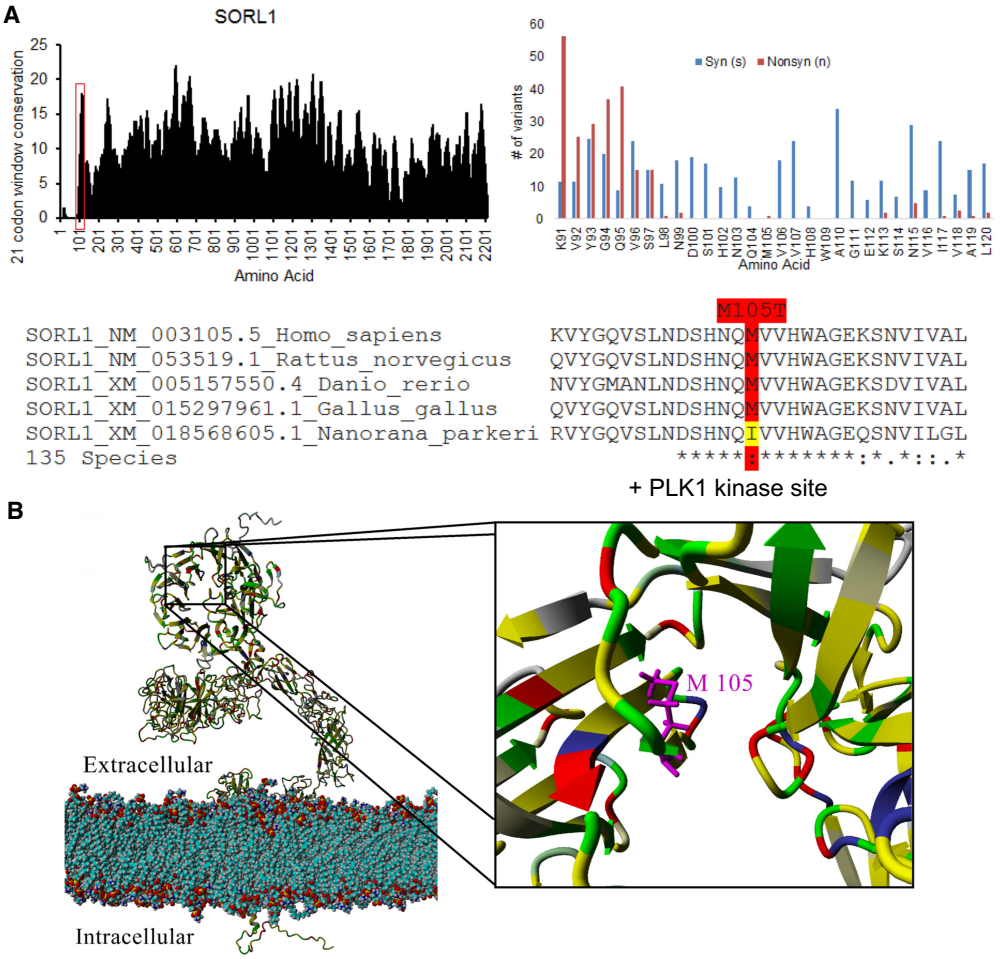

Figure 3. Molecular modeling of the effect of the M105T variant on SORL1. (*A*) Conservation analysis of the *SORL1* gene sequence was performed across open reading frame sequences of 135 species. Scores at each codon were assessed with 100% conservation receiving a score of 1, with addition of a score for codon selection (score of 0 if dN-dS of site is below mean, 0.25 for sites with values above the mean to one standard deviation above the mean, 0.5 for sites greater than one standard deviation but below two standard deviations, 1 for sites greater than two standard deviations). A score of 2 is maximal, suggesting an amino acid that is 100% conserved with codon wobble indicative of a high selection rate at the position. The values were then placed on a 21-codon sliding window (combining values for 10 codons before and after each position) to identify conserved motifs within the gene. (*B*) Model of SORL1 protein (assessed with YASARA2). Colors are based on 135 species alignments fed into ConSurf such that colors indicate (gray) not conserved, (yellow) conserved hydrophobic, (red) conserved polar acidic, (blue) conserved polar basic, (green) conserved hydrophilic. Note that the M105T variant leads to a predicted gain of a PLK1 kinase target site in SORL1.

(Louwersheimer et al. 2017), which is wider than the age differences between the family members we genotyped, suggesting that this segregation analysis may not be completely informative. Considering all of the evidence, we returned this variant to the patient as a VUS (it could also be considered a "likely risk variant"). Modeling suggests M105T is a highly conserved residue (Fig. 3A) in which change to a threonine may create a PLK1 kinase site that may disrupt function (Fig. 3B) (discussed further in Supplemental ACMG Pathogenicity Evidence Details).

APOE ε4 *with* TREM2, AKAP9, *and* GBA *Risk Variants.* In two cases with EOAD beginning in the late 40s, we observed a risk allele in *TREM2* and one or two *APOE* ε4 alleles. The first was *TREM2* (NM_018965.3:c.140G > A, p.(Arg47His)) (Guerreiro et al. 2013; Jonsson et al. 2013) with one *APOE* ε4 allele. This *TREM2* variant was returned as an "established risk variant." Second, we observed *TREM2* (NM_018965.3:c.259G > A, p.(Asp87Asn)) (Guerreiro et al. 2013) (see Supplemental ACMG Pathogenicity Evidence Details) with two *APOE* ε4 alleles. This *TREM2* variant was returned as a "likely risk variant."

In an African–American patient with features of both EOAD and FTD, we observed two *APOE* ε4 alleles along with a variant in *AKAP9* previously reported to increase risk in African–Americans (NM_005751.4:c.7638A > G, p.(Ile2546Met)) (Logue et al. 2014). In this case, despite being observed in only one study with replication, the specificity of this variant disease association to African–American ethnicity and additional functional data (Ikezu et al. 2018) provided enough evidence to return this as an "established risk variant."

A patient with EOAD with onset in the mid-50s harbored *GBA* (NM_000157.3:c.1448T > C, p.(Leu483Pro) [previous nomenclature, p.(Leu444Pro)]) and two *APOE ε4* alleles, originally associated with Lewy body disorders (Mata et al. 2008), but later also with mixed dementia with Lewy bodies and AD (Tsuang et al. 2012; Nalls et al. 2013). Because of this and a recent association with accelerated cognitive decline (Liu et al. 2016), we returned this as a "likely risk variant."

VPS13C Predicted Loss of Function with APOE *ε4.* A patient with mixed symptoms of AD and FTD with onset in the late 60s harbored *VPS13C* (NM_020821.2:c.10954C > T, p.(Arg3652*)) and two *APOE* ε4 alleles. A patient with EOAD with onset in the late 40s had *VPS13C* (NM_020821.2:c.1988delC, p.(Thr663fs*2)), a variant in *PLCD1* (NM_006225.3:c.631C > T, p.(Arg211Trp)), and one *APOE* ε4 allele. Only *APOE* ε4 was reported back to these patients because of uncertain contribution of the other variants to the phenotype. Homozygous loss of *VPS13C* is associated with early-onset Parkinson's (Schormair et al. 2018). We do not know the significance of the observation of one loss-of-function allele here, although unpublished studies have reported an association between heterozygous loss of function in *VPS13C* and FTD (see Supplemental ACMG Pathogenicity Evidence Details). *PLCD1* was proposed as a candidate gene for AD in one study (Shimohama et al. 1998). Observing two loss-of-function variants in *VPS13C* in this small cohort leads us to speculate that heterozygous loss-of-function variants in *VPS13C* may contribute to early-onset dementia.

Variants of Uncertain Significance or Research Interest

Five other patients harbored interesting—but speculative—VUSs or combinations of variants of interest for future research. These include (1) a patient with possible CADASIL and a haplotype of uncertain significance with two variants in *NOTCH3* (NM_000435.2:c.133G > C, p.(Asp45His) and NM_000435.2:c.154G > A, p.(Gly52Arg)), (2) a patient with a VUS in *MAPT* (NM_005910.5:c.1174A > G, p.(Ile392Val)), (3) a patient with an *APOE* ε4 allele and a variant in both *ADAM10* (NM_001110.3:c.359T > C, p.(Ile120Thr)) and *TTC3* (NM_001001894.2:c.5557G > A, p.(Val1893Met)), (4) a patient with an *APOE* ε4 allele, and a variant in both *SORL1* (NM_003105.5:c.1247G > A, p.(Arg416Gln)) and *MME* (NM_007289.2:c.1241A > G, p.(Tyr414Cys)), and (5) a patient with variants in *TM2D3* (NM_078474.2:c.206C > T, p.(Pro69Leu)) and *TNK1* (NM_001251902.1:c.393C > G, p.(His131Gln)). Furthermore, one patient harbored a secondary pathogenic variant in *KCNQ1* (NM_000218.2:c.1552C > T, p.(Arg518*)). We expand on these cases in the Supplemental ACMG Pathogenicity Evidence Details.

Quantitative Enrichment of Multiple Alleles

Because we observed so many cases harboring multiple established alleles, we asked if this effect was statistically enriched over a control population recruited from the same

geographical area, with controls reporting a family history of dementia excluded. These controls are healthy unaffected parents from a childhood disease study in which de novo mutations are the most common cause of disease (Bowling et al. 2017), making these parents reasonably representative population controls. Genome sequencing data from these controls was produced in the same manner (depth, sequencing technology, processing pipelines, etc.) as the primary data for this study. We set criteria for qualifying variants as follows: (1) *TREM2* or *GBA* missense or loss-of-function variants with CADD > 20 and population frequency <0.5% in both gnomAD (Lek et al. 2016) and TOPMed Bravo (NHLBI 2018), (2) *ABCA7*, *SORL1*, *TBK1*, or *GRN* loss-of-function variants with CADD > 20 and population frequency <0.5%, (3) the specific *PLD3* and *AKAP9* variants observed here because their associations are for single alleles, (4) missense-only variants with CADD > 20 and population frequency <0.01% for *SORL1*, *CSF1R*, *APP*, *PSEN1*, *PSEN2*, and *MAPT*, (5) expansion carriers in *C9orf72*, and (6) *APOE* ε4 alleles. We recognize that this may contain bias because these filtering criteria were selected after analysis of cases. However, we attempted to mitigate this by selecting reasonable thresholds that would catch variants not identified in this study but that would still have been considered if they had been identified. For example, we did not observe any variants meeting these criteria in *TBK1* or *GRN* but included them here because of their important role in disease. We also included *C9orf72* carriers without information on if any are present in the control population, but this is a reasonable assumption (see Supplemental ACMG Pathogenicity Evidence Details).

Variants meeting the criteria described are highly enriched in cases, whereas unaffected family members are intermediate between cases and population controls (Fig. 4A). Intriguingly, there is no enrichment of *APOE* ε4 alleles in the absence of other qualifying alleles in this cohort for cases versus population controls, although the distribution of *APOE* ε4 alleles is detectably different between population controls and unaffected family members (Fig. 4B). In contrast, the presence of *APOE* ε4 alleles in combination with another qualifying variant is highly enriched in cases (and also enriched in unaffected family members), regardless of whether Mendelian variants are included in the calculation (Fig. 4C) or excluded (Fig. 4D). The odds ratios for *APOE* ε4 alleles in combination with another qualifying variant in cases without a Mendelian cause versus population controls suggests that the presence of rare variants increases odds ratios approximately multiplicatively over those typically reported for *APOE* ε4 alone (typically reported: ~2.5 for one *APOE* ε4 allele, with a rare variant, 5.5; 10–15 for two *APOE* ε4 alleles, with a rare variant, 39.1), see Supplemental ACMG Pathogenicity Evidence Details on *APOE*) (Fig. 4D).

Because we observed a clear enrichment of multiple qualifying alleles in cases versus controls, we asked if multiple qualifying alleles were associated with a lower age of onset. Age at onset did not significantly correlate with number of implicated alleles (numbers based on qualifying variants used for Fig. 4), but trends were in the direction of lower age of onset with more implicated alleles (Spearman's $r = -0.2189$, $P = 0.2288$ for all alleles [Supplemental Fig. 1A], Spearman's $r = -0.3664$, $P = 0.0856$ when excluding cases with Mendelian contributors [Supplemental Fig. 1B]). Because the small number of samples in this study may lead to a false-negative observation, we analyzed data from the Alzheimer's Disease Sequencing Project (2208 cases), and age at enrollment did significantly correlate with implicated alleles based on the same filtering criteria (Spearman's $r = -0.2072$, $P < 0.0001$ [Supplemental Fig. 1C]).

DISCUSSION

One key theme in this study was the frequent observation of multiple possible contributory alleles. We even observed this in multiple cases with clear, highly penetrant, pathogenic

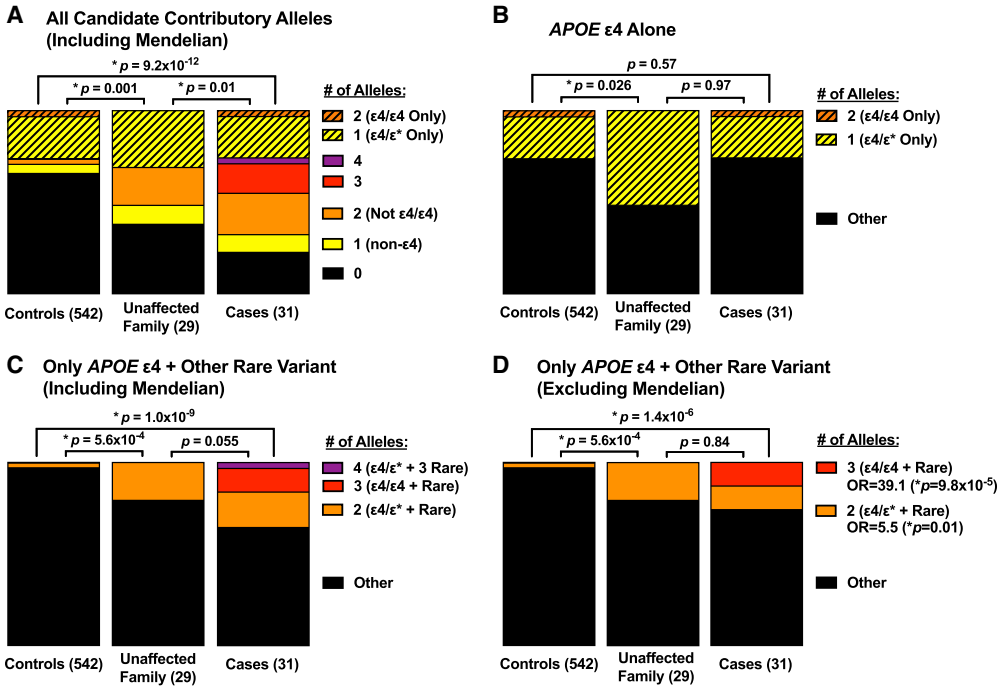

Figure 4. Multiple variants in neurodegeneration-associated genes are often observed in early-onset dementia, with a critical role for rare variants acting in combination with *APOE* ε4. Note that for all panels, ε4/ε* indicates either ε4/ε3 or ε4/ε2 (mostly ε4/ε3). Also for all panels, cases $N = 31$ (32 probands excluding 1 sibling from an affected sibling pair), controls $N = 542$, and unaffected family members of cases $N = 29$. All comparisons are by exact conditional Cochran–Armitage trend test unless otherwise specified. (*A*) Qualifying candidate alleles associated with neurodegeneration (see text for criteria) are highly enriched in cases ($P = 9.2 \times 10^{-12}$). Unaffected family members are intermediate between cases and controls ($P = 0.01$ vs. cases, $P = 0.001$ vs. controls). (*B*) Presence of *APOE* ε4 alone, in the absence of any other qualifying variants, is not enriched in cases ($P = 0.57$). Unaffected family members show a significantly different *APOE* ε4 allele distribution from controls ($P = 0.026$) but not cases ($P = 0.97$). (*C*) Presence of *APOE* ε4 along with at least one qualifying rare variant (including Mendelian variants) is highly enriched in cases ($P = 1.0 \times 10^{-9}$). Enrichment over controls is also observed in unaffected family members ($P = 5.6 \times 10^{-4}$), but unaffected family members are not statistically distinguishable from cases ($P = 0.055$). (*D*) Presence of *APOE* ε4 along with at least one qualifying rare variant (excluding Mendelian variants) is highly enriched in cases ($P = 1.4 \times 10^{-6}$) and also enriched in unaffected family members ($P = 5.6 \times 10^{-4}$), but unaffected family members are not distinguishable from cases ($P = 0.84$). The odds ratio for presence of one *APOE* ε4 allele along with one qualifying rare variant in cases versus controls is 5.5 ($P = 0.01$ by Fisher's exact test, 95% CI 1.2–19.3). The odds ratio for presence of two *APOE* ε4 alleles along with one qualifying rare variant in cases versus controls is 39.1 ($P = 9.8 \times 10^{-5}$ by Fisher's exact test, 95% CI 5.3–447.5).

variants despite a small cohort size. The degree to which additional alleles contribute in dominant cases cannot be assessed without larger cohorts to evaluate effects on age of onset or other variables. However, given that other studies have made similar observations in ALS/FTD (van Blitterswijk et al. 2012, 2014; Pottier et al. 2015; Giannoccaro et al. 2017; Farhan et al. 2018), this phenomenon clearly warrants further investigation.

In cases in which a dominant pathogenic variant was not present, there was enrichment for multiple established alleles outside of *APOE* ε4 homozygosity alone contributing to disease risk. Although the odds ratios for this observation are compelling (5.5 for an established rare risk variant along with one *APOE* ε4 allele and 39.1 for an established rare risk variant along with two *APOE* ε4 alleles; Fig. 4D) we recognize that the small sample size in this study limits their interpretability. Future studies with larger sample sizes, and ideally future studies

with a prospective design, are required to more precisely determine the effect of presence of multiple established risk alleles on lifetime risk for dementia.

The observation that unaffected family members exhibited either an intermediate number of possibly contributory variants or an enriched number compared to population controls, but not distinguishable from cases, is expected for moderate-penetrance variants and highlights why we strongly emphasize not pursuing testing for unaffected family members when we return risk variant information. Possible explanations for why unaffected family members do not exhibit segregation include known incomplete penetrance of risk alleles, other genetic contributors not currently annotated as conferring risk, environmental influences, and variable age of onset (i.e., some currently unaffected family members may go on to develop disease in the future, as age-of-onset spreads with risk alleles can be as high as 24 yr [Louwersheimer et al. 2017]).

Every case with a moderately penetrant risk variant established by case-control studies identified in this cohort also harbored one or two *APOE* ε4 alleles, emphasizing the importance of *APOE* ε4. In addition, although detectable differences were minimal between cases and unaffected family members, no unaffected family members harbored two *APOE* ε4 alleles along with a qualifying rare variant. That observation, along with the high odds ratio versus population controls (39.1; Fig. 4D), suggests a well-established rare variant contributor along with two *APOE* ε4 alleles confers a high level of risk, and that individuals harboring such a combination of variants should be closely monitored.

Future efforts in the analysis of large cohorts should include analysis of level of risk when rare risk variants are present—for example, by incorporation of signal from rare variation in established risk genes into polygenic risk scores. Several groups have begun developing polygenic risk scores for AD (Escott-Price et al. 2015; Desikan et al. 2017), but these scores are based solely on common variation. This is, of course, a reasonable approach because it maximizes reproducibility, as considering rare variants could lead to an overtrained model. However, although rare variants are rare individually, aggregation approaches may provide replicable and meaningful signals if incorporated for key genes in which rare variants are now established to confer risk for AD, such as *ABCA7*, *SORL1*, and *TREM2*. Similarly, although large FTD genetic studies are not as progressed as those for AD, we can begin to consider genes in which variation in a polygenic risk score may be informative for FTD, such as *TBK1* (Cirulli et al. 2015), *MFSD8* (Geier et al. 2019), *DPP6*, *UNC13A*, and *HLA-DQA2* (Pottier et al. 2019).

This study adds support for an oligogenic model in cases of early-onset dementia, emphasizing that comprehensive genetic approaches are valuable for research studies and could hold value clinically in the future. Mendelian diagnostic yield in this population was 28%, with an additional 22% of patients harboring risk-increasing variants that, in combination with *APOE* ε4, likely account for most of the genetic contribution to their symptoms. Genome sequencing is able to identify relevant variation in conditions with high genetic heterogeneity, nonspecific phenotypes, or established risk factors that do not follow a clear Mendelian pattern, and allowed for the identification of cryptic genotype–phenotype relationships that likely would have been missed by panel testing. However, we note that for future studies, substantially more samples could be evaluated for lower cost using arrays with high coverage of neurodegeneration associated variants and/or targeted sequencing of more comprehensive panels with known genes associated with neurodegenerative diseases (both of which are commercially available) if the goal is not to discover new variants, but to more precisely measure the level of risk when multiple established risk variants are present. This lower cost would be helpful for case-control studies but would be especially important in a prospective study design in which the hit rate for carriers of multiple risk variants outside of *APOE* ε4 homozygosity is expected to be low. For example, only 3.0% of controls in this study harbor multiple contributory alleles meeting the criteria we described for inclusion in

Figure 4D outside of *APOE* ε4 homozygosity (which contributes an additional 3.3%). In addition to potential prospective studies that would benefit from rare variant information, efforts are currently ongoing to recruit patients for clinical trials based on *APOE* ε4 status (Langbaum et al. 2019). Incorporation of information about established rare variant risk factor status could be helpful for efforts such as these.

In addition to the research value of this study, it had value for patient care as well—for example, by allowing for referral of families to the Dominantly Inherited Alzheimer's Network and studies such as the ALLFTD study (a natural history study of FTD) if an actionable variant was identified. Furthermore, although return of risk variants to patients is controversial, in this study we qualitatively observed that patients and their caregivers were generally appreciative of receiving findings even if they were not definitive. We acknowledge that our study conditions were ideal for such an observation: The clinicians involved were deeply knowledgeable about genetic contributions to dementia and the role of risk variants, and the patients were informed at study enrollment and at return of results about limitations of risk variant observations if they were identified. Furthermore, participants had access to genetic counseling services through the study to help clarify the potential significance of results. Future studies to quantitatively evaluate clinician, patient, and caregiver responses to risk variant information are needed. We conclude that application of more comprehensive genetic assessment (including genome sequencing, if appropriate) could aid in interpretation of early-onset dementia cases currently and will continue to grow in utility as future studies designed to both more comprehensively define levels of risk based on the presence of multiple risk alleles and also to study clinician, patient, and caregiver responses to such information are conducted.

METHODS

Genome Sequencing

Genome sequencing was performed at the HudsonAlpha Institute for Biotechnology on Illumina HiSeq X or NovaSeq platforms using paired end 150-base pair reads. Mean depth was 38× with an average of 89.0% of bases covered at 20× (a full sequencing coverage table is available in Supplemental Table 2). Sequencing libraries were prepared by Covaris shearing, end repair, adapter ligation, and PCR using standard protocols. Library concentrations were normalized using KAPA qPCR prior to sequencing. All sequencing variants returned to patients were validated by Sanger in a CAP/CLIA laboratory.

Data Processing and Quality Control

Demuxed FASTQs were aligned with bwa-0.7.12 (Li and Durbin 2009) to hg19. BAMs were sorted and duplicates were marked with Sambamba 0.5.4 (Tarasov et al. 2015). Indels were realigned, bases were recalibrated, and gVCFs were generated with GATK 3.3 (McKenna et al. 2010). gVCFs were batch-called with GATK 3.8. KING 2.1.2 (Manichaikul et al. 2010) was used for sex checks on VCFs, for validation of known familial relationships, and to check for unknown familial relationships (none of which was identified).

C9orf72 Expansion Testing

Samples from 30 of 32 patients were tested for pathogenic *C9orf72* repeat expansion alleles by GeneDx. Two patients did not have sufficient material for testing, but both lacked symptoms consistent with a *C9orf72* repeat expansion and also had another likely explanation of symptoms: One had a pathogenic *APP* variant and another harbored both one *APOE* ε4 allele and a *TREM2* established risk allele.

Genomic Data Analysis

The HudsonAlpha-developed Codicem application (http://envisiongenomics.com/codicem-analysis-platform/) was used to analyze and support the interpretation of the variant data (described elsewhere [Holt et al. 2019]). Although this software package was used for analysis, it would not be necessary to use this package to reproduce this work. Simple filtering for population allele frequencies (i.e., gnomAD [Lek et al. 2016] and TOPMed Bravo [NHLBI 2018]), in silico deleteriousness scores (i.e., CADD [Kircher et al. 2014], PolyPhen-2 [Adzhubei et al. 2010], and SIFT [Ng and Henikoff 2003]), and gene lists relevant to the phenotype of interest would recapitulate our findings using any suitable software package or even by a command line interface.

In addition to searching for single-nucleotide variants and small indels, we also searched for large copy-number variations using four callers (DELLY [Rausch et al. 2012], ERDS [Zhu et al. 2012], CNVnator [Abyzov et al. 2011], and BIC-seq2 [Xi et al. 2016]), but did not identify any relevant to patient phenotypes (including absence of APP duplications).

SORL1 Structural Modeling

SORL1 structural modeling and evolutionary conservation analysis was assessed using a previously published sequence-to-structure-to-function workflow (Prokop et al. 2017).

Statistics

The exact conditional Cochran–Armitage trend test was calculated using the CATTexact 0.1.0 package and Fisher's exact test using fisher.test in R 3.4.1. Spearman correlations were calculated with GraphPad Prism 8.2.1.

Return of Results

Results meeting criteria for return were delivered to patients by clinicians in the UAB Memory Disorders Clinic through letters written by a genetic counselor. Letters included information on the variant, associated disease, recurrence risk, and management recommendations. Patients were given the option to have a genetic counselor present for return of results via phone or videoconference or to follow up with a genetic counselor after delivery of results. Primary results were provided only to probands. Although a secondary result was identified in only one participant who was a patient, we did also offer nonpatient participants (family members) receipt of actionable secondary findings (ACMG 59) if such a result had been identified. Family members of patients that received diagnostic results were provided with information to seek out clinical genetic counseling and targeted testing for familial variants if they desired.

ADDITIONAL INFORMATION

Data Deposition and Access

All data from participants enrolled as a part of this study, including more detailed phenotype data for the cases described here, are available on the National Institute on Aging Genetics of Alzheimer's Disease Data Storage (NIAGADS) site (https://www.niagads.org/) under project NG00082. Data from control subjects not enrolled as a part of this study are available under dbGaP (https://www.ncbi.nlm.nih.gov/gap/) accession phs001089.v3.p1, which contains data generated by the Clinical Sequencing Exploratory Research (CSER) Consortium established by the National Human Genome Research Institute (NHGRI). Funding support for CSER was provided through cooperative agreements with the NHGRI and the National Cancer Institute (NCI) through grant numbers U01 HG007301 (Genomic Diagnosis in

Children with Developmental Delay). Information about CSER and the investigators and institutions who comprise the CSER consortium can be found at https://cser-consortium.org. ADNI (Alzheimer's Disease Neuroimaging Initiative, part of the ADSP genomes batch call) and ADSP data are available at NIAGADS under projects NG00066 and NG00067 and on dbGaP under accession phs000572.v7.p4 (see Supplemental Extended Acknowledgements for full list of ADNI and ADSP contributors and funding sources). All variants described here are deposited in ClinVar (http://www.ncbi.nlm.nih.gov/clinvar/) under accession numbers VCV000018089.2, VCV000031151.1, VCV000666796.1, VCV000666796.1, VCV000093125.1, VCV000068126.1, VCV000120224.1, VCV000426729.2, VCV000195350.1, VCV000261064.1, VCV000136339.1, VCV000004288.3, VCV000447782.1, VCV000447788.1, VCV000003131.6, and submission SUB6475528.

Ethics Statement

This study was approved by UAB IRB protocol X161202004, "Evaluation of Genomic Variants in Patients with Neurologic Diseases." All participants described provided explicit written consent for publication.

Acknowledgments

We thank Alissa Mackiewicz from the HudsonAlpha Foundation for assistance in securing funding, Jennifer Mahaffey at UAB for assistance with the IRB application, Mackenzie Fowler at UAB for assistance with participant recruitment, the Clinical Services Laboratory and the Genomic Services Laboratory at HudsonAlpha for DNA isolations, library generation, quality control, and sequencing, the Codicem software development team at HudsonAlpha for genome analysis software, David Bick at HudsonAlpha for helpful discussions about ACMG guidelines, and Dominique Campion at the University of Rouen for correspondence indicating the absence from both cases and controls of the M105T variant in *SORL1* in the data set published in Bellenguez et al. (2017).

Authors' Contributions

J.N.C., G.M.C., R.M.M., and E.D.R. designed the study. J.N.C. and R.M.M. secured funding. J.N.C. and E.D.R. wrote the IRB protocol. E.C.M. coordinated all aspects of patient interaction. J.N.C., M.D.A., B.A.M., and B.N.L. analyzed genomes with input from M.C., E.C.M., and E.D.R. M.D.A. coordinated *C9orf72* testing. J.N.C., D.E.G., J.M.J.L., J.W.P., E.G.G., J.M.H., and J.S.N. conducted other analyses. M.C. wrote clinical letters and provided genetic counseling. M.L.T provided phenotype information for controls. J.S.Y. accessed ADSP and supervised E.G.G. E.A.W. supervised J.M.H., J.S.N., and the software development team. E.D.R., D.S.G. and M.N.L. recruited participants and returned results. G.M.C. supervised D.E.G., J.M.J.L., and M.L.T. J.N.C. wrote the manuscript, with edits by E.C.M., M.C., M.D.A., B.A.M., B.N.L., J.W.P., E.G.G., J.M.H., E.A.W., G.M.C., and E.D.R. All authors approved the final manuscript.

Competing Interest Statement

The authors have declared no competing interest.

Received July 30, 2019; accepted in revised form October 25, 2019.

Funding

Funding for genomes sequenced at HudsonAlpha was generously provided by the Daniel Foundation of Alabama and donors to the HudsonAlpha Foundation Memory and Mobility Fund.

REFERENCES

Abyzov A, Urban AE, Snyder M, Gerstein M. 2011. CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res* **21:** 974–984. doi:10.1101/gr.114876.110

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. 2010. A method and server for predicting damaging missense mutations. *Nat Methods* **7:** 248–249. doi:10.1038/nmeth0410-248

Akimoto C, Volk AE, van Blitterswijk M, Van den Broeck M, Leblond CS, Lumbroso S, Camu W, Neitzel B, Onodera O, van Rheenen W, et al. 2014. A blinded international study on the reliability of genetic testing for GGGGCC-repeat expansions in *C9orf72* reveals marked differences in results among 14 laboratories. *J Med Genet* **51:** 419–424. doi:10.1136/jmedgenet-2014-102360

Bellenguez C, Charbonnier C, Grenier-Boley B, Quenez O, Le Guennec K, Nicolas G, Chauhan G, Wallon D, Rousseau S, Richard AC, et al. 2017. Contribution to Alzheimer's disease risk of rare variants in *TREM2*, *SORL1*, and *ABCA7* in 1779 cases and 1273 controls. *Neurobiol Aging* **59:** 220.e221–220.e229. doi:10.1016/j.neurobiolaging.2017.07.001

Bis JC, Jian X, Kunkle BW, Chen Y, Hamilton-Nelson KL, Bush WS, Salerno WJ, Lancour D, Ma Y, Renton AE, et al. 2018. Whole exome sequencing study identifies novel rare and common Alzheimer's-associated variants involved in immune response and transcriptional regulation. *Mol Psychiatry* doi:10.1038/s41380-018-0112-7

Blauwendraat C, Wilke C, Simón-Sánchez J, Jansen IE, Reifschneider A, Capell A, Haass C, Castillo-Lizardo M, Biskup S, Maetzler W, et al. 2018. The wide genetic landscape of clinical frontotemporal dementia: systematic combined sequencing of 121 consecutive subjects. *Genet Med* **20:** 240–249. doi:10.1038/gim.2017.102

Bowling KM, Thompson ML, Amaral MD, Finnila CR, Hiatt SM, Engel KL, Cochran JN, Brothers KB, East KM, Gray DE, et al. 2017. Genomic diagnosis for children with intellectual disability and/or developmental delay. *Genome Med* **9:** 43. doi:10.1186/s13073-017-0433-1

Campion D, Charbonnier C, Nicolas G. 2019. *SORL1* genetic variants and Alzheimer disease risk: a literature review and meta-analysis of sequencing data. *Acta Neuropathol* **138:** 173–186. doi:10.1007/s00401-019-01991-4.

Chesi A, Staahl BT, Jovičić A, Couthouis J, Fasolino M, Raphael AR, Yamazaki T, Elias L, Polak M, Kelly C, et al. 2013. Exome sequencing to identify de novo mutations in sporadic ALS trios. *Nat Neurosci* **16:** 851–855. doi:10.1038/nn.3412

Cirulli ET, Lasseigne BN, Petrovski S, Sapp PC, Dion PA, Leblond CS, Couthouis J, Lu YF, Wang Q, Krueger BJ, et al. 2015. Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. *Science* **347:** 1436–1441. doi:10.1126/science.aaa3650

Cruts M, Theuns J, Van Broeckhoven C. 2012. Locus-specific mutation databases for neurodegenerative brain diseases. *Hum Mutat* **33:** 1340–1344. doi:10.1002/humu.22117

Dashnow H, Lek M, Phipson B, Halman A, Sadedin S, Lonsdale A, Davis M, Lamont P, Clayton JS, Laing NG, et al. 2018. STRetch: detecting and discovering pathogenic short tandem repeat expansions. *Genome Biol* **19:** 121. doi:10.1186/s13059-018-1505-2

Desikan RS, Fan CC, Wang Y, Schork AJ, Cabral HJ, Cupples LA, Thompson WK, Besser L, Kukull WA, Holland D, et al. 2017. Genetic assessment of age-associated Alzheimer disease risk: development and validation of a polygenic hazard score. *PLoS Med* **14:** e1002258. doi:10.1371/journal.pmed.1002258

Dolzhenko E, van Vugt J, Shaw RJ, Bekritsky MA, van Blitterswijk M, Narzisi G, Ajay SS, Rajan V, Lajoie BR, Johnson NH, et al. 2017. Detection of long repeat expansions from PCR-free whole-genome sequence data. *Genome Res* **27:** 1895–1903. doi:10.1101/gr.225672.117

Escott-Price V, Sims R, Bannister C, Harold D, Vronskaya M, Majounie E, Badarinarayan N, GERAD/PERADES, IGAP consortia, Morgan K, et al. 2015. Common polygenic variation enhances risk prediction for Alzheimer's disease. *Brain* **138:** 3673–3684. doi:10.1093/brain/awv268

Farhan SMK, Gendron TF, Petrucelli L, Hegele RA, Strong MJ. 2018. OPTN p.Met468Arg and ATXN2 intermediate length polyQ extension in families with C9orf72 mediated amyotrophic lateral sclerosis and frontotemporal dementia. *Am J Med Genet B Neuropsychiatr Genet* **177:** 75–85. doi:10.1002/ajmg.b.32606

Geier EG, Bourdenx M, Storm NJ, Cochran JN, Sirkis DW, Hwang JH, Bonham LW, Ramos EM, Diaz A, Van Berlo V, et al. 2019. Rare variants in the neuronal ceroid lipofuscinosis gene *MFSD8* are candidate risk factors for frontotemporal dementia. *Acta Neuropathol* **137:** 71–88. doi:10.1007/s00401-018-1925-9

Giannoccaro MP, Bartoletti-Stella A, Piras S, Pession A, De Massis P, Oppi F, Stanzani-Maserati M, Pasini E, Baiardi S, Avoni P, et al. 2017. Multiple variants in families with amyotrophic lateral sclerosis and frontotemporal dementia related to *C9orf72* repeat expansion: further observations on their oligogenic nature. *J Neurol* **264:** 1426–1433. doi:10.1007/s00415-017-8540-x

Goldman JS, Farmer JM, Wood EM, Johnson JK, Boxer A, Neuhaus J, Lomen-Hoerth C, Wilhelmsen KC, Lee VM, Grossman M, et al. 2005. Comparison of family histories in FTLD subtypes and related tauopathies. *Neurology* **65:** 1817–1819. doi:10.1212/01.wnl.0000187068.92184.63

Guerreiro R, Wojtas A, Bras J, Carrasquillo M, Rogaeva E, Majounie E, Cruchaga C, Sassi C, Kauwe JS, Younkin S, et al. 2013. TREM2 variants in Alzheimer's disease. *N Engl J Med* **368:** 117–127. doi:10.1056/NEJMoa1211851

Holt JM, Wilk B, Birch CL, Brown DM, Gajapathy M, Moss AC, Sosonkina N, Wilk MA, Anderson JA, Harris JM, et al. 2019. VarSight: prioritizing clinically reported variants with binary classification algorithms. *BMC Bioinformatics* **20:** 496. doi:10.1186/s12859-019-3026-8.

Ikezu T, Chen C, DeLeo AM, Zeldich E, Fallin MD, Kanaan NM, Lunetta KL, Abraham CR, Logue MW, Farrer LA. 2018. Tau phosphorylation is impacted by rare *AKAP9* mutations associated with Alzheimer disease in African Americans. *J Neuroimmune Pharmacol* **13:** 254–264. doi:10.1007/s11481-018-9781-x

Jonsson T, Stefansson H, Steinberg S, Jonsdottir I, Jonsson PV, Snaedal J, Bjornsson S, Huttenlocher J, Levey AI, Lah JJ, et al. 2013. Variant of *TREM2* associated with the risk of Alzheimer's disease. *N Engl J Med* **368:** 107–116. doi:10.1056/NEJMoa1211103

Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. 2014. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* **46:** 310–315. doi:10.1038/ng.2892

Lambert MA, Bickel H, Prince M, Fratiglioni L, Von Strauss E, Frydecka D, Kiejna A, Georges J, Reynish EL. 2014. Estimating the burden of early onset dementia; systematic review of disease prevalence. *Eur J Neurol* **21:** 563–569. doi:10.1111/ene.12325

Langbaum JB, Karlawish J, Roberts JS, Wood EM, Bradbury A, High N, Walsh TL, Gordon D, Aggarwal R, Davis P, et al. 2019. GeneMatch: a novel recruitment registry using at-home APOE genotyping to enhance referrals to Alzheimer's prevention studies. *Alzheimers Dement* **15:** 515–524. doi:10.1016/j.jalz.2018.12.007

Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, et al. 2016. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536:** 285–291. doi:10.1038/nature19057

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25:** 1754–1760. doi:10.1093/bioinformatics/btp324

Liu G, Boot B, Locascio JJ, Jansen IE, Winder-Rhodes S, Eberly S, Elbaz A, Brice A, Ravina B, van Hilten JJ, et al. 2016. Specifically neuropathic Gaucher's mutations accelerate cognitive decline in Parkinson's. *Ann Neurol* **80:** 674–685. doi:10.1002/ana.24781

Logue MW, Schu M, Vardarajan BN, Farrell J, Bennett DA, Buxbaum JD, Byrd GS, Ertekin-Taner N, Evans D, Foroud T, et al. 2014. Two rare *AKAP9* variants are associated with Alzheimer's disease in African Americans. *Alzheimers Dement* **10:** 609–618 e611. doi:10.1016/j.jalz.2014.06.010

Louwersheimer E, Cohn-Hokke PE, Pijnenburg YA, Weiss MM, Sistermans EA, Rozemuller AJ, Hulsman M, van Swieten JC, van Duijn CM, Barkhof F, et al. 2017. Rare genetic variant in *SORL1* may increase penetrance of Alzheimer's disease in a family with several generations of APOE-ε4 homozygosity. *J Alzheimers Dis* **56:** 63–74. doi:10.3233/JAD-160091

Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. 2010. Robust relationship inference in genome-wide association studies. *Bioinformatics* **26:** 2867–2873. doi:10.1093/bioinformatics/btq559

Mata IF, Samii A, Schneer SH, Roberts JW, Griffith A, Leis BC, Schellenberg GD, Sidransky E, Bird TD, Leverenz JB, et al. 2008. Glucocerebrosidase gene mutations: a risk factor for Lewy body disorders. *Arch Neurol* **65:** 379–382. doi:10.1001/archneurol.2007.68

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20:** 1297–1303. doi:10.1101/gr.107524.110

Nalls MA, Duran R, Lopez G, Kurzawa-Akanbi M, McKeith IG, Chinnery PF, Morris CM, Theuns J, Crosiers D, Cras P, et al. 2013. A multicenter study of glucocerebrosidase mutations in dementia with Lewy bodies. *JAMA Neurol* **70:** 727–735. doi:10.1001/jamaneurol.2013.1925

Ng PC, Henikoff S. 2003. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res* **31:** 3812–3814. doi:10.1093/nar/gkg509

NHLBI UoMa. 2018. The NHLBI Trans-Omics for Precision Medicine (TOPMed) Whole Genome Sequencing Program. BRAVO variant browser.

Pottier C, Bieniek KF, Finch N, van de Vorst M, Baker M, Perkersen R, Brown P, Ravenscroft T, van Blitterswijk M, Nicholson AM, et al. 2015. Whole-genome sequencing reveals important role for *TBK1* and *OPTN* mutations in frontotemporal lobar degeneration without motor neuron disease. *Acta Neuropathol* **130:** 77–92. doi:10.1007/s00401-015-1436-x

Pottier C, Ren Y, Perkerson RB, Baker M, Jenkins GD, van Blitterswijk M, DeJesus-Hernandez M, van Rooij JGJ, Murray ME, Christopher E, et al. 2019. Genome-wide analyses as part of the international FTLD-TDP whole-genome sequencing consortium reveals novel disease risk factors and increases support for immune dysfunction in FTLD. *Acta Neuropathol* **137:** 879–899. doi:10.1007/s00401-019-01962-9

Prokop JW, Lazar J, Crapitto G, Smith DC, Worthey EA, Jacob HJ. 2017. Molecular modeling in the age of clinical genomics, the enterprise of the next generation. *J Mol Model* **23:** 75. doi:10.1007/s00894-017-3258-3

Rademakers R, Baker M, Nicholson AM, Rutherford NJ, Finch N, Soto-Ortolaza A, Lash J, Wider C, Wojtas A, DeJesus-Hernandez M, et al. 2011. Mutations in the colony stimulating factor 1 receptor (*CSF1R*) gene cause hereditary diffuse leukoencephalopathy with spheroids. *Nat Genet* **44:** 200–205. doi:10.1038/ng.1027

Raghavan NS, Brickman AM, Andrews H, Manly JJ, Schupf N, Lantigua R, Wolock CJ, Kamalakaran S, Petrovski S, Tosto G, et al. 2018. Whole-exome sequencing in 20,197 persons for rare variants in Alzheimer's disease. *Ann Clin Transl Neurol* **5:** 832–842. doi:10.1002/acn3.582

Rausch T, Zichner T, Schlattl A, Stütz AM, Benes V, Korbel JO. 2012. DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics* **28:** i333–i339. doi:10.1093/bioinformatics/bts378

Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, et al. 2015. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* **17:** 405–424. doi:10.1038/gim.2015.30

Sassi C, Ridge PG, Nalls MA, Gibbs R, Ding J, Lupton MK, Troakes C, Lunnon K, Al-Sarraj S, Brown KS, et al. 2016. Influence of coding variability in APP-Aβ metabolism genes in sporadic Alzheimer's disease. *PLoS One* **11:** e0150079. doi:10.1371/journal.pone.0150079

Schormair B, Kemlink D, Mollenhauer B, Fiala O, Machetanz G, Roth J, Berutti R, Strom TM, Haslinger B, Trenkwalder C, et al. 2018. Diagnostic exome sequencing in early-onset Parkinson's disease confirms *VPS13C* as a rare cause of autosomal-recessive Parkinson's disease. *Clin Genet* **93:** 603–612. doi:10.1111/cge.13124

Shimohama S, Kamiya S, Fujii M, Ogawa T, Kanamori M, Kawamata J, Imura T, Taniguchi T, Yagisawa H. 1998. Mutation in the pleckstrin homology domain of the human phospholipase C-δ1 gene is associated with loss of function. *Biochem Biophys Res Commun* **245:** 722–728. doi:10.1006/bbrc.1998.8307

Sims R, van der Lee SJ, Naj AC, Bellenguez C, Badarinarayan N, Jakobsdottir J, Kunkle BW, Boland A, Raybould R, Bis JC, et al. 2017. Rare coding variants in *PLCG2*, *ABI3*, and *TREM2* implicate microglial-mediated innate immunity in Alzheimer's disease. *Nat Genet* **49:** 1373–1384. doi:10.1038/ng.3916

Stabile C, Taglia I, Battisti C, Bianchi S, Federico A. 2016. Hereditary diffuse leukoencephalopathy with axonal spheroids (HDLS): update on molecular genetics. *Neurol Sci* **37:** 1565–1569. doi:10.1007/s10072-016-2634-6

Steinberg KM, Yu B, Koboldt DC, Mardis ER, Pamphlett R. 2015. Exome sequencing of case-unaffected-parents trios reveals recessive and de novo genetic variants in sporadic ALS. *Sci Rep* **5:** 9124. doi:10.1038/srep09124

Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P. 2015. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31:** 2032–2034. doi:10.1093/bioinformatics/btv098

Tsuang D, Leverenz JB, Lopez OL, Hamilton RL, Bennett DA, Schneider JA, Buchman AS, Larson EB, Crane PK, Kaye JA, et al. 2012. GBA mutations increase risk for Lewy body disease with and without Alzheimer disease pathology. *Neurology* **79:** 1944–1950. doi:10.1212/WNL.0b013e3182735e9a

van Blitterswijk M, van Es MA, Hennekam EA, Dooijes D, van Rheenen W, Medic J, Bourque PR, Schelhaas HJ, van der Kooi AJ, de Visser M, et al. 2012. Evidence for an oligogenic basis of amyotrophic lateral sclerosis. *Hum Mol Genet* **21:** 3776–3784. doi:10.1093/hmg/dds199

van Blitterswijk M, Mullen B, Wojtas A, Heckman MG, Diehl NN, Baker MC, DeJesus-Hernandez M, Brown PH, Murray ME, Hsiung GY, et al. 2014. Genetic modifiers in carriers of repeat expansions in the *C9ORF72* gene. *Mol Neurodegener* **9:** 38. doi:10.1186/1750-1326-9-38

Vissers LE, de Ligt J, Gilissen C, Janssen I, Steehouwer M, de Vries P, van Lier B, Arts P, Wieskamp N, del Rosario M, et al. 2010. A de novo paradigm for mental retardation. *Nat Genet* **42:** 1109–1112. doi:10.1038/ng.712

Wingo TS, Lah JJ, Levey AI, Cutler DJ. 2012. Autosomal recessive causes likely in early-onset Alzheimer disease. *Arch Neurol* **69:** 59–64. doi:10.1001/archneurol.2011.221

Xi R, Lee S, Xia Y, Kim TM, Park PJ. 2016. Copy number analysis of whole-genome data using BIC-seq2 and its application to detection of cancer susceptibility variants. *Nucleic Acids Res* **44:** 6274–6286. doi:10.1093/nar/gkw491

Zhu M, Need AC, Han Y, Ge D, Maia JM, Zhu Q, Heinzen EL, Cirulli ET, Pelak K, He M, et al. 2012. Using ERDS to infer copy-number variants in high-coverage genomes. *Am J Hum Genet* **91:** 408–421. doi:10.1016/j.ajhg.2012.07.004